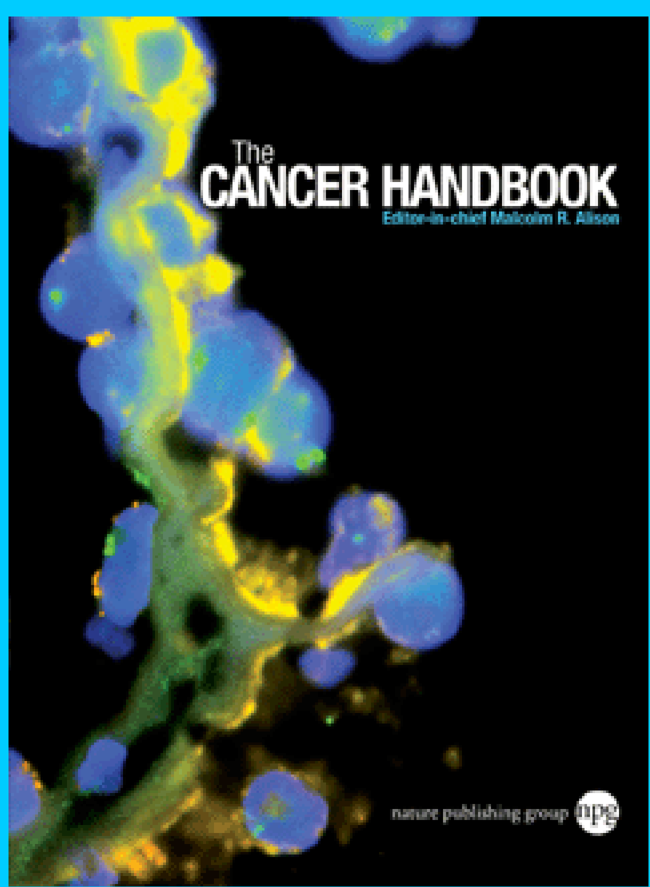
The Cancer Handbook



2003.3.18

huangzhiman

For www.dnathink.org

The Cancer Handbook

Contents

Section A - The Molecular Basis of Cell and Tissue Organisation
Section B - The Causation and Prevention of Cancer
Section C - Diagnostic Imaging and Image-Guided Intervention
Section D - Systemic Oncology
Section E - Pre-clinical Models for Human Cancer
Section F - The Treatment of Human Cancer

Section A - The Molecular Basis of Cell and Tissue Organization

1. Cell and Tissue Organisation
2. Regulation of the Cell Cycle
3. Overview of Oncogenesis
4. Inherited Predispositions to Cancer
5. DNA Viruses
6. RNA Viruses
7. Genetic Instability and DNA repair
8. Telomerase
9. Apoptosis
10. Signalling by steroid receptors
11. Signalling by cytokines
12. Signalling by tyrosine kinases
13. Signalling by TGF beta
14. Wnt Signal Transduction
15. Extracellular Matrix: the networking solution
16. Invasion and Metastasis
17. Angiogenesis
18. Cell Proliferation in Carcinogenesis

back to top

Section B - The Causation and Prevention of Cancer

1. Identifying Cancer Causes through Epidemiology
2. Mechanisms of Chemical Carcinogenesis

3. The Formation of DNA Adducts
4. Physical Causes of Cancer
5. Non-Genotoxic Causes of Cancer
6. Infectious Agents and Cancer
7. Short-term Testing for Genotoxicity
8. Cancer Bioassays for Pharmaceuticals a Regulatory Perspective
9. Molecular Epidemiology of Cancer
10. Dietary Genotoxins and Cancer
11. Tobacco Use and Cancer
12. Occupational Causes of Cancer
13. Anti-Genotoxins and Cancer
14. Intervention and Chemoprevention

back to top

Section C - Diagnostic Imaging and Image-Guided Intervention

Part I- Imaging Modalities in Diagnosis and Monitoring

1. Plain Film Radiography
2. Computed Tomography
3. Ultrasound
4. Magnetic Resonance Imaging
5. Nuclear Medicine Studies
6. Mammography

Part II - Diagnostic and Therapeutic Interventional Procedures

1. Percutaneous Biopsy
2. Transcatheter Therapy
3. Direct Percutaneous Tumour Therapy

back to top

Section D - Systemic Oncology

1. Introduction to the Diagnosis of Cancer
2. Skin
3. Oral cavity & major and minor salivary glands

4. Respiratory tract
5. Upper Gastrointestinal tract
6. Lower gastrointestinal tract
7. Liver, gall bladder and extrahepatic bile ducts.
8. Pancreas
9. Endocrine organs
10. Breast
11. Female reproductive system
12. Urinary tract
13. Male reproductive system
14. Lymph nodes, spleen and bone marrow.
15. Bones and joints.
16. Soft Tissues
17. Pleura and peritoneum
18. Heart
19. Neuromuscular System
20. Eye and ocular adnexa
21. Ear
    Appendum. Myeloid Leukemias and related Neoplasms

back to top

Section E - Pre-clinical Models of Human Cancer

1. Advantages and limitations of models for human cancer
2. Basic Tissue Culture
3. Transgenic technology in the study of oncogenes and Tumor Suppressor Genes
4. Gene knockouts
5. Human tumours in animal hosts
6. Mammary tumour induction in animals as a model for human breast cancer
7. Mathematical models in cancer research
8. Models for tumour growth and differentiation
9. Angiogenesis models
10. Models for tumour cell adhesion and invasion
11. Tumour metastasis models
12. Models for tumour cell-endothelial cell interactions
13. Modelling Tumor Tissue Interactives
14. Models for drug development and drug resistance
15. Models for immunotherapy and cancer vaccines
16. Gene therapy models
17. Models for epithelial carcinomas
18. Models for haematological malignancies

19. Models for melanomas and sarcomas
20. Models for CNS malignancies
21. Models for endocrine cancer

back to top

Section F - The Treatment of Human Cancer

1. Mechanisms of action of cancer chemotherapeutic agents
2. Drug Resistance & Reversal
3. Molecular mechanisms of radiotherapy
4. Antibodies and recombinant cytokines
5. Genetic and Cellular Vaccines
6. Differentiation Therapy
7. Chemoprevention
8. Antisense & ribozyme therapy
9. Hormonal Therapy
10. Antiangiogenic Therapy
11. Targeting the Extracellular Matrix
12. Growth Factor Receptor Blockade
13. Signal Transduction pathway targeting
14. Cell cycle and DNA repair as targets for therapy
15. Gene Therapy - tumour suppressor replacement/oncogene suppression
16. Genetic Prodrug Activatron Therapy (GPAT)
17. Stem cell transplantation
18. Novel surgical strategies in the Management of Cancer
19. Translational research (overview of phase I, II and III clinical trials)

back to top

¡¡

Description:

The Cancer Handbook is a major new reference work that provides a comprehensive overview of cancer, biology and medicine. All aspects of scientific and clinical information in cancer research and medicine are explored in this expansive resource for a wide audience including medical and life science students, as well as researchers, scientists and clinicians in the cancer field and related biomedical areas.

The Cancer Handbook bridges the gap between the basic science of cancer and clinical diagnosis and

treatment providing the ¡®horizontal connections¡¯ between all major areas of cancer research. It stands out from existing oncology textbooks and reference works in that it focuses on all general aspects of cancer research. Over 100 chapters cover fields including: **molecular basis of cancer**, **causation and prevention**, **diagnostic and investigative procedures**, **systemic oncology**, **pre-clinical models**, and the **treatment** of human cancer. Highly structured and cross-referenced, this work has been written with an emphasis on clarity of style, with plenty of diagrams and data redrawn in a consistent and accessible format.

The Handbook also includes an **extensive glossary**. Chapters are referenced with **key sources** and **further readings** pointing the way to the next level of detail.

Available as both a **2-volume print set, and a fully searchable online database**, *The Cancer Handbook* is an essential reference for undergraduate and graduate students in the field of medicine, clinical physicians, oncologists, and other research scientists in the field of cancer as well as molecular and cellular biology.

Editor-in-Chief: Malcolm R. Alison

The Cancer Handbook

Editors

Advisory Editors

- **Professor Robert Weinberg**
  MIT, Whitehead Institute for Biomedical Research
  Cambridge, MA, USA

- **Professor Jerry Shay**
  University of Texas, Southwestern Medical Center
  Dallas, TX, USA
- **Professor Lance Liotta**
  National Cancer Institute, Division of Clinical Sciences
  Bethesda, MD, USA

Editor-in-Chief

- **Professor Malcolm R. Alison**, PhD, DSc, FRCPath;
  Reader in Pathology, Imperial College School of Medicine
  Hammersmith Hospital, Department of Histopathology, United Kingdom.

Section Editors

- **Professor William J. Gullick**
  The University of Kent at Canterbury, The Department of Biosciences
  Canterbury, Kent

- **Dr David. S. Salomon**

National Cancer Institute, Laboratory of Tumour Immunology and Biology, Division of Basic Sciences
Bethesda, MD, USA

- **Dr Nigel Gooderham**
  Imperial College School of Medicine, Section on Molecular Toxicology, Division of Biomedical Sciences
  South Kensington, London

- **Professor Michael A. Bettmann**
  Dartmouth-Hitchcock Medical Center, Department of Radiology
  Lebanon, New Hampshire, USA

- **Dr Larry M. Weiss**
  City of Hope National Medical Center, Division of Pathology
  Duarte, CA, USA

- **Professor Garth L. Nicolson**
  Institute of Molecular Medicine
  Huntington Beach, CA, USA

- **Dr Joerg Haier**
  University Hospital Muenster, Department of Radiology
  Muenster, Germany

- **Professor Nick Lemoine**
  Imperial College School of Medicine, ICRF Molecular Oncology Unit, Departmentt of Cancer Medicine
  London

- **Professor James Abbruzzese**
  University of Texas MD Anderson Cancer Center, Department of Gasterointestinal and Digestive Diseases
  Houston, TX, USA

Authors

Over 100 chapter authors have been selected by the editors and advisors as experts in their topics, and drawn from the international biomedical community.

Chapter 1

Cell and Tissue Organisation

Emma Marshman, Catherine Booth and Christopher S. Potten
Paterson Institute for Cancer Research, Manchester, UK

CONTENTS

- Tissue Classification
- Cell Proliferation and its Control
- Cellular Hierarchies
- Cell Organisation in Specific Tissues
- Control of Tissue Organisation in Development
- Cancer Development and Tissue Organisation
- Conclusions

To understand the changes that take place during cancer development, it is important to understand the basic principles of cell and tissue organisation and the mechanisms that control growth and structure.

TISSUE CLASSIFICATION

Groups of cells that are similar in structure, function and embryonic origin are referred to as tissues. The tissues of the body can be divided into four main groups as follows:

Epithelial Tissue

Epithelial tissue covers most of the free surfaces of the body, both internal and external, and often invaginates to form specialized structures such as glands. For example, it forms the outer layer of skin and the lining of the gastrointestinal tract and breast ducts. In addition to providing physical protection, epithelial cells control permeability, provide sensation and produce specialized secretions from glands, e.g. mucus, hormones and enzymes. Taking all the surface linings and their associated glands and structures together, epithelial tissues make up the major part of total body mass.

Connective Tissue

Connective tissue, or mesenchyme, protects and supports the body and its organs. Types of mesenchymal tissue include cartilage, bone and adipose tissue. The reticuloendothelial system is often considered a type of connective tissue. Reticuloendothelial cells are the defensive and oxygen-supplying cells of the body and are mostly derived from bone marrow precursor cells. The reticuloendothelial cells or haematopoietic cells are distributed throughout the body as free cells in blood and lymph or make up organs such as the spleen and lymph nodes.

Muscle Tissue

Muscle tissue is responsible for movement, such as skeletal movement, but also movement of food, blood and secretions. To carry out this function, muscle cells possess organelles and properties distinct from those of other cells which makes them capable of powerful contractions that shorten the cell along the longitudinal axis. There are three types of muscle tissue: skeletal, cardiac and smooth muscle. The contraction mechanism is similar in all three, but they differ in their internal organisation.

Nervous Tissue

Nervous tissue is specialized for the conduction of electrical impulses from one region of the body to another. Neural tissue consists of two basic cell types, neurons and supporting cells called glial cells. About 98% of the neural tissue in the body is concentrated in the brain and spinal chord with the rest making up the peripheral nervous system.

Since each tissue is made up of a number of specialized cell types that maintain tissue structure and function, there must be exquisite control over cell numbers to maintain the integrity of the tissue. The ability to respond to cell loss (via damage or senescence) varies in the different tissues, since not all cells have the same capacity for regeneration. Tissues can therefore be classified into

three groups depending on this cell replacement capability, as follows:

Rapidly Self-renewing Tissues

In tissues such as the skin, the intestine and the haematopoietic system, there is continuous cell loss either by surface abrasion, by damage or because the cell has aged. This cell loss has to be compensated for by cell production (proliferation), otherwise the tissue would begin to shrink (or expand if proliferation exceeds cell loss). Thus, the number of cells produced by cell division precisely balances cell loss in order for the tissue to maintain its size and mass.

Conditionally Renewing Tissues

In tissues such as the liver, breast, prostate and connective tissue, there is little or no replacement under normal circumstances. However, there is potential for regenerative proliferation under conditions in which the tissue's integrity is significantly compromised, e.g. damage or disease, or in response to hormonal influences.

Essentially Non-renewing Tissues

In some tissues, e.g. the female germ line and the central nervous system, there is little or no cell replacement or capacity for regeneration in the adult.

CELL PROLIFERATION AND ITS CONTROL

Control of cell division within a tissue is particularly important in rapidly self-renewing tissues when proliferation must balance cell loss. Although the exact mechanisms used by tissues to sense the need to increase or decrease cell division are unclear, it is obvious that proliferation must be regulated by a complex network of signals and messages including growth factors, cytokines and hormones. These messages can be produced by the cells themselves (autocrine regulation), may be produced by neighbouring cells of either similar or unrelated cell types (paracrine regulation, e.g. epithelial–mesenchymal interactions), and by circulating hormones (systemic endocrine regulation) as illustrated in **Figure 1**.

Some of the network of signals that control tissue homeostasis may prevent overproduction where necessary or arrest the cell cycle if a cell is damaged. In the latter scenario DNA damage is detected and the cell cycle arrested as it reaches specific checkpoints. These checkpoints will be discussed in more detail in the chapter Regulation of the Cell Cycle. At the checkpoint, the defective DNA can either be repaired or, if too severe, the cell may commit suicide in a process referred to as apoptosis. Apoptosis will be discussed in detail in the chapter *Apoptosis*.

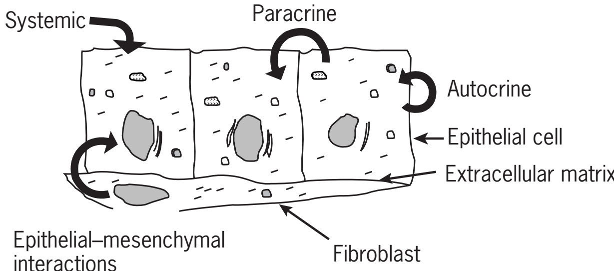


Figure 1 Signal mechanisms involved in cellular communication.

Malfunctions of Control

The first stages of cancer formation are associated with malfunctions in the control mechanisms described above, in such a way that the critical balance between cell proliferation and cell loss by differentiation or apoptosis is disturbed or deregulated. If this balance is shifted in favour of proliferation, the tissue will expand in a progressive and eventually uncontrolled manner, distorting the tissue structure and function. The balance may only need to be shifted slightly in favour of proliferation for a cancer to develop. Cancer development will be further described towards the end of this chapter.

CELLULAR HIERARCHIES

At the bottom of all the hierarchies in the body are the embryonic stem cells. Embryonic stem cells are referred to as totipotent, i.e. they are capable of differentiating into all types of tissue. Embryonic cells may separate and each form complete embryos, e.g. twins. Human embryonic stem cells have recently been isolated from embryonic tissue and can be maintained as undifferentiated cells in laboratory cultures under certain carefully controlled conditions. Even after 4–5 months in culture, these cells are still able to form types of cells from all three embryonic germ layers; including gut epithelium (endoderm), cartilage, bone, smooth muscle and striated muscle (mesoderm) and neural epithelium and embryonic ganglia (ectoderm) (Thomson *et al*., 1998). Studies with embryonic stem cells will give valuable information about the mechanisms controlling differentiation and organisation and may ultimately allow us to grow replacements for tissues or even organs that have been damaged by disease.

In development, these embryonic stem cells are abundant; however, as the animal ages the cellular potency becomes more and more restricted (the capability for wide gene expression becomes more restricted) until ultimately stem cells only remain in tissues capable of regeneration.

Such stem cells are found at the point of origin of cell production within an adult tissue and can produce a steady stream of cells (Potten, 1992). These daughter cells, termed dividing transit cells, can expand their numbers via further cell divisions and mature into functional differentiated cells, called simple transit cells. Simple transit cells are eventually lost from the tissue at the end of their functional lifespan. The linear evolution in the adult animal tissue is therefore organized into a hierarchy or 'family tree' with the cells responsible for cell production at the bottom and the functional cells at the top. The specialization process involved in the progression from the bottom to the top of the hierarchy, termed differentiation, represents a change in the pattern of gene expression which may be the consequence of changes either in the internal programming of the cell or of the external stimuli that affect the cell.

At the bottom of the adult hierarchy, and ultimately responsible for cell replacement in renewing tissues, are the pluripotent stem cells (capable of producing many but not all differentiated cell lineages, i.e. they are not totipotent). In many cases, these cells cannot be identified by a common marker or a single property. Instead, cells are classed as stem cells if they exhibit or have the potential to exhibit the following properties:

1. stem cells are undifferentiated (relative to the cells in the tissue);
2. stem cells are capable of proliferation;
3. stem cells are capable of self-maintenance;
4. stem cells can produce differentiated progeny;
5. stem cells can regenerate the tissue after damage.

When a stem cell divides, under normal circumstances, it is thought to generate a daughter that is another stem cell (thereby maintaining itself) and one daughter that will move up the hierarchy towards differentiation. Although this situation remains to be conclusively proven, it is certainly the average situation that must occur in an adult tissue. Whether the determinants of such division are intrinsic to the stem cell itself or are influenced by the surrounding environment also remains to be determined.

If stem cell numbers need to increase or decrease in response to external stimuli, this asymmetric form of cell division will switch to symmetrical division in which either two stem cell daughter or two nonstem cell daughters are produced. Stem cell expansion will inevitably increase cellular production (i.e. speed up regeneration, generate hyperplasia), whereas stem cell removal will reduce or remove cellular production (depending on how many stem cells remain in the tissue), e.g. generate aplasia or hypoplasia.

The next steps in the life of a nonstem cell daughter, particularly in a rapidly renewing tissue, are the amplification of cell numbers. The daughter cells divide a number of times and are known as transit amplifying cells. During this time the cells gradually appear to lose their stem cell properties and acquire a more mature phenotype until, after a given number of divisions, they are fully differentiated cells. These transit amplifying cells are therefore generally a short-lived phenotype, although during the early cell generations they may be called upon to behave as stem cells in a trauma situation in some tissues. With successive divisions they eventually lose this ability. The later-generation differentiated cells then perform the function for which they were generated, gradually senesce and die. This is also therefore a form of programmed cell death.

The advantage of such an organisation is that only a few stem cells are needed to maintain a whole tissue. Generally, these stem cells have a slow cell cycle time which allows for genetic housekeeping, i.e. time to repair any genetic damage. Small numbers of stem cells followed by around five generations of transit amplifying cells create an environment in which the greatest risk of introducing a mutation (during division) is in the transit cells (which are ultimately lost from the tissue) rather than in the long-lived stem cells. In conditionally renewing tissues the organisation is less clear. Although stem cells must exist, it is possible that they are normally quiescent or are cycling very slowly, and are only activated by trauma or hormonal stimuli.

The progression from stem cell to differentiated cell could be preprogrammed but is more likely to be controlled by extrinsic factors. An organized hierarchy obviously experiences (and/or is able to respond to) different control signals at different stages. This can be aided by a physical organisation, such that there is a spatial distribution within the hierarchy controlled by a series of microenvironments or niches. A gradient of controlling factors probably exists along the maturation axis.

Particularly important in the microenvironment is the basement membrane upon which epithelial cells sit. This basement membrane is a highly organized extracellular matrix (ECM) made up of proteins such as collagen and laminins. The effects of the matrix are primarily mediated by cell adhesion molecules such as integrins and cadherins which are families of cell surface receptors. Cell adhesion molecules help to connect the exterior of the cell with the interior of the cell in two ways: by transducing signals initiating from the extracellular interactions and by mediating structural linkages between the cytoskeleton and the ECM of other cells (Horwitz and Werb, 1998). These processes will be further described in the chapters *Wnt Signal Transduction* and *Extracellular Matrix: The Networking Solution.*

CELL ORGANISATION IN SPECIFIC TISSUES

To illustrate the points made in the previous section, the stem cells and hierarchies of a number of tissues will be described in more detail.

Haematopoietic System

The hierarchical organisation of the continually renewing cells in the bone marrow has been extensively studied. All mature blood cells in the body are derived from a small number of stem cells that reside in the bone marrow in a process called haematopoiesis. Over 10^{11} new cells are produced daily to maintain homeostasis since the majority of mature blood cells are short-lived. In addition, normal daily cell replacement must also be sporadically increased to fight infection or to compensate for blood loss.

The haematopoietic lineage is shown in **Figure 2**. The most primitive stem cell of the bone marrow is the pluripotent stem cell which has the capability to produce all the different cell types of the blood. To add another level of complexity, this pluripotent stem cell may itself be part of a stem cell hierarchy. Myeloid and lymphoid stem cells are produced from the pluripotent stem cell population. The myeloid stem cell then goes on to produce a number of progenitor cells which are the precursors of the six types of mature functional myeloid cells: erythrocytes, thrombocytes, eosinophils, macrophages, mast cells and neutrophils. These cells have different functions within the immune system and in the blood. There may be further as yet unknown subdivisions in the stem cell hierarchy. The lymphoid stem cell produces a number of lymphoid progenitors which mature into B and T lymphocytes to provide defence against pathogens or toxins.

Although mature blood cells can be distinguished from each other, stem cells and progenitor cells have no specific distinguishing features under the microscope. Identification of early progenitor cells and stem cells is also made difficult by the low incidence of these cells in blood. For example, pluripotent stem cells are thought to make up only 0.01–0.1% of total bone marrow cells (Heyworth *et al*., 1997). Functional assays have been devised, the first of which was described by Till and McCullogh (1961). This method involves transplantation of some healthy bone marrow cells into mice whose own bone marrow has been destroyed by irradiation. The transplanted cells produce colonies of differentiated haematopoietic cells in the spleen which can be counted. In addition to functional assays, external markers have been used to identify progenitor cells. Myeloid and lymphoid stem cells and early progenitor cells can be separated from blood by antibodies that react to specific antigens only present on these cell types, e.g. CD34 antigen which is expressed on 0.5–5% of human bone marrow cells. Methods for separation of pluripotent stem cells using specific markers are under development.

In the bone marrow, stem cells and their progeny are exposed to a number of different stimuli including physical interactions with other cells mediated by cell adhesion molecules, interactions with extracellular matrix molecules such as collagen and fibronectin and exposure to growth-stimulatory and growth-inhibitory chemicals called cytokines. There are over 15 cytokines involved in haematopoiesis and these are produced by a number of cell types including the mature cells themselves, e.g. neutrophils, B and T cells, as well as by fibroblasts and bone marrow stromal cells providing autocrine and paracrine regulation (Heyworth *et al*., 1997). All these signals coordinate the self-renewal and differentiation of the stem cells and the formation of the mature cell types.

The role of cytokines in determining which type of cell (e.g. mast cell or neutrophil) an early progenitor cell differentiates into is highly complex. Some cytokines have many target cells, whereas others are much more restricted. Interleukin (IL-3), for example, can stimulate stem cells

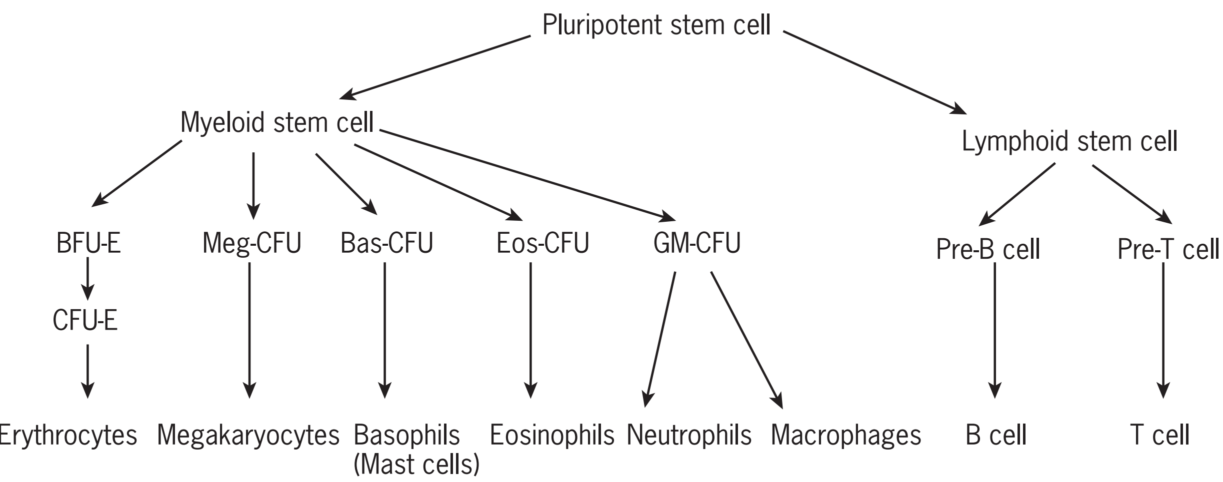


Figure 2 Haematopoietic cell lineage.
Abbreviations: BFU-E, erythroid burst-forming unit; CFU-E, erythroid colony-forming unit; Meg-CFU, megakaryocytic colony-forming unit; Bas-CFU, basophilic colony-forming unit; Eos-CFU, eosinophilic colony-forming unit; GM-CFU, granulocyte macrophage colony-forming unit.

to produce myeloid progenitor cells and can also stimulate myeloid progenitor cells to produce a number of mature cell types (Dexter, 1993). Another example is granulocyte–macrophage colony-stimulating factor (GM-CSF) which acts on the granulocyte–macrophage progenitor cell and the eosinophil progenitor cell to produce neutrophils, macrophages and eosinophils. In contrast, some growth factors have direct effects on only one cell population, e.g. erythropoietin, which acts only on the erythroid progenitor cell to produce erythrocytes. Other cytokines mainly influence the maturation of cells rather than the proliferation of progenitor cells, e.g. IL-5 and eosinophil development.

Originally it was thought that haematopoiesis was regulated solely by modulation of the production of these directly acting cytokines, e.g. stem cells would be acted upon by IL-3 to produce progenitor cells and then specific cytokines would be made to induce maturation of the progenitor cells into whichever specific cells were required by the bone marrow. It is now known that control is exerted at a more complex level such that a certain growth factor alone will not have effects on a particular cell type; however, when it is combined with another factor proliferation or maturation can be induced. For example, lymphoid stem cells will not respond to macrophage colony-stimulating factor or IL-1 alone, but are stimulated in the presence of a combination of these two growth factors.

The haematopoietic cell lineage has illustrated the complex communication network required for the differentiation of relatively unknown stem cells into the specific cells of the blood.

Small Intestine

The epithelium of the small intestine provides another example of a self-renewing tissue which has been studied for many years. The tissue organisation of the small intestine is different from that described for the haematopoietic system, being highly polarized and structured. However, the regulation mechanisms are equally complicated and largely unknown at present.

In the small intestine, epithelium covers finger-like projections called villi and flask-shaped crypts located at the base of the villi which are embedded in the connective tissue (see **Figure 3**). Epithelial cells are produced in the lower part of the crypt and migrate up the crypt on to the villi and are continuously shed from the villus tip. In common with the haematopoietic system, migration from the early precursor cells is accompanied by differentiation and specialization. Cells differentiate into three functional cell types as they move up the crypt: the predominant enterocyte, the mucus-secreting goblet cell and the peptide hormone-secreting enteroendocrine cells. In addition, a number of cells migrate down to the base of the crypt to become the fourth cell type, the Paneth cells. Paneth cells secrete a number of proteins including lysozyme, which is thought to play a role in fighting bacterial infection.

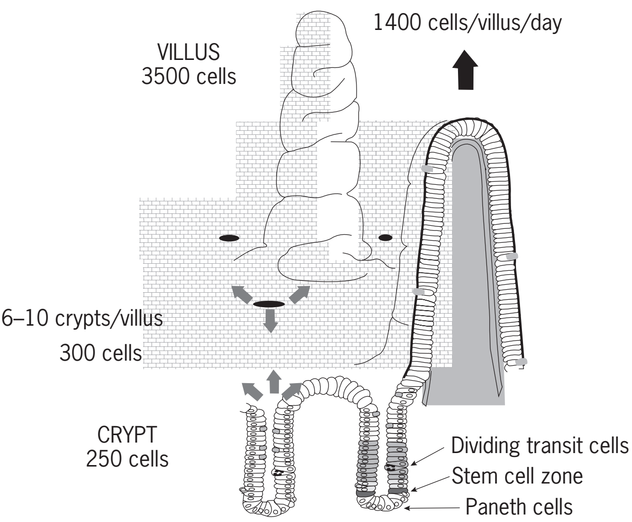


Figure 3 Organisation of small intestinal epithelium.

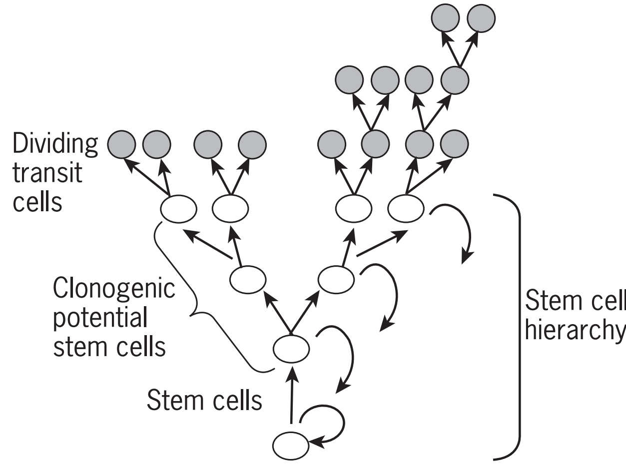


Figure 4 Proposed stem cell model for the small intestine.

Replacement of cells shed at the villus tip must be balanced by cell production in the crypt, at a rate of about 10^{10} cells per day in humans (**Figure 3**) (Potten, 1992). Cell replacement is achieved by stem cells located amongst or just above the Paneth cells at the base of the crypt. Unfortunately, there are no markers for intestinal stem cells and at present, characterization studies can only be carried out by disturbing the system and observing the outcome. A stem cell model has been proposed based on clonal regeneration studies following radiation or drug exposure (**Figure 4**). The proposed model suggests that there are 4–6 ancestor or functioning stem cells per crypt (Potten, 1998). These stem cells are very sensitive to toxic insults (e.g. radiation and some chemotherapeutic agents) and are unable to repair damaged DNA. If damaged they readily initiate apoptosis and die. This sensitivity may reflect the need to avoid repopulation of the crypt with cells containing damaged DNA, and thereby preserves the integrity of the tissue. Stem cells that die, however, are

easily replaced by the other surviving stem cell members or by their immediate daughter cells, which make up the second tier of the hierarchy. The second tier stem cells have a better repair capacity and, if not required to regenerate the first tier (such as in a normal situation), they are displaced into the transit compartment. If this second tier is destroyed, a third tier may also exist that contains about 20 even more resistant stem cells with the best repair capacity. These three tiers therefore make up a population of around 30–40 potential stem cells–cells that are acting as stem cells or retain the ability to act as a stem cell if required. Since each of these cells can regenerate a clonal population (a crypt), they are also termed clonogenic cells. Above the level of clonogenic stem cells there are about 124 dividing transit cells which have no stem-cell attributes. These proliferative cells move or are displaced at a rate of 1–2 cell positions per hour from the crypt on to the villus (Potten, 1992, 1998).

Regulation of cell proliferation in the gut is not fully understood. However, a large number of factors are known to be involved, including growth factors, cytokines and ECM molecules. The epidermal growth factor (EGF) family is one group of substances known to stimulate proliferation and includes epidermal growth factor itself and TGF-α (Potten *et al*., 1997). In contrast, the TGF-β family of growth factors have been associated with negative regulation or inhibition of crypt cell proliferation (see also the chapter *Signalling by TGF-β*). In common with growth factors, *in vitro* studies suggest that some interleukins have stimulatory effects (e.g. IL-4) and some have inhibitory effects (e.g. IL-11 and IL-6).

The ECM underlying the epithelium plays a role in a number of key processes, one of which is cell migration. The process of migration is not fully understood and it was initially thought that cells moved in tandem with underlying connective tissue. More recent studies suggest that cells 'walk' over stationary ECM which contains a number of adhesion molecules such as E-cadherin, laminin, fibronectin, tenascin and collagen. Migration is thought to involve decreased cell attachment to one or more of these adhesion molecules, since adhesion molecule expression patterns vary along the crypt/villus axis. The stationary nature of stem cells may be due to their strong anchorage to the stroma. For example, fibronectin, which is a particularly 'sticky' adhesion molecule, is abundant in the crypt whereas tenascin which is less adhesive is predominant on the villus. In addition, movement is controlled by the expression/availability of integrins, epithelial cell receptors for these adhesion molecules – a cell can only be influenced by adhesion molecule levels within the basement membrane if it expresses the appropriate receptors.

Epidermis

The outermost layer of the skin, the epidermis, is another example of a self-renewing tissue. The epidermis is predominantly made up of keratinocytes (about 80% of the total). Other epidermal cell types exist with specific functions: melanocytes give the skin its pigmentation and afford some protection against ultraviolet light, Merkel cells sense fine mechanical events and Langerhans cells form part of the body's immune system.

The outer surface of the epidermis is called the stratum corneum and is composed of a layer of thin, dead keratinocytes. These cells bear little resemblance to normal keratinocytes, since by the time they reach the surface their nucleus and internal organelles have disappeared and they are reduced to thin plates of keratin. Keratins are a family of insoluble proteins that form intermediate filaments within cells and confer mechanical strength. Surface keratinocytes represent the final mature functional differentiated cells of the skin. These cells are continually being shed or lost and therefore perpetual cell replacement is required to maintain epidermal function.

Below the stratum corneum are three other epidermal cell layers: the granular layer, the spinous layer and the basal layer. These epidermal regions are depicted in **Figure 5**, although there are many more layers of cells than are shown in the diagram. In common with the small intestine and the haematopoietic system, stem cells are responsible for the regenerative potential of skin. These stem cells are located within the basal layer (Lavkar and Sun, 1983). Studies carried out on mouse epidermis suggest that 5–12% of cells in the basal layer are stem cells (Potten, 1992). Transitory dividing cells produced from these stem cells make up about 50% of the basal layer with the remaining basal layer cells being postmitotic and having no proliferative characteristics. These cells are committed to terminal differentiation and achieve this as they slip out of the basal layer and migrate into the spinous layer, where they flatten. From the spinous layer, cells progress up into the granular layer until they reach the stratum corneum where they are eventually shed. The stem-cell progeny generate a discrete column of cells, from basal cell to keratinized cell, arranged in a hexagonal pattern and called an epidermal proliferative unit (Potten, 1981). It has been estimated that it takes the human keratinocyte between 26 and 42 days to travel from the basal layer to the outermost cornified layer and therefore it takes 1–2 months for the epidermis to replace itself completely.

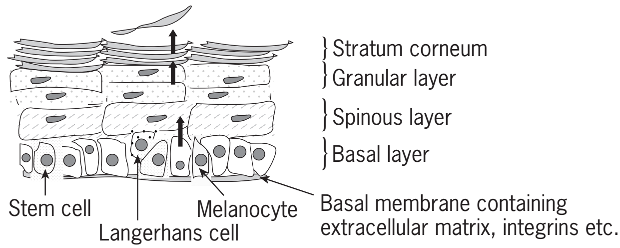


Figure 5 The murine epidermal proliferative unit.

In common with the small intestine, the underlying ECM plays a key role in basal layer processes. It has been suggested that the ECM mediates adhesion, regulates terminal differentiation and aids cell movement upward from the basal layer. When basal keratinocytes become committed to undergo terminal differentiation, their ability to adhere to components of the ECM decreases and upward cell migration occurs (Jones and Watt, 1993). Populations of putative stem cells that are greater than 90% pure have been isolated on the basis of their adhesive properties.

The epithelial cells of the skin, and indeed other sites of the body, are able to form a barrier due to a number of functionally and structurally distinct epithelial cell junctions, including tight junctions, gap junctions, desmosomes and hemidesmosomes. Tight junctions seal neighbouring cells together to stop water-soluble molecules leaking between the cells and confine transport proteins either to the outward-facing membrane (apical) or to the inner membranes (basolateral) to control the passage of certain chemicals (e.g. glucose transport in the small intestine). In contrast, gap junctions are involved in cell–cell signalling. Gap junctions are intercellular channels made up of connexin proteins that allow inorganic ions and other small water-soluble molecules to pass directly from the cytoplasm of one cell to the cytoplasm of another, thereby coupling the cells both metabolically and electrically. To maintain mechanical strength, cells are linked together with desmosomes. Desmosomes consist of a dense plaque of intracellular attachment proteins (including plakoglobin and desmoplakins) which are associated with rope-like intermediate keratin filaments that form a continuous network throughout the tissue. Hemidesmosomes, or half-desmosomes, connect the basal surface of epithelial cells to the underlying basement membrane such as that which separates the epidermis and the dermis. In addition to aiding attachment, hemidesmosomes have also been found to be important in modulating the organisation of the cytoskeleton, proliferation and differentiation. These effects are mediated by integrins which transduce signals from the ECM to the interior of the cell as described earlier. Absence or defects of hemidesmosomal proteins can result in devastating blistering skin diseases.

Breast

The breast or mammary gland is an example of a conditionally renewing tissue in that cell replacement is generally limited except under certain conditions, e.g. pregnancy.

The organisation of the breast changes during three developmental phases. The first stage occurs in the foetus where mammary glands arise as buds from the epidermis which elongate to form simple, branched ducts. At puberty, there is rapid extension and branching of the ducts which terminate in globular structures called terminal end buds. These terminal end buds and terminal ducts then go on to form lobules of alveolar buds.

The third phase of mammary development occurs during pregnancy and lactation and at this stage the breast can be considered to be morphologically mature and functionally active. The alveolar buds and lobes subdivide further, giving rise to large clusters of alveolar lobes. During lactation, the clusters of alveolar lobes become distended and form secretory alveoli lined with alveolar cells which produce milk. After cessation of lactation, involution of the breast occurs where the secretory cells of the alveoli degenerate and disappear. Similarly, after the menopause, there is progressive involution of the ductal and glandular components of the breast. The connective tissue supporting the breast also degenerates with loss of stromal cells and collagen fibres.

In the normal breast, the ducts and lobes of the mammary gland are separated from the stroma by a basement membrane. This basement membrane is lined with two cell types, an outer lining of myoepithelial cells containing myofilaments and an inner lining of epithelial cells. As described in the section regarding the epidermis, the epithelial cells of the breast are connected together with desmosomes whereas myoepithelial cells connect to the basement membrane with hemidesmosomes.

Studies in rodent mammary glands indicate that epithelial cell types and alveolar cells arise from stem cell populations capable of generating the fully differentiated lactating mammary gland. These stem cells are thought to be present in the basal cell layer of ducts and end buds, although little more is known about their identity (Rudland *et al*., 1997). It has been suggested that stem cells can give rise to either ductal epithelial cells in a reversible manner or myoepithelial cells in an irreversible manner. Alveolar cells are thought to be derived from ductal epithelial cells.

As with the other tissues described, the differentiation of cells produced by breast stem cells is strictly controlled. Unlike the haematopoietic system where differentiation is controlled mainly by paracrine and autocrine secretions, the breast is also subject to control by circulating hormones secreted by the pituitary, ovary and adrenal glands. For example, during each menstrual cycle at about the time of ovulation, there is an increase in lobular size and epithelial cell vacuolization under the influence of oestrogens and rising progesterone. When menstruation occurs, the fall in hormone levels causes lobular regression. Similarly in pregnancy, oestrogens and progesterone stimulate proliferation and development, and prolactin released by the pituitary gland activates the production of alveolar cells. Additionally, lactation is triggered by the release of oxytocin, which causes contraction of the smooth muscle components of the myoepithelial cells surrounding the alveoli leading to milk expulsion.

Local growth hormones are also important since the growth promoting effects of oestrogen are believed to be mediated by TGF-α and insulin-like growth factor-1 (IGF-1) which increase epithelial cell growth and inhibit

myoepithelial cell differentiation (Rudland *et al.*, 1997). Additionally, production of basic fibroblast growth factor (bFGF) by breast stem cells may regulate their own growth and that of myoepithelial cells in an autocrine/paracrine fashion.

Liver

The liver is another example of a conditionally renewing tissue. Epithelial cells make up the majority of the liver, particularly hepatocytes and also biliary duct epithelium. The hierarchial organisation of the liver, however, is not fully understood. Unlike the tissues mentioned above, where one population of stem cells is responsible for cell replacement, it seems likely that a two-compartment system is operative in this tissue (Alison, 1998). First, in the event of damage, hepatocytes are able to regenerate themselves very efficiently. This is unusual because the ability to regenerate is normally characteristic of undifferentiated stem cells, and hepatocytes are thought to be fully differentiated. A second compartment is activated if the function of surviving hepatocytes is drastically impaired and involves generation of new hepatocytes from stem cells. Liver stem cells, believed to be located in biliary ductules, produce oval cells. Oval cells are then thought to differentiate into the functional cell types, e.g. hepatocytes. The proposed model is illustrated in **Figure 6**.

Others

As indicated above, knowledge about the organisation of tissues ranges from the haematopoietic system and small intestine, which have been extensively studied, to the liver and breast, where information is more limited. Recently, a number of developments have been made in understanding the hierarchy of other tissue types. For example, mesenchymal stem cells have been isolated from human bone marrow. These cells replicate as undifferentiated cells and have the potential to differentiate into lineages of mesenchymal tissues including bone, cartilage, fat, tendon, muscle and marrow stroma (Pittenger *et al.*, 1999). Neuronal stem cells have also been discovered in foetal brain which can develop into neurons and glial cells if certain growth factors are present *in vitro* (Vogel, 1999).

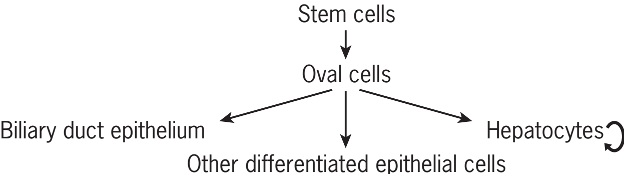


Figure 6 A proposed model for liver cell generation and regeneration.

CONTROL OF TISSUE ORGANISATION IN DEVELOPMENT

Regulation of tissue organisation in development is likely to involve a number of highly complicated mechanisms; however, one group of genes called homeobox genes has already been found to play a significant role in this process. Homeobox genes are a family of regulatory genes encoding transcription factors (homeoproteins) that can activate or repress the expression of a large number of target genes and so determine cell fate and general pattern formation. One major family of homeobox genes termed Hox genes control the identity of various regions along the body axis. These Hox genes are activated in sequence such that early Hox genes that control hindbrain development, followed by activation of Hox genes that control the thoracic region and late genes that control the lumbo-sacral region. Perturbing the expression of these factors can induce gross changes in tissue, organ and even limb development. For example, synpolydactyly, an inherited disease characterized by hand and foot malformation, is caused by expansions of the *HOXD13* gene.

In addition to developmental regulation, certain homeobox genes are also involved in inducing differentiation in renewing tissues. For example, the *cdx* genes are involved in controlling intestinal epithelial cell differentiation, possibly by transducing signals from laminin-1 in the underlying mesenchyme, and *HOXA9* and *PBX1* are some of the many genes involved in the control of haematopoietic differentiation. Altered expression of any of these can suppress differentiation and ultimately lead to tumour formation.

CANCER DEVELOPMENT AND TISSUE ORGANISATION

The chapter *Overview of Oncogenesis* will discuss the mechanisms involved in carcinogenesis in detail, but briefly the process is thought to involve a number of steps. First, a cell experiences a mutation that may or may not influence its immediate behaviour. This cell may then be more susceptible to subsequent mutations and, over time, gradually accumulate enough damage such that the normal control or 'braking' mechanisms is perturbed. This gradual accumulation of mutations is therefore known as the multistage model of carcinogenesis and explains why cancer is generally a disease of old age unless, for example, the primary mutation is an inherited disorder.

Within a tissue experiencing this process, the first observable histological stage is hyperplasia or cellular overgrowth, although this term must be used carefully since tissue regeneration in response to wounding is itself a form of hyperplasia. Hyperplasia can therefore be benign in addition to cancerous. Since in the adult hyperplasia can only occur in proliferating tissues it is not surprising that

almost all cancers arise in rapidly renewing or conditionally renewing tissues. In each case there is a malfunction in cellular homeostasis and cell production exceeds cell loss.

The origin of cell production, and the only permanent resident of a renewing tissue, is the stem cell. Cancers can therefore be thought of as stem cell diseases (transformation of a maturing cell would have no long-term effect since even if it divides a few times, each cell is ultimately lost from the tissue in a relatively short time frame). An expansion of stem cell numbers can therefore lead to hyperplasia. Normally such an expansion would be detected by the tissue and the excess stem cell removed, via apoptosis. However, if this does not occur, cellular output will be dramatically increased. For example, in the colon expression of the anti-apoptotic gene *bcl-2* may allow the survival of a single extra stem cell in an intestinal crypt (Potten *et al.*, 1997). This alone can lead to 128 extra cells being produced by that one crypt (owing to the expansion by the transit amplifying cells). As the animal ages these excess stem cells persist and may experience further mutations (e.g. in apoptosis regulation such as by *p53*, growth factor signal transduction such as in *SMAD* and *ras*, DNA repair by mismatch repair enzymes such as MSH2, or in cellular adhesion such as changed integrin or E-cadherin expression), thereby increasing cancer risk. These mutations generally occur in three vital areas – regulation of cell division in the renewing population (restraint), DNA repair (such that the normal DNA is not maintained) and interactions with the extracellular environment (cells or matrix). Together these will subvert the normal differentiation process and allow unrestrained tissue growth without the accompanying levels of cell death, followed by invasion and metastasis into other tissue sites.

CONCLUSIONS

The organisation of cells and tissues has been discussed in development and in the normal adult and we have attempted to highlight the complex nature of the regulation processes that control cell proliferation, differentiation and regeneration. Cancer development provides us with an excellent example of the devastating effects observed when these processes are subverted and emphasizes the need for such exquisitely controlled mechanisms.

REFERENCES

Alison, M. (1998). Liver stem cells: a two compartment system. *Current Opinion in Cell Biology*, **10**, 710–715.

Dexter, T. M. (1993). Synergistic interactions in haemopoiesis: biological implications and clinical use. *European Journal of Cancer*, **29A**, S6–S9.

Graham, G. J. and Pagnell, I. B. (1992). The haematopoietic stem cell: properties and control mechanisms. *Seminars in Cellular Biology*, **3**, 423–434.

Heyworth, C. M., *et al*. (1997). Growth factors and the regulation of haematopoietic stem cells. In: Potten, C. S. (ed.), *Stem Cells*. 423–446 (Academic Press, London).

Horwitz, A. R. and Werb, Z. (1998). Cell adhesion and the extracellular matrix: recent progress and emerging themes. *Current Opinion in Cell Biology*, **10**, 563–565.

Jones, P. H. and Watt, F. M. (1993). Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell*, **73**, 713–724.

Lavker, R. M. and Sun, T.-T. (1983). Epidermal stem cells. *Journal of Investigative Dermatology*, **81**(1S), 121–127.

Pittenger, M. F., *et al*. (1999). Multilineage potential of adult human mesenchymal stem cells. *Science*, **284**, 143–147.

Potten, C. S. (1981). Cell replacement in epidermis (keratopoiesis) via discrete units of proliferation. *International Review of Cytology*, **69**, 271–317.

Potten, C. S. (1992). Cell lineages. In: McGee, J. O'D., *et al*. (eds), *Oxford Textbook of Pathology*, Vol. 1, 43–52 (Oxford University Press, Oxford).

Potten, C. S. (1998). Stem cells in gastrointestinal epithelium: numbers, characteristics and death. *Philosophical Transactions of the Royal Society of London*, **353**, 821–830.

Potten, C. S., *et al*. (1997). The intestinal epithelial stem cell: the mucosal governor. *International Journal of Experimental Pathology*, **78**, 219–243.

Rudland, P. S., *et al*. (1997). Mammary stem cells in normal development and cancer. In: Potten, C. S. (ed.), *Stem Cells*. 147–232 (Academic Press, London).

Thomson, J. A., *et al*. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, **282**, 1145–1147.

Till, J. E. and McCulloch, E. A (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, **14**, 213–222.

Vogel, G. (1999). Harnessing the power of stem cells. *Science*, **283**, 1432–1134.

FURTHER READING

Cillo, C., Faiella, A., Cantile, M., and Boncinelli, E. (1999). Homeobox genes and cancer. *Experimental Cell Research*, **248(1)**, 1–9.

D'Andrea, A. D. (1994). Haematopoietic growth factors and the regulation of differentiative decisions. *Current Opinion in Cell Biology*, **6**, 804–808.

Loeffler, M. and Potten, C. S. (1997). Stem cells and cellular pedigrees – a conceptual introduction. In: Potten, C. S. (ed.), *Stem Cells*. 423–446 (Academic Press Ltd., London).

Schwarzbauer, J. (1999). Basement membrane: putting up the barrier. *Current Biology*, **9**, R242–244.

Stappenbeck, T. S., Wong, M. H., Saam, J. R., Mysorekar, I. U. and Gordon, J. I. (1998). Notes from some crypt watchers: regulation of renewal in the mouse intestinal epithelium. *Current Opinion in Cell Biology*, **10**, 702–709.

Chapter 2

Regulation of the Cell Cycle

Arthur B. Pardee
Dana-Farber Cancer Institute, Boston, MA, USA

CONTENTS

- Overview
- The Normal Cell Cycle
- Molecular Biology of the Cycle
- Regulation of Cycle Phases
- Checkpoints, Mutations and Cancer
- Cancer Therapy and the Cycle
- Acknowledgement

OVERVIEW

In each of us are about 50 trillion living cells, all of which originated from only one cell, a fertilized egg. As we developed into adults this cell divided into two cells, these into four, and so forth, at least 45 times. The orderly process by which one cell becomes two is named the cell cycle. This cycle is fundamental not only for understanding cell growth, but also for replacement of cells lost by damage, as in wound healing and from the normal wear and tear of our bodies. The cell cycle is evidently tightly regulated, because we usually make new cells only when they are needed. Indeed, cancers arise when cell growth control is defective. 'Cancer is a wound that does not heal.'

One should remember that cells in most tissues are not usually progressing through the cycle, but are at rest, happily performing their specialized functions in support of the whole organism. But as exceptions, bone marrow, intestinal epithelial and some other cells are constantly dividing. A cell has a life cycle. It is formed, eventually becomes worn and dies by a programmed cell-death mechanism called apoptosis. Thereafter, nearby cells grow and divide to replace it. Cell numbers are balanced by proliferation versus apoptosis. After a cell becomes cancerous the balance is perturbed in favour of proliferation. These facts can be overlooked because much research is performed with cells put into culture and under conditions that permit proliferation.

THE NORMAL CELL CYCLE

History of Cell Cycle Biology

Before the cell cycle, microscopy revealed an interval of about 1 day between successive divisions of one cell into two cells. Until about 50 years ago, no changes could be observed during most of this interval, until about 1 hour before division when chromosomes which contain the hereditary material become visible and are equally partitioned between the two daughter cells, a complex process termed mitosis that takes place through M phase.

Production of two cells from one requires duplication of all of the myriad molecules that compose each cell. The most evidently duplicated molecules are deoxyribonucleic acid (DNA), the heredity-containing material in chromosomes. DNA does not duplicate throughout the cycle, but only during several hours in mid-cycle. This period is named the S phase, for DNA synthesis. Other molecules are duplicated at different times throughout the cycle. These findings about DNA synthesis (Howard and Pelc, 1951) showed that the cycle is organized as a series of events, and created the present framework of its four phases: a 'gap' (G_1 phase) during which a cell prepares for DNA synthesis, DNA synthesis (S phase), preparation for mitosis (G_2 phase) and the mitotic M phase, after which the cell divides and two new cycles commence (**Figure 1**). For a historical summary of biology of the cycle, see Baserga (1985).

Quiescence

Commencing by considering normal animal cells, most of the cells within us are in a quiescent state (G_0 phase). They have left their cycling during the G_1 state, so in quiescent cells DNA has not yet duplicated. But quiescent cells differ from G_1 cells in many other properties, in particular lacking molecules required for growth. This fact told us that the molecular switch that controls growth versus quiescence, and that is defective in cancers, is to be found in G_1 phase (Pardee, 1989).

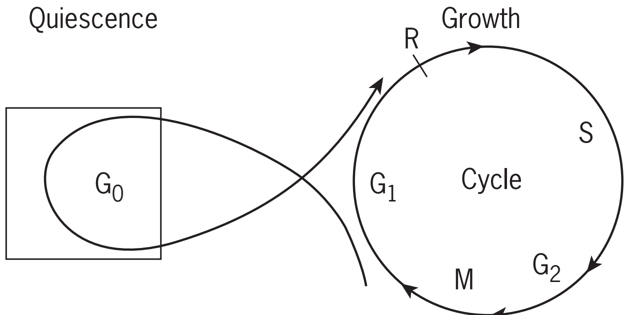


Figure 1 The basic cell cycle. The consecutive phases of the cycle, entry from G_0 and exit from G_1 to quiescence and differentiation are indicated.

G_1 Phase

When cells are activated to proliferate they advance from G_0 to G_1 phase, during several hours after several growth factor proteins are provided in their environment. These include epidermal growth factor (EGF) and insulin-like growth factor (IGF-1), which must overcome inhibitions by crowding of cells and the negative factor TGF-β. Growth factors and nutrients must be supplied from the blood in an organism. To grow cells outside the body, in tissue culture, a nutrient medium is required, in which growth factors are usually supplied by adding blood serum. Cells complete their cycle and then become quiescent after growth factors have been removed.

The length of time that cells in a culture spend in G_1 phase is highly variable, e.g. from 6 to 24 h, unlike the fairly uniform time they spend in the other phases. Many other synthetic biochemical processes take place in G_1 phase (see below).

S Phase

The requirement for growth factors to pass through G_1 phase is lost at the restriction point (R), located shortly before cells start to synthesize DNA. At the beginning of S phase, enzymes involved in DNA duplication increase, and they move into the nucleus where DNA is duplicated, from the surrounding cytoplasm where proteins are synthesized. Then at specific times during the next 6–8 h the DNAs of the perhaps 40 000 genes located on 23 pairs of chromosomes are replicated, each according to a timed program. For example the dihyrofolate reductase gene replicates quickly in very early S phase, but other genes are duplicated at other specific times throughout S phase.

G_2 Phase

After DNA synthesis is completed, several hours are required before initiation of mitosis, presumably to produce needed enzymatic machinery. Many G_2 products are unknown; a terminal one is the maturation promoting factor (MPF).

M Phase and Cell Division

Mitosis requires less than 1 h, and is subdivided into four main stages, in which the duplicate chromosomes pair and condense, and a mitotic 'machinery' consisting mainly of microtubule proteins segregates them equally between the two daughter cells. At completion of M phase, proteins of the mitotic apparatus are destroyed. The daughter cells then become separated, and each can repeat the cycle processes.

MOLECULAR BIOLOGY OF THE CYCLE

Signalling Molecules

Comparisons of growth of mammalian cancer and normal cells in culture revealed in 1974 that the basis of cancer's deranged growth control is located in G_1 phase, shortly before initiation of DNA synthesis (Pardee, 1989). In the same year, genetic studies of the cycle were intiated; research with cycle-controlling yeast mutants led to the discovery of numerous cycle-regulatory genes (Hartwell and Kastan, 1994). Biochemistry and molecular biology soon followed, with the identification of new genes and key enzymes; in particular proteins named cyclins that activate these kinases were discovered by Hunt and Ruderman (see review by Murray and Hunt). These rise and fall during the cycle because of periodic changes in their synthesis and destruction (Minshull *et al.*, 1989). Cyclin-dependent kinases (cdks) that phosphorylate proteins required for cell cycle progression were identified (Nurse *et al.*, 1998). Several proteins that inhibit these kinases and that vary during the cycle were discovered later. This involvement of both positively and negatively acting molecules illustrates the Ying–Yang principle of dynamic opposing actions, frequently seen in biology.

G_1 Phase Kinases, Cyclins and Inhibitors

We will outline the main steps of growth activation and control in G_1 phase, but this process is too complex to describe here fully (**Figure 2**) (see Murray and Hunt; Andreef). In summary, a biochemical network regulates the critical process of controlling cell growth during G_1 phase. Numerous nutrients including sugars, salts, vitamins and essential amino acids are required for cell growth (Baserga, 1985). Externally supplied growth factors start the cell cycle, from G_0 into G_1 phase. They initiate a multi-step cascade of signals that ultimately

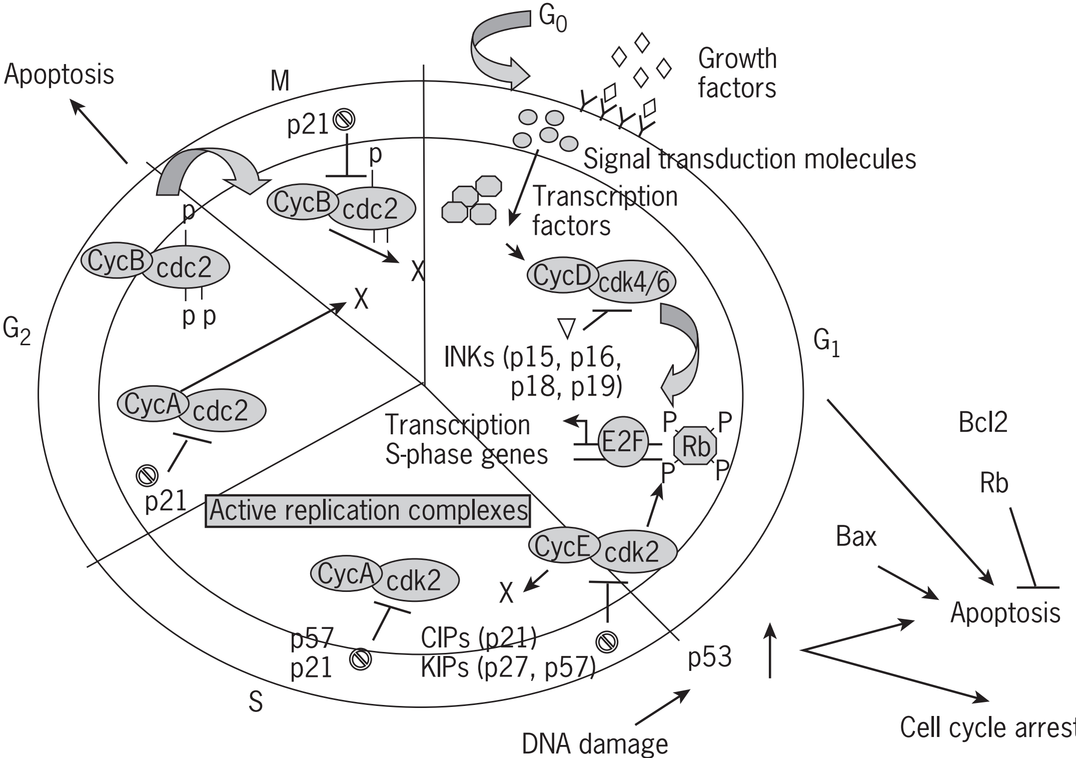


Figure 2 Cell-cycle control molecules. Some of the many molecules that provide growth-regulating signals throughout the cycle are shown, and are discussed throughout this chapter. (Adapted from Ford and Pardee, 1999.)

activates genes to produce their messenger RNAs and proteins, and which culminates in the starting up of DNA synthesis.

The growth factors bind extracellulary to their specific receptor proteins that traverse the membrane surrounding each cell (**Figure 3**). These receptors conduct the external signal to the interior of the cell, and there they activate the receptor's special kinase. These then turn on a cascade of signals involving other proteins including Ras, Fos, Myc and MAP kinases. The Ying–Yang principle is again involved, as illustrated by phosphorylations catalysed by PI-3 kinase that are balanced by dephosphorylations catalysed by the PTEN phosphatase enzyme. The activation of G_1 phase results in expression of at least 100 genes.

The discovery of cyclins, which are the key proteins regulating transition through the cycle (Roberts, 1999), was soon followed by discoveries of multiple cdks. Their complexes with cyclins catalyse stages of cell cycle progression (**Figure 4**). As cells proceed through the cycle, four major cyclins are produced sequentially (D, E, A and B), and they activate several cyclin-dependent kinases. Central is cyclin D which increases in early to mid G_1 phase and regulates cyclin-dependent kinases cdk4 and cdk6 (Sherr, 1996). Cyclin D/cdk triggers the synthesis of cyclin E in late G_1 phase, which in turn activates cdk2, cyclin A production and DNA synthesis.

Phosphorylations are also regulatory, in addition to the synthesis of cyclins. Yet another kinase, CAK, activates the cyclin-dependent kinases. Furthermore, a major role is played by relocalization of cyclin–cdk to the active nuclear compartment within a cell during the cell cycle.

Further investigations revealed yet other proteins whose role is to block activities by binding to cyclin–cdk complexes. These are a family named inhibitors of kinases (INKs). They counterbalance the cyclin's activation of cdks, to affect cycling, development and tumorigenesis (Sherr, 1996). The inhibitory proteins block cyclin D–cdk activities. p27 blocks cell progression, is high in quiescent cells and decreases during late G_1 to release cdk–cyclin activities. Inhibition of cyclins by the cdk inhibitor p21 has often been demonstrated to be induced under various growth-arresting conditions.

In the next step, activated cdks phosphorylate proteins that are essential for progression of the cell cycle. The retinoblastoma tumour suppressor (pRb), absent in retinoblastomas, releases a gene-activating protein named E2F when it is phosphorylated. If this is prevented, E2F is not active, cyclin E is not synthesized, and cells cannot pass through the R point. Additionally, proteasomes' activity of destruction of key inhibitory proteins is vital for passing each checkpoint in the cycle (Koepp *et al.*, 1999). The proteasome is a biochemical machine, composed of protein subunits, that chews up proteins including cyclins

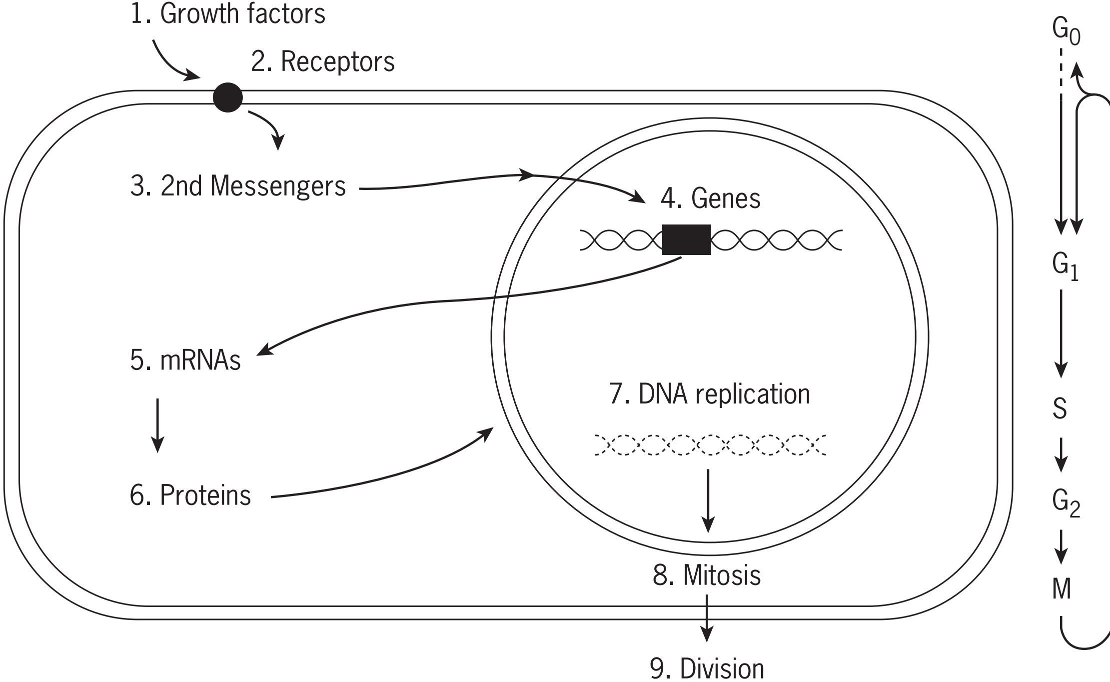


Figure 3 The path to cell proliferation. Growth factors initiate a signalling cascade that takes a cell through the cycle, indicated at the right.

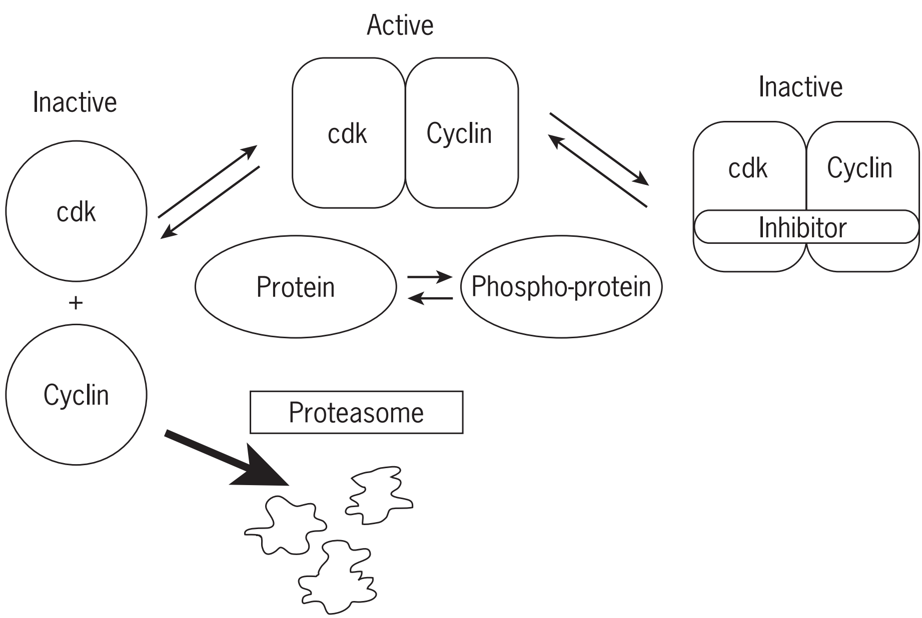


Figure 4 Interactions of cdk with cyclin and inhibitory protein. A cyclin-dependent kinase (cdk) acquires the enzymatic activity to phosphorylate substrate proteins when it binds a cyclin, and this protein complex is inactivated by binding of an inhibitory protein. These reactions are readily reversible and they depend upon phosphorylation of cdk. Cyclins are irreversibly degraded by proteasomes after they are no longer needed.

after they become chemically labelled and targeted for removal.

Unlike peptide growth factors, steroid hormones do not initiate cytoplasmic signalling pathways, but move directly to the nucleus where they activate genes. The sex hormone oestrogen binds to its receptor protein in the nucleus of breast cells, and this in turn binds to and activates growth-stimulating target genes.

Entry into S Phase

Increased cyclin D and E overcome inhibition of cdk activity, and pRb is phosphorylated. This releases E2F and activates genes involved in initiating S phase, including enzymes of DNA synthesis. An example is DNA polymerase, whose transcription is regulated at G_1/S phase by a complex of proteins that contains pRb-like p107, cyclin A and kinase. It is worth noting that most major cell cycle processes are catalysed by large complexes composed of many proteins.

Progression through S phase depends upon cyclin A kinase. Early in S phase, cyclins D and E are degraded by proteasomes. Degradation also removes E2F, which is necessary to prevent programmed cell death (apoptosis) of S phase cells (Lees and Weinberg, 1999).

G_2 Phase and Entry into Mitosis

Mitosis depends upon completion of S phase, and events in G_2 phase are preparatory for it. The complex molecular basis for onset of mitosis was explosively discovered in the early 1980s. Ruderman found that fertilization of oocytes triggers activation of cyclin mRNAs. Hunt discovered that the amounts of cyclin proteins oscillate during the cell cycle, rising during DNA replication and early mitosis and falling at the end of mitosis. Injection of isolated cyclin A into quiescent oocytes drove the cells into M phase. At this time also, Nurse identified the cdk kinase cdc2 as essential for entry into M phase. Unbound cdc2 by itself was inactive. Newly formed cyclin B was shown to bind to and activate cdc2, establishing the first molecular mechanism to explain cell-cycle progression. Then destruction of cyclin B, involving a specialized multi-subunit anaphase-promoting complex, is essential for completion of the cycle.

Research with cell free systems has permitted detailed biochemical investigations of mitosis, showing for example that cyclin B binds to cdc2 (**Figure 2**). This activation of cdc2 kinase is necessary for progression into and through mitosis. The kinase is regulated by a variety of proteins that include cyclin B, phosphatases and kinases and by its subcellular localization. Cyclin B1 begins to accumulate in S phase and increases through G_2. It forms a complex with cdc2, which primes cdc2 phosphorylation. The complex is, however, still inactive, owing to other phosphorylations on cdc2. During G_2 phase, a kinase's (wee1) activity is greater than that of the phosphatase cdc25, and this imbalance keeps cyclin B–cdc2 inactive. At the G_2–M boundary, wee1 is degraded, allowing cdc25 to activate the complex. Furthermore, during G_2 the cyclin B–cdc2 complex resides in the cytoplasm, and at the G_2–M boundary it is rapidly relocated to the nucleus, where it phosphorylates the nuclear membrane protein laminin, which causes the nuclear membrane to break down. Thereafter, chromosomes condense and mitosis proceeds. These many phosphorylations are important for the massive morphological changes that are necessary for a cell to divide.

The Next Cycle – Licensing for DNA Synthesis

A process named licensing permits only one DNA replication per cycle. DNA synthesis cannot be reinitiated until after mitosis is completed. pRb is a critical determinant in preventing DNA reduplication. Perhaps related is the breakdown and reformation during mitosis of the membrane around the nucleus. This permits interaction of molecules from the nucleus and cytoplasm. Degradation of cyclin B by proteasomes is necessary for the start of S phase in the following cycle. Licensing can be disrupted: cells that have lost the cdk inhibitor p21 undergo multiple rounds of DNA synthesis without mitosis, and this process is also activated by anticancer agents. Staurosporin can eliminate the dependence of DNA synthesis on the prior M phase.

Cell Ageing

The normal cell cycle outlined above is modified by various conditions. One of these is cell age. The cycle in early embryo cells is very rapid. It lacks G_1 phase and the corresponding growth-controlling G_1 checkpoint. Mature human cells slow their cycle as they become older, and they cease growing, in G0 or G_1 phase, after about 50 cycles, as initially shown by Hayflick (Baserga). A cdk inhibitor was first discovered in ageing cells by its increase before final arrest of cycling. A progressive shortening of the telomeric DNA, located at the ends of chromosomes, after each cycle is proposed to provide a biological 'clock' for cell ageing. (See chapter on *Telomerase*.)

REGULATION OF CYCLE PHASES

Checkpoints

Entry into and exit from S and M phases are very carefully regulated events. Checkpoint is a name given (Hartwell and Kastan, 1994) to the set of identified cycle-regulatory steps: G_1 restriction point (and the similar START in yeast) and the G_1/S and G_2/M blocks resulting from DNA

damage. Cell-cycle checkpoints are based upon pathways and feedback mechanisms ensuring that a phase of cell cycle does not begin until the preceding phase has been completed with high fidelity. If a checkpoint fails, programmed cell death (apoptosis) or genomic instability ensues. Such failures are important steps in the progression from normal to cancerous cells.

A surveillance system is engaged to make the choice between cell growth and quiescence (Pardee, 1989). When extracellular stimulation by growth factors or nutrients is inadequate, cells cannot pass beyond a specific point in late G_1 phase, in mammalian cells named the restriction point (R). Instead they revert to quiescence (G_0). The final steps that are needed to pass R require synthesis of an unstable protein, proposed to be cyclin E. Under inadequate conditions, synthesis does not keep up with loss, and so this protein cannot be accumulated to an amount sufficient to move the cell into S phase. Once beyond the R point, cells are committed to divide and they no longer require the extracellular growth factors during the remainder of the cell cycle. Restriction point control is defective in cancer cells, and this independence releases cancer cells to continue growing under conditions that keep normal cells in the quiescent state (Pardee, 1989).

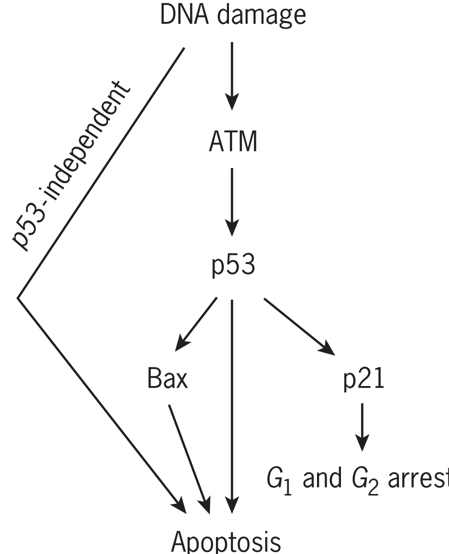


Figure 5 A molecular sequence from DNA damage to apoptosis. A variety of conditions that make a cell unnecessary, such as irreversible damage to it, initiate p53-dependent and independent signalling pathways that lead to apoptosis.

The DNA Damage-induced G_1 Checkpoint

After DNA is damaged, other checkpoint controls delay entry into the next phase of the cell cycle. One such major checkpoint is at the G_1 to S transition, which prevents cells from beginning DNA synthesis until the damaged DNA is repaired. Several proteins, in particular p53, have been implicated in this checkpoint mechanism (**Figure 5**). Individuals who are mutated in the ataxia telangiectasia gene, *ATM*, are very sensitive to X-rays and have a high incidence of tumorigenesis. In response to DNA damage, ATM phosphorylates and increases the level of the p53 protein, a tumour suppressor that is mutated in more than 50% of cancers (Levine, 1997). p53 causes cells to arrest at the G_1–S boundary, which is at least partly due to its production of p21, one of the proteins that inhibits cyclin–cdk complexes.

The DNA Damage-induced S and G_2–M Checkpoints

Within several minutes of exposure to DNA-damaging agents, such as X-rays, mammalian cells in S phase exhibit a dose-dependent reduction in DNA synthesis. Less is known about the mechanism of this S phase checkpoint than about those in G_1 and G_2.

DNA damage also induces a G_2–M checkpoint, as described by Tolmach. This checkpoint delay gives time for DNA repair before the cell goes through mitosis. If repair is not completed in this interval, the cells progress into mitosis without repairing all the DNA damage (Fingert *et al*., 1988), and this results in death or mutations of surviving daughter cells which can thereby become cancerous. This molecular G_2–M checkpoint mechanism is a complex network of phosphorylations and dephosphorylations catalysed by several enzymes and proteins that are moved between cytoplasm and nucleus. Basically, a block in activation of cyclin B–cdc2 prevents the movement of cells into mitosis.

M phase Checkpoints

Mitosis properly segregates chromosomes into the daughter cells. Accurate segregation depends on proper chromosome alignments on and attachment to the mitotic spindle, which is composed of microtubule proteins. A checkpoint ensures that this segregation process occurs correctly. As little as one double strand break in DNA, or depletion of deoxynucleotide building blocks, activates the checkpoint control and stops cells at the G_2–M boundary. This control mechanism delays completion of mitosis until all the chromosomes are attached to the mitotic spindle. The mechanism blocks progression through mitosis if chromosomes are misaligned, and assembly of the microtubules that guide the chromosomes can be inhibited by anticancer drugs such as taxol. Mutations of mitotic checkpoint genes are found in human cancers.

Checkpoints and Programmed Cell Death (Apoptosis)

Apoptosis is a highly regulated process that eliminates physiologically unneeded cells and those that are damaged beyond repair (see the chapter *Apoptosis*). Activated checkpoints give time for a cell to repair its damaged DNA, but if the damage is not soon corrected the cells will initiate apoptosis. This mechanism therefore may prevent the mutations that cause cancer (Sellers and Fisher, 1999). Checkpoint genes, including *p53*, called 'the guardian of the genome,' are involved in causing apoptosis, as is *bax* and other members of the *bcl-2* family (**Figure 5**). Different cells show various responses to damage and drugs, partly because they express different members of the Bcl-2 family (see the chapter *Apoptosis*).

The cyclin A–kinase complex necessary for S phase progression is inhibited when cells are treated with X-rays, and this can result in apoptosis because of the inability of this complex to remove the apoptotic G_1–S factor E2F (Lees and Weinberg, 1999).

Cell Ageing and the Cycle

The elimination of cell ageing is named immortalization. It is an important step in cancer progression, although it does not cause cancer-associated changes (see Hanahan and Weinberg). One way in which human cells can be immortalized is by inserting the gene for the enzyme telomerase, which restores the ageing cell's telomere lengths. Telomerase is also involved in the G_2–M checkpoint.

Cancer is a major cause of death in the elderly. Its incidence increases rapidly, killing about 10% of people between ages 75 and 85 versus 1% between ages 45 and 55. Yet in spite of these epidemiological facts, there is 60% under-representation of cancer patients 65 years or older in treatment trials, few studies have specifically focused on persons over age 65 and many pathological and molecular investigations do not include age as a determinant variable.

CHECKPOINTS, MUTATIONS AND CANCER

The general sequential organisation and duration of the cycle are preserved in cancer, but checkpoint controls are defective (Pardee, 1989; Hartwell and Kastan, 1994). Modifications in cancers are found at many levels of growth regulation, some of which have already been mentioned. The main defect is misregulation of growth initiation at the R point. Furthermore, since checkpoints ensure that mutations are kept low in normal cells, defective checkpoints increase the mutation rate in cancer cells and result in progressive loss of control and emergence of neoplastic disease.

Mutations are causal for cancer; the disease is based upon them. Mutations are found in many genes in advanced cancers. Some of these change cell-cycle controls, including creating a supply of nutrients through angiogenesis, modulating DNA repair, apoptosis, immortalization and metastatic capability (see Hanahan and Weinberg).

The minority of cancer-prone mutations are hereditary. In these cases, a mutated gene on one of a pair of chromosomes is inherited. If, later in life, a mutation occurs of this gene in its partner chromosome, a cancer cell can be produced. Several inherited diseases that are associated with cancer susceptibility have defective checkpoint control. Li–Fraumeni syndrome is a hereditary disease characterized by cancers arising in close relatives. It is a result of a germline mutation in the *p53* gene that abrogates the G_1 checkpoint. Ataxia telangiectasia is characterized by acute cancer predisposition and also other major dysfunctions. Cells from AT patients in culture exhibit severely impaired G_1, S and G_2 checkpoint functions. As mentioned, the *ATM* gene is activated in response to DNA damage and is necessary for activating p53. Another cancer, retinoblastoma, involves mutations in the *Rb* gene, and produces childhood retinal tumours. Survivors have a high risk of developing secondary cancers, particularly osteosarcoma. The BRCA-1 and -2 mutations are associated with hereditary breast cancer; they modify cell cycling and DNA repair. Several other genetic diseases, including Bloom's syndrome, Fanconi anaemia and Nijmegen breakage syndrome, are associated with defects in cell cycle checkpoints and cancer susceptibility.

The majority of cancer-related mutations arise throughout life. For example, cyclin A levels often become abnormally high in cancer cells, and contribute to tumorigenesis. The cyclin D1 and E genes are amplified and over-expressed in many human cancers.

Carcinogenesis can also be caused by viruses such as SV40 and papillomavirus. They introduce their genes that produce proteins that bind to and eliminate the functions of p53 and Rb, thereby bypassing G_1–S, and to a lesser extent G_2, checkpoint controls.

Cancers are often associated with environmental mutagens, such as are produced by smoking. Repeated exposures can produce the several different mutations that are required to cause a cancer. Master mutations can activate growth-promoting oncogenes or loss or inactivation of the tumour-suppressor genes that limit growth. As an example, many cancers have lost or mutated the *p53* gene. One consequence of this mutation is survival of the cancer cell, because p53-dependent checkpoints are eliminated and the programmed cell death mechanism is diminished. Another consequence is that the mutation rate is increased, termed genomic instability (see the chapter *Genomic Instability and DNA Repair*).

The mutations of a half-dozen or more cellular genes is required for tumour formation (Kinzler and Vogelstein, 1996). This number of events is very unlikely in normal cells, whose rate of mutation is approximately 10^{-7} per gene

per cell duplication. Therefore, mutations of genes that increase the overall mutation rate are frequent early events in tumorigenesis. Defects of checkpoint controls in cancers, including hereditary ones, create mutations which are likely to be misrepaired and are progressive because of error-prone repair mechanisms in cancer cells. For example, mutations of BRCA1 or BRCA2 cause *p53* abnormality, which leads to breast cancer (Tseng *et al*., 1997). Other good examples are mutations in colon cancer that cause defective repair of damaged DNA and thereby create genomic instability.

Substances that modify checkpoint controls can change the rate of appearance of mutations, and therefore the progression of cancer. For cells in tissue culture, very high doses of caffeine or related compounds bypass the G_2–M checkpoint, and as a consequence most damaged cells die. These results demonstrate the protective role of the G_2–M checkpoint against damage-induced chromosomal aberrations (Fingert *et al*., 1988). However, chromosomal abnormalities may appear in the few surviving cells.

CANCER THERAPY AND THE CYCLE

Classical Chemotherapies

Currently applied therapies are aimed at killing cancer cells with cytotoxic agents that are applied in combinations. They can prolong the lives of patient with some kinds of cancer but have little effect against others, and all too often the cancer reappears within a few years. One drug provided alone is generally ineffective, because some cancer cells survive this treatment and so the cancer reappears. Multiple drugs are necessary for effectiveness, but this multi-targeting is limited by toxicity to normal cells.

Molecular differences between cancer and normal cells are subtle (see Hanahan and Weinberg). They are mainly related to defective controls of cell growth and survival. Many clinically applied drugs preferentially kill the cycling cancer cells relative to the generally quiescent normal cells, which are essential for survival of the individual. But some kinds of normal cells are cycling, and so the drugs are toxic to the patient. Several cell cycle events provide targets for therapy (**Figure 6**). (See the section on *The Treatment of Human Cancer*.)

Antagonists of Growth Factors

Sex hormones stimulate the growth of some breast, ovarian and prostate cancer cells. Blocking these hormones' action can kill these cells. Tamoxifen is chemically related to the sex hormone oestrogen, with which it competes for binding to oestrogen receptors in a cell. Since tamoxifen blocks the stimulation by oestrogen and does not activate growth, it is in fact inhibitory.

Blocking S Phase

Cornerstones of standard chemotherapy are inhibitors of DNA synthesis. These are small molecule antagonists

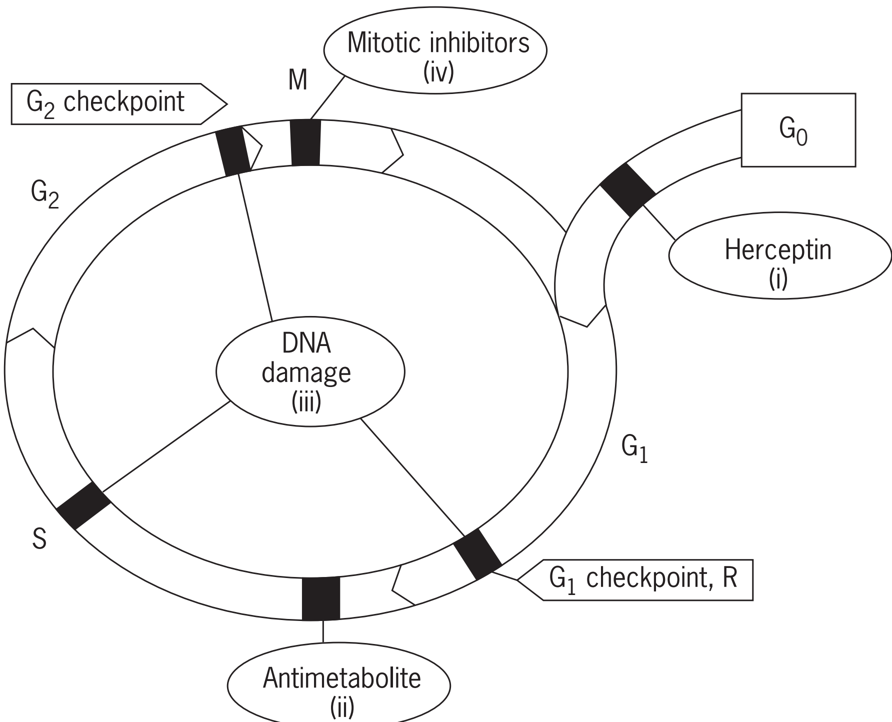


Figure 6 Current therapeutic approaches. Various therapeutic methods that are discussed in the text are summarized here.

structurally similar to metabolic compounds required in the synthesis of DNA and cell survival. As examples, fluorouracil is structurally very similar to uracil, which is needed for DNA synthesis, and methotrexate is an analogue of the vitamin folic acid, also essential for DNA synthesis.

DNA-damaging Agents

Agents that damage DNA are lethal. Examples are X-rays and clinically applied alkylating compounds such as cytoxan and cisplatin. They are more effective against cancer versus noncancer cells because the latter generally are not growing. Also, the normal cells more effectively repair damage during checkpoint delays, before the lethal event of passage of the damaged cell through mitosis and consequent partitioning of damaged DNA between the daughter cells.

Mitotic Inhibitors

Several clinically applied drugs upset the mitotic mechanism in cycling cells, and thereby are lethal. These include taxol from the yew tree and alkaloid toxins from the vinca plant. These currently used compounds, and also experimental epothilones derived from microorganisms, cause lethal mitotic arrest of cycling cells. Their targets are the microtubule proteins, which guide chromosomes through their mitotic separation. Purified plant and microbial products such as these very frequently are starting points for finding anticancer drugs.

A novel drug such as epothilone enters the clinic every few years. Another current example is the antimetabolite gemcytabine, which during S phase is incorporated into newly forming DNA where it arrests continuation of lengthening of the molecule. Difficulties in introducing novel drugs have roots not only in drug discovery, but to a great extent in complex legal requirements for meeting safety standards. These require extremely extensive clinical trials, which with the many costs of doing business, require hundreds of millions of dollars to develop one drug.

Cycle Activators as New Targets

Discovery in cancer cells of over-activated growth-signalling pathways provides possibilities for chemotherapy at every step. Drugs targeted against these reactions are being applied clinically and are in clinical trials.

Tuning Down External Stimuli

Some tumours secrete self-stimulating growth factors into their environment, which also can affect nearby cells. A fascinating example is stimulation by a tumour of the production of new blood vessels, angiogenesis. This process creates the blood supply essential for nourishment of the tumour. Secretion by the tumour of a growth factor VEGF stimulates this production of blood vessel cells, and of blood vessels from their assembly. Both antiangiogenic drugs and also antibodies that neutralize VEGF are being developed as anticancer agents (Boehm-Viswanathan, 2000).

About one third of breast tumour cells lack oestrogen receptors. Unlike those discussed above these are not stimulated by oestrogen or inhibited by tamoxifen, and so they are generally treated with classical anticancer agents following surgery. They are stimulated to grow by EGF, because too many EGF family receptors are on their surface. The monoclonal antibody (herceptin) made against these receptors is effective against some of these cancers, especially when applied in combination with the drugs taxol and doxorubicin.

Targets in the G_1 Phase Signal Transduction Pathway

The molecules that transmit growth signals from a cell's membrane receptors to its nucleus during G_1 phase (**Figure 2**) provide numerous targets for cancer treatment, now under investigation (Adams and Kaelin, 1998; Kaelin, 1999). One major participant is Ras, a small protein that must be positioned against the inner surface of the cell membrane to interact with growth factor receptors. Enzymes must chemically modify Ras for it to occupy this position, and so drugs are being developed that prevent this modification and thereby block the signalling pathway.

Signalling events require numerous kinases that modify the activities of other proteins by addition of phosphates to them. Kinase inhibitors can arrest cell growth and cause death of tumour cells (Shapiro and Harper, 1999). Specific inhibitors of critical cyclin-dependent kinases are being developed. An inhibitor has already demonstrated high efficacy in the treatment of chronic myelogenous leukaemia, a malignancy characterized by the activation of Abl kinase (Drucker and Lydon, 2000).

S Phase Lethality

A cell initiates a sequence of molecular events culminating in apoptosis during S phase unless certain molecules that initiated DNA synthesis are first inactivated. Applying a molecular analogue of part of the G_1–S factor E2F blocks the degradation of E2F and causes apoptosis (Lees and Weinberg, 1999).

Modulating Checkpoints

Several novel potential therapies are being developed (**Figure 7**).

Mitotic catastrophe

The loss of G_2 cell cycle checkpoints can increase tumour-cell sensitivity to chemotherapy. Furthermore, these cells often cannot take refuge at the G_1 checkpoint owing to the loss of p53 or other G_1 checkpoint molecules, whereas

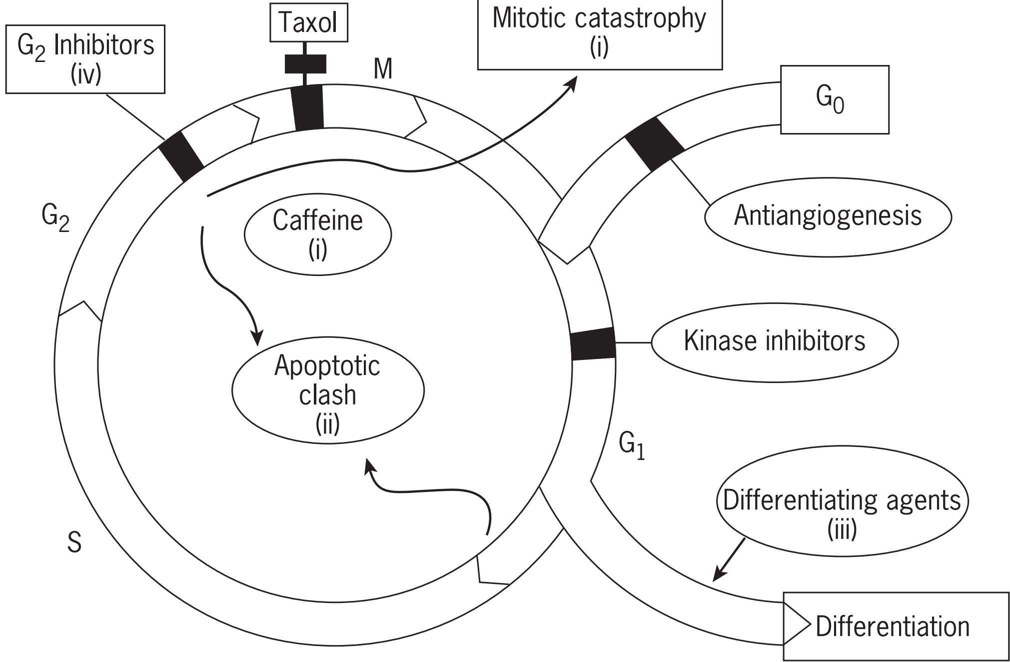


Figure 7 Potential therapies. Several potential therapies described in the text are illustrated.

cells with the normal checkpoint may still recover. Following DNA damage and the arrest at the G_2 checkpoint, some compounds can cause cycling cancer cells to move on through a lethal mitotic cell division. This process was recently aptly renamed 'mitotic catastrophe'. Post-treatment with a caffeine analogue enhanced cytotoxicity of drugs to cancer cells implanted in mice, but not to the mice (Fingert *et al*., 1988). These agents, however, proved to be toxic to humans at doses that abrogate the G_2 checkpoint, which limits their therapeutic use. Other inhibitors that can eliminate the G_2 checkpoint such as the inhibitors of cyclin-dependent kinases, flavopiridol and UCN-0101, are currently undergoing clinical trials (Shapiro and Harper, 1999).

Clash Hypothesis

Remarkably effective synergistic killing of a variety of human cancer cells was found with the combined application together of two molecules derived from plant sources, β-lapachone and taxol (Li *et al*., 1999). These combined drugs killed nearly all of several kinds of cancer cells in culture, at concentrations that did not show major lethality when the drugs were applied singly. Several kinds of human cancer cells growing in mice were destroyed when the drugs were applied together but not if they were applied separately. The tumours did not reappear for at least 2 months. Very importantly, the mice showed no signs of toxicity under these conditions, so there was a very high therapeutic index. Hence there are strong indications of clinical utility of this drug combination.

The mechanism of this powerful synergistic and tumour-specific lethality is being investigated. One hypothesis is based upon the proposal of 'clashing' checkpoint signals; apoptosis is caused by the production of two simultaneous molecular checkpoint signals created by growth conditions (Evan and Littlewood, 1998; Blagosklonny, 1999). Since β-lapachone causes G_1 arrest whereas taxol gives a G_2–M arrest, such clashes of conflicting molecular signals might also be created by this combination of drugs. They could selectively cause apoptotic death of cancer cells which already have defective checkpoint and apoptotic mechanisms.

Differentiation Therapy

An alternative to toxic cancer therapies is to restore the normal cell's properties. Cells in the blood are created by growth and then differentiation of precursor stem cells, followed by R point arrest of the matured cells and their eventual death. Leukaemic cells are mutated blood cells that do not undergo this terminal differentiation and death, and instead they continue to proliferate. Drugs have been discovered that recreate this differentiation-growth arrest process. Retinoids (vitamin A derivatives) are used in this way to treat promyelocytic leukaemia.

Newer drugs such as Saha show potential to inhibit leukaemias through differentiation therapy (Richon *et al*., 1999). This approach is being generalized to make solid tumours differentiate as well. The drugs function by turning on differentiation-related genes, a process that is activated by addition of acetyl groups to the histone

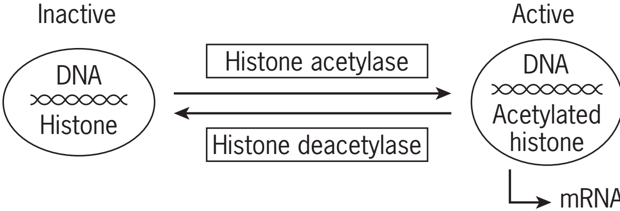


Figure 8 Control by histone acetylation. Histones that surround the DNA in chromatin block activities of genes. Acetylation of the histones permits them to produce their messenger RNAs.

proteins associated with DNA in chromatin (**Figure 8**). This process is catalysed by the enzyme histone acetylase, and it is reversed by deacetylase enzymes, which produce an inactive structure. These changes from the acetylated to the deacetylated state function as an on–off switch for regulation of gene expression. The differentiating drugs shift this balance by blocking deacetylase, thereby the activating acetylation dominates. Thus, drugs that permit histone acetylation, or that decrease the closely connected DNA methylation, are approaches for re-expression of tumour-suppressor genes such as *BRCA1*, *p16* and *p21* that are silenced in cancers by these processes.

Selective Protection of Normal Cells

Chemotherapy of cancer is limited by toxicity to normal cells. With traditional chemotherapy, dose-limiting side effects emerge, including toxicity to bone marrow and gastrointestinal tract, dermatological toxicity and cardiotoxicity. Therefore, selective protection of normal cells against chemotherapeutic drugs could improve the therapeutic index (the ratio of doses that affect cancer versus host), permitting the application of higher drug concentrations (see Blagosklonny and Pardee). Defective checkpoint mechanisms in cancer cells can be the basis of such a selective survival of normal cells (**Figure 6**).

Until recently, the mainstream approach for cancer treatment was directed to finding synergistic combinations in which all the drugs are toxic against cancer cells, and so combinations with an independently inactive drug were considered inappropriate. However, a high therapeutic index, with less toxicity to normal cells, was found in clinical trials when drugs that blocked entry of normal cells into M phase were administered before subsequently adding taxol. This prevented taxol lethality in M phase. This antagonism was translated in the clinic as a decrease in side effects to normal cells.

G_1 checkpoint arrest in normal but not in cancer cells is produced by low, nonlethal concentrations of compounds such as cycloheximide that slow protein synthesis. Later addition of a toxic S-phase-specific agent cannot kill them. In contrast, independence of cancer cells from this cycle arrest causes them to enter S phase, where they are killed while they are synthesizing DNA. This idea, proposed in 1975, has been revived in a new form. Low doses of doxorubicin or etoposide induce p53- and p21-dependent growth arrest of normal cells without cell death, but these drugs do not arrest cancer cells. This pretreatment thereby abolished the cytotoxicity otherwise caused by later addition of microtubule-active drugs (paclitaxel, vincristine, epothilones). Protection of cells with normal checkpoint was achieved, whereas no protection was observed in cancer cells lacking p53 or p21 (see Blagosklonny and Pardee).

Novel inhibitors of the cell cycle are being developed as lethal drugs against solid cancers and leukaemia cells. However, these are active also against normal cells; proliferating bone marrow and epithelial cells are particularly vulnerable. Thus, searches for compounds that reversibly inhibit proliferation of these normal cells will be especially valuable to protect the individual. For example, although two compounds similarly inhibited protein kinase C, UCN-01 was selected as the drug to develop because it had higher cytotoxicity to cancer. For a selective growth arrest of normal cells, dependent on protein kinase C, one would choose the less toxic inhibitor GF109203X. Other strategies utilizing the retention of checkpoints in normal cells to protect them versus tumour cells are discussed (see Blagosklonny and Pardee). Since proliferation of normal cells is highly regulated, the search for such inhibitors should produce surprises.

In summary, defects in cancers of various molecular mechanisms that control cell growth, differentiation and apoptosis have recently been discovered. These differences from normal cells provide novel targets for therapy, some of which are being developed and tested.

ACKNOWLEDGEMENT

This work was supported by NIH Grant CA RO1 61253.

REFERENCES

Adams, P. D. and Kaelin, W. G., Jr (1998). Negative control elements of the cell cycle in human tumors. *Current Opinion in Cell Biology*, **10**, 791–797.

Blagosklonny, M. V. (1999). A node between proliferation, apoptosis, and growth arrest. *Bioessays*, **21**, 704–709.

Boehm-Viswanathan, T. (2000). Is angiogenesis inhibition the Holy Grail of cancer therapy? *Current Opinion in Oncology*, **12**, 89–94.

Druker, B. J. and Lydon, N. B. (2000). Lessons learned from the development of an Abl tyrosine inhibitor for chronic myelogenous leukemia. *Journal of Clinical Investigation*, **105**, 3–7.

Evan, G. and Littlewood, T. (1998). A matter of life and cell death. *Science*, **281**, 1317–1322.

Fingert, H. J., *et al*. (1988). *In vivo* and *in vitro* enhanced antitumor effects by pentoxifylline in human cancer cells treated with thiotepa. *Cancer Research*, **48**, 4375–4381.

Hartwell, L. H. and Kastan, M. B. (1994). Cell cycle control and cancer. *Science*, **266**, 1821–1828.

Howard, A. and Pelc, S. R. (1951). Nuclear Incorporation of P32 as demonstrated by autoradiography. *Experimental Cell Research*, **2**, 178–187.

Kaelin, W. G., Jr (1999). Choosing anticancer drug targets in the postgenomic era. *Journal of Clinical Investigation*, **104**, 1503–1506.

Kinzler, K. W. and Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell*, **87**, 159–170.

Koepp, D. M., *et al*. (1999). How the cyclin became a cyclin: regulated proteolysis in the cell cycle. *Cell*, **97**, 431–434.

Lees, J. A. and Weinberg, R. A. (1999). Tossing monkey wrenches into the clock: new ways of treating cancer. *Proceedings of the National Academy of Sciences of the USA*, **96**, 4221–4223.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323–331.

Li, C. J., *et al*. (1999). Potent inhibition of tumor survival *in vivo* by β-lapachone plus taxol: combining drugs imposes different artifical checkpoints. *Proceedings of the National Academy of Sciences of the USA*, **96**, 13369–13374.

Minshull, J., *et al*. (1989). The role of cyclin synthesis, modification and destruction in the control of cell division. *Journal of Cell Science, Supplement*, **12**, 77–97.

Nurse, P., *et al*. (1998). Understanding the cell cycle. *Nature Medicine*, **4**, 11030–1106.

Pardee, A. B. (1989). G_1 events and regulation of cell proliferation. *Science*, **246**, 603–608.

Richon, V. M., *et al*. (1999). A class of hybrid polar inducers of transformed cell differentiation inhibits histone deactylase. *Proceedings of the National Academy of Sciences of the USA*, **95**, 3003–3007.

Roberts, J. M. (1999). Evolving ideas about cyclins. *Cell*, **97**, 129–132.

Shapiro, G. I. and Harper, J. W. (1999). Anticancer drug targets: cell cycle and checkpoint control. *Journal of Clinical Investigation*, **104**, 1645–1653.

Sherr, C. J. (1996). Cancer cell cycles. *Science*, **274**, 1672–1677.

Tseng, S. L., *et al*. (1997). Allelic loss at BRCA1, BRCA2, and adjacent loci in relation to TP53 abnormality in breast cancer. *Genes Chromosomes Cancer*, **20**, 377–382.

FURTHER READING

Andreeff, M., *et al*. (2000). Cell proliferation, differentiation and apoptosis. In: Holland, J., *et al*. (eds), *Cancer Medicine*, 5th edn. 17–32. (B. C. Decker, Hamilton, Ontario).

Baserga, R. (1985). *The Biology of Cell Reproduction.* (Harvard University Press, Cambridge, MA).

Blagosklonny, M. V. and Pardee, A. B. (2001). Exploiting cancer cell cycling for selective protection of normal cells. *Cancer Research*, **11**, 4301–4305.

Ford, H. L. and Pardee, A. B. (1999). Cancer and the cell cycle. *Journal of Cellular Biochemistry*, **75**, 166–172.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Murray, A. W. and Hunt, T. (1993). *The Cell Cycle, an Introduction.* (Freeman, San Francisco).

Sellers, W. R. and Fisher, D. E. (1999). Apoptosis and cancer drug targeting. *Journal of Clinical Investigation*, **104**, 1655–1661.

Chapter 3

Overview of Oncogenesis

Julie L. Boerner, Jacqueline S. Biscardi and Sarah J. Parsons
University of Virginia, Charlottesville, VA, USA

CONTENTS

- Properties of Neoplastic Cells
- Oncogenes
- Tumour Suppressors
- Molecular Mechanisms of Cancer
- Clinical Correlations

PROPERTIES OF NEOPLASTIC CELLS

Normal cells are exquisitely attuned to their environment and respond to external cues via tightly regulated signalling pathways that either trigger or repress growth. In order for a cell to undergo mitogenesis, a growth-promoting signal from the extracellular environment must first initiate a cascade of events within the cell that results in activation of genes that stimulate cell division. With few exceptions, most of the cell populations within an adult organism are terminally differentiated and no longer proliferate. Cancer arises when a cell, for a variety of reasons, escapes the normal constraints placed on its growth and begins to divide in an unregulated fashion.

Factors that Promote Growth

Extracellular factors that stimulate growth include peptide growth factors such as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF), which bind tyrosine kinase receptors located on the cell surface. Cytokines such as growth hormone, interleukins and prolactin also bind cell surface receptors which are not kinases themselves but are able to transduce their signals via interaction with separate tyrosine kinase molecules. Another class of growth factors bind serpentine receptors that couple to intracellular pathways via heterotrimeric G proteins. Lastly, steroid hormones such as oestrogen, which bind intracellular receptors, also have mitotic activity. All classes of receptors are capable of triggering a cascade of signalling events culminating in mitotic activity, or proliferation, of the target cell.

Mutation of Growth Regulatory Genes in Cancer

In tumour cells, molecules that regulate signalling pathways which stimulate growth are often mutated, resulting in a constant 'on' signal to the cell. These molecules can be positive for growth regulation, as cancer-causing oncogenes, or negative for growth regulation, as protective tumour-suppressor genes. In addition, tumour cells often develop their own autocrine loops for growth, wherein a growth factor required for the activation of a pro-growth signalling pathway is constantly produced and secreted into the extracellular milieu. One example of such an autocrine loop is found in breast tumour cells, which are able to produce the growth factor TGF-α. TGF-α binds and activates the EGF receptor, thereby triggering mitotic pathways within the cell.

Cell Death

Apoptosis, or programmed cell death, is yet another process that is subverted by a tumour cell (Jacotot *et al.*, 2000). In a normal cell, a series of 'checkpoints' must be met before the cell permits itself to divide. If irreparable damage to its DNA is present, the cell undergoes apoptosis, thus ensuring that its mutated DNA is not transmitted to progeny cells. The molecules that regulate this process of apoptosis are often themselves mutated in cancer cells, which are then able to escape the checks and balances that a normal cell must undergo before it can divide. (See chapter on *Apoptosis*.)

Cell–Cell Interaction

Cells can also respond mitogenically to cues from other cells. Normal cells are growth-inhibited by contact with other cells and form a monolayer when grown in culture. Cancer cells, on the other hand, form foci, or piled-up accumulations of constantly dividing cells; foci result as a consequence of loss of contact inhibition. Molecules called cellular adhesion molecules (CAMs) and cadherins are expressed on the surface of cells and negatively regulate growth. Cadherin molecules on adjacent cells bind one another in a calcium-dependent manner; this binding prevents cells from entering the mitotic cycle (Christofiori

and Semb, 1999). Further intracellular signalling occurs via the catenin family of molecules, which link the cadherins to the cytoskeleton and to the transcription machinery. The negative regulation normally provided by the interaction of these molecules is frequently lost in tumour cells. For example, epithelial cell cadherin (E-cadherin) is mutated, absent, or reduced in expression in a variety of human tumours. In cell culture systems, loss of the E-cadherin gene leads to loss of cell–cell contacts and increases in cell motility and invasiveness. Aberrant phosphorylation of the catenins can lead to loss of proper cell–cell contacts, which is thought to be a step in the acquisition of invasive properties of the cancer cell. Moreover, the *APC* tumour-suppressor gene, which is mutated in human cancer, is known to associate with β-catenin. Cancer-causing mutations in *APC* involve the portion of the molecule that binds β-catenin. Thus, loss of these tumour-suppressor genes and their appropriate interactions with cadherins and catenins relieves the constraints of contact inhibition, a hallmark of tumour cells.

Cell–Substratum Interactions

Pro-growth cues can also come from the extracellular matrix or substratum (such as the basement membrane) on which cells grow (Miyamoto *et al*., 1998). Proteins such as fibronectin, a component of the ECM, bind integrin receptors on the cell surface. The integrin receptors then cooperate with growth factor receptors to trigger mitogenic pathways via activation of signalling cascades involving several different kinase molecules. Alternative, tumour-specific isoforms and unique combinations of integrins are often present in tumour cells, thereby providing additional means by which growth signals can be initiated. (See also chapter *Extracellular Matrix: The Networking Solution*.)

Angiogenesis

Tumours also exhibit extensive vascularization which increases as the tumour grows (Folkman, 1992). This outgrowth of new blood vessels is termed angiogenesis and is not seen in normal adult animals except in the cases of wound healing and pregnancy, where new tissues such as placenta are formed. In the absence of new blood vessels, a tumour is able to grow to a maximum size of approximately 1 mm in diameter, the distance that oxygen and nutrients are able to diffuse into the tumour (Kurschat and Mauch, 2000). Vascularization thus allows the tumour to grow larger. (See also chapter on *Angiogenesis*.)

Migration and Metastasis

Tumour cells often have the ability to migrate away from the original tumour site and grow in distant parts of the body (Kurschat and Mauch, 2000). This ability to metastasize requires that the tumour cells produce proteases that degrade the basement membrane of blood vessels through which the tumour cells will travel. Metastatic tumour cells are also able to escape immune surveillance of the host organism and then to grow again in another part of the body. Factors produced at these distal sites are thought to provide a favourable environment for the continued proliferation of the tumour cells. Evidence for this comes from the finding that specific types of tumours have a propensity to metastasize to the same sites, e.g. prostate cancer commonly targets bone. Moreover, tumour cells also have the ability to affect the underlying layer of stroma, or fibroblast cells, stimulating them to produce growth factors and cytokines that enhance tumour growth. **Figure 1** depicts how a cell breaks away from the primary tumour and generates a secondary neoplasm within the organism. (See also chapter on *Invasion and Metastasis*.)

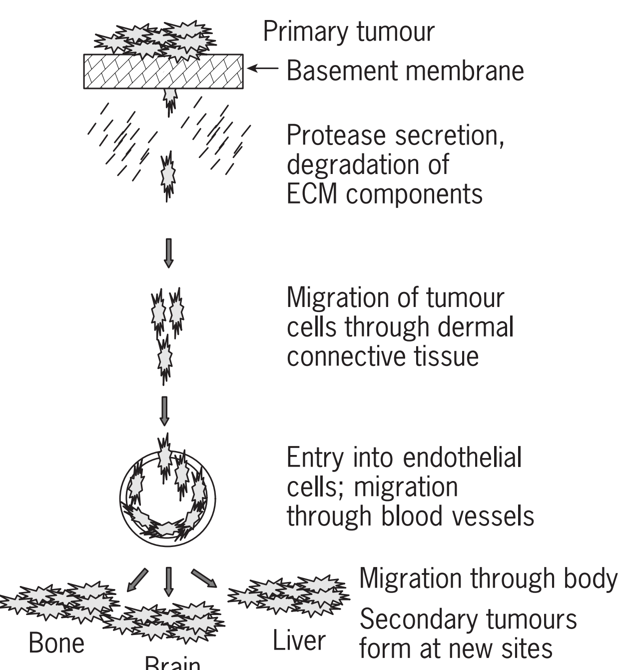


Figure 1 Metastatic progression. Individual cells within the primary tumour upregulate production of specific proteases, which gives the cell the ability to degrade ECM or basement membrane. Cells then break away from the primary tumour and begin to migrate. Migrating cells adhere to connective tissue and invade dermal tissue. Cells thus enter the vascular system by migrating between endothelial cells and moving through blood vessels. Lastly, tumour cells reach secondary sites where conditions are favourable for their continued growth.

Conversion of a Normal Cell to a Tumour Cell

A cell becomes converted from a normal to a neoplastic cancer cell when the regulation of one or more of the above

processes is lost. Loss of regulation occurs when mutations arise in two broad families of genes that regulate growth: oncogenes, which act as positive signals for growth, and tumour-suppressor genes, which act as brakes or checkpoints on a cell's progression through the cell cycle. These mutations may be caused by environmental, chemical or biological agents or events that result in irreversible alterations in the genome of a cell, so that progeny cells also carry the same mutations that allow for uncontrolled growth. This is the first step on a pathway that can eventually lead to an aggressive, metastatic tumour. Fortunately, organisms possess several means of dealing with environmental insults and genetic alterations. More than one genetic 'hit,' or error, is required before an actual tumour is able to arise, as will be discussed later.

ONCOGENES

Historical Perspective

In 1911, Peyton Rous laid the groundwork for the oncogene theory of cancer, a theory that became the basis for all modern cellular signalling and genetic research. He identified a spindle-celled sarcoma in chickens that was transplantable from one bird to another, using a filtrate of the tumour (Rous, 1911; Weiss *et al*., 1985). The infectious agent responsible for the tumours was later found to be the Rous sarcoma virus (RSV). Thus, a cancer causing agent had been discovered, but the means by which the virus induced tumours was still unclear.

Further insight into the process of oncogenesis was provided in 1914, by Theodor Boveri who hypothesized that cells in cancer tissue contain defective chromosomes. Working with double-fertilized sea urchin embryos, Boveri observed that the resulting aberrant chromosomes were passed on to progeny cells (Boveri, 1925). This finding led him to believe that cells of malignant tumours have damaged chromosomes and that a neoplastic cell can arise from a normal cell and pass its altered genome on to progeny cells. Thus, a tumour cell is in some way defective and has lost the properties of a normal cell. Environmental insults were also implicated in causing cancers. In 1918, Yamagiwa and Ichikawa showed that continual irritation of the normal epithelium of rabbit ears caused papilloma-like growths and metastasis.

Building on the early work of Rous, Shope provided further evidence for the viral basis of oncogenesis by his demonstration that a papilloma-like growth was transmissible from animal to animal (Weiss *et al*., 1985). In 1951, Gross showed that mice inoculated with leukaemic extracts developed neoplasms. From these extracts, the Gross murine leukaemia virus was isolated (Weiss *et al*., 1985). Seven years later, Temin and Rubin showed that infection of cultured chicken fibroblasts with the Rous sarcoma virus caused neoplastic transformation of the cells (Weiss *et al*., 1985). Martin and others later identified the oncogenic portion of the RSV genome as v-*src*, the viral *src* oncogene. These early results suggested a transmissible mechanism for tumour initiation. Thus, as early as the beginning of the twentieth century, a cellular/genetic model of oncogenesis had been postulated.

The Oncogene Hypothesis

The best-known theory of oncogenesis, however, is a relatively recent one. In 1982, Bishop and Varmus hypothesized that cancer-causing genes, or oncogenes, that are carried by tumour-inducing viruses have normal counterparts that are present in the genomes of all vertebrate cells (Bishop, 1982). These normal genes are termed proto-oncogenes. Evidence for this hypothesis came from hybridization studies, where radiolabelled v-*src* DNA was found to bind, or hybridize, to its complementary counterpart (c-*src*) in normal avian cellular DNA. The v-*src* and c-*src* genes encode a tyrosine kinase, an enzyme that transfers phosphate from ATP to the amino acid tyrosine found in cellular proteins. These phosphorylations have profound effects on cell growth. Similar studies eventually led to the identification of a family of viral oncogenes, which can be transmitted by either DNA or RNA viruses. DNA viruses either can cause lytic infection leading to the death of the cellular host, or can replicate their DNA along with that of the host genome and promote neoplastic transformation of the cell. DNA viruses encode various proteins which, along with environmental and genetic factors, help to initiate and maintain the neoplastic state. RNA tumour viruses, on the other hand, integrate DNA copies of their RNA genomes into the genome of the host cell. Since the viral genomes contain transforming oncogenes, they induce cancerous transformation of the host cell.

Mechanism of Acquisition of Cellular Sequences by RNA Tumour Viruses

Multiple lines of evidence indicate that viral oncogenes arise when an RNA virus integrates its genome near the coding sequence for a proto-oncogene and incorporates the proto-oncogene's DNA into its own genetic material during the virus replication cycle. Through multiple rounds of infection and genome replication, deletions and other mutations occur in the proto-oncogene, conferring on the gene tumorigenic properties. Ensuing infection of a normal cell by an RNA tumour virus carrying such an oncogene causes malignant transformation of that cell. Although this process rarely occurs in human tumours, many of the same

Table 1 Examples of human oncogenes

Oncogene	Originally identified in	Mechanism of activation in human tumours	Location	Associated human cancers
src	Rous sarcoma virus	Overexpression, C-terminal deletion	Cytoplasmic	Breast, colon, lung carcinomas
myc	Avian myelocytomatosis virus	Translocation	Nuclear	Burkitt lymphoma
abl	Abelson murine leukaemia virus	Translocation	Cytoplasmic	Chronic myeloid leukaemia
Ha-*ras*	Harvey murine sarcoma virus	Point mutation	Cytoplasmic	Bladder cancer
K-*ras*	Kirsten murine sarcoma virus	Point mutation	Cytoplasmic	Colon, lung carcinomas
erbB (EGFR)	Avian erythroblastosis virus	Overexpression, deletion	Cytoplasmic	Breast carcinoma, glioblastoma

genes 'captured' by animal retroviruses are altered in human cancers. These alterations take the form of base pair changes, insertions, amplifications and translocations, which result in a protein product that no longer responds normally to growth-regulatory cues. Only one allele of the gene needs to be mutated for the oncogenic effect. Thus, oncogenes are described as carrying dominant mutations. For example, Ras in human tumours is often found to be mutated at a single amino acid residue (Wittinghofer, 1998), whereas the oncogene *abl* is activated by chromosomal translocation (Heisterkamp *et al*., 1985). In the case of c-Src, a negative regulatory site present in normal c-*src* is mutated in a small subset of colon cancers, thus rendering the protein constitutively active (Irby *et al*., 1999), whereas in human breast cancer, overexpression of the normal c-Src protein appears to play a role in the deregulation of cell growth (Biscardi *et al*., 1999).

Proto-oncogenes can be classified as either cytoplasmic or nuclear, depending on where in the cell they are localized. Many of the cytoplasmic proto-oncogenes code for tyrosine kinase molecules, enzymes that are able to phosphorylate substrate proteins on tyrosine residues and that are known to be essential for controlling the signalling cascades that regulate mitosis. Others, such as Ras, transmit cellular growth signals by binding guanine nucleotides in the form of GTP or GDP. Ras is often found mutated at single sites such that it is constantly bound to GTP, which causes the molecule to be constitutively active. Mutations in Ras are found in approximately 30% of human cancers (Wittinghofer, 1998). Serine–threonine kinases, such as the Raf family of kinases, are the targets of Ras and constitute another family of proto-oncogenes that regulate proliferation. Nuclear oncogenes such as *myc* regulate gene transcription. **Table 1** lists a few examples of the better-known oncogenes, their subcellular localization and mechanism of oncogenic activation. Although these oncogenes are defined as cancer causing genes, it is important to note that the introduction of a single activated oncogene into a cell does not result in neoplastic transformation. At least two active oncogenes, or an activated oncogene and an inactivated tumour suppressor, are required for tumour formation.

TUMOUR SUPPRESSORS

Tumour-suppressor genes are defined as recessive genes, i.e. they must sustain mutations or deletions of both alleles in order to contribute to cancer formation and progression. This definition implies that one functional allele of the tumour-suppressor gene is sufficient for normal cell function. Patients with familial cancers frequently inherit one normal and one abnormal allele of the tumour-suppressor gene from their parents. If the second, normal allele is lost, the protective effect of the gene product no longer exists. Therefore, introduction of a wild-type copy of the gene back into the tumour should inhibit further tumour growth. Unfortunately, putative tumour-suppressor genes shown to be inactivated in cancer are not sufficient by themselves to restore normal cell function. Thus, whether such genes are actually tumour suppressors remains a debated question.

Discovery and Identification

The origin of the concept of tumour-suppressor genes (or anti-oncogenes) came from cell fusion studies dating back to the early 1900s. These studies revealed that when one tumour cell is fused with another and the fused product is introduced into mice, tumour formation results (Sager, 1989). However, when a tumour cell is fused with a normal cell and introduced into mice, the fusion blocks tumour formation. These observations suggested that some activity must be present in the normal cell that inhibits transformation.

Retinoblastoma (Rb) Gene

The first tumour-suppressor gene identified was the *Rb* gene, which is associated with the childhood illness of retinoblastoma (Knudson, 1971). In an epidemiological study, Knudson and colleagues noticed that bilateral retinoblastoma occurred frequently within the same family, whereas unilateral retinoblastoma did not appear to be a

genetically inherited disease. In families with bilateral retinoblastoma, karyotyping techniques were used to detect homozygous loss of chromosome 13q, a defect that was transmitted to offspring. Homozygous loss was found to be necessary but not sufficient for the formation of retinoblastoma, since not every family member with the loss of both alleles developed the disease. Later, the gene responsible for development of the disease was cloned and termed *Rb* for retinoblastoma. Reintroduction of this gene into cultured retinoblastoma tumour cells reversed the malignant phenotype, suggesting that the gene was indeed a tumour suppressor (Bookstein *et al*., 1989).

Tumour Suppressors in Colon Cancer

Since the cloning of *Rb*, many other tumour-suppressor genes have been identified (**Table 2**). Several of the most notable are a group of tumour-suppressor genes that were identified by studying progressive stages of colon cancer. They include the 'adenomatous polyposis coli' (*APC*) gene, the 'deleted in colon cancer' (*DCC*) gene and the 'mutated in colon cancer' (*MCC*) gene (Peddanna *et al*., 1996). *APC* maps to chromosome 5q21 and is mutated in 70% of patients with a hereditary form of colon cancer, termed familial adenomatous polyposis (FAP). Also mapping to chromosome 5q is the *MCC* gene, which was found to be mutated in 55% of all colon cancers studied. *DCC* was mapped to chromosome 18 and is deleted in 73% of colon cancers. *APC* and *DCC* code for proteins that play roles in regulating cell adhesion in normal cells. It is speculated that loss of these genes can lead to increases in cell motility, a key characteristic of metastasis.

p53

The p53 protein is involved in sensing DNA damage and regulating cell death (Marx, 1993). In normal cells, when DNA damage is sensed by p53, the cell cycle is arrested to permit DNA repair. Upon completion of this process, the cell progresses through the mitotic cycle. If repair fails to occur, p53 initiates the process of apoptosis, or programmed cell death. Thus, normal cells with genetic defects die. If p53 is not present in the cell (via gene deletion) or is mutated to be nonfunctional, DNA damage is not repaired, and the cell progresses through the cell cycle, transmitting its damaged DNA to its progeny. p53 is so important to the maintenance of 'healthy' DNA that it is mutated or deleted in over 70% of human cancers, including osteosarcomas, rhabdomyosarcomas and carcinomas of the breast, colon, lung and prostate.

BRCA1 and *BRCA2*

Another more recently identified tumour-suppressor gene, *BRCA1*, was found to be linked to an increased risk of hereditary breast cancer (Zheng *et al*., 2000). Loss of chromosome 17q had long been known to occur in familial breast cancer. The *BRCA1* gene mapped to chromosome 17q, but it was not until 1993 that it was identified and cloned. Many heritable mutations were identified in *BRCA1* from breast cancer patients and include an 11-bp deletion, a 1-bp insertion, a stop codon and a missense substitution. However, this may be an underestimation of its involvement in oncogenesis, as mutations and inactivating events, such as promoter methylation, also may regulate *BRCA1*

Table 2 Tumour-suppressor genes and their function and associated cancers

Name	Function in normal cells	Associated cancers
p53	Cell cycle regulator	Colon and others
BRCA1	Cell cycle regulator, genomic integrity and chromatin structure	Breast, ovarian, prostate and others
BRCA2	Genomic integrity	Breast, ovarian, prostate and others
PTEN	Tyrosine and lipid phosphatase	Prostate, glioblastomas
APC	Cell adhesion	Colon
DCC	Cell adhesion	Colon
MCC	Undetermined	Colon
p16-INK4A	Cell cycle regulator	Colon and others
MLH1	Mismatch repair	Colon and gastric cancers
MSH2	Mismatch repair	Colon and gastric cancers
DPC4	Cell death regulator	Pancreatic
Wt1	Cell death regulator	Wilms' tumour
NF1	Regulator of GTPases	Astrocytomas
NF2	Cell adhesion	Astrocytomas
VHL	Ubiquitination	Renal
PTC	Regulator of hedgehog signalling	Thyroid
TSC2	Cell cycle regulator	Breast and renal
TSG101	Cell cycle regulator	Renal and leukaemia

expression. Such gene regulation events are still being defined and are difficult to identify by screening techniques. Another *BRCA* family member, *BRCA2*, also has been cloned. This gene localizes to 13p12–13, and mutations within it correlate with breast cancer occurrence. *BRCA1* and 2 also are mutated or deleted in about 33 and 34% of sporadic breast tumours, respectively.

PTEN

PTEN, a gene encoding a phosphoprotein and phospholipid phosphatase, was first identified in glioblastoma patients who had sustained deletions of chromosome 10q23 (Li *et al*., 1997). *PTEN* is mutated in 31% of glioblastomas, 100% of prostate cancers and 6% of breast cancers. Interestingly, deletion of *PTEN* in gliomas segregates independently of mutations in p53, i.e. tumours containing *PTEN* mutations do not contain p53 mutations (Liu *et al*., 1997). However, *PTEN* deletions/mutations do correlate with amplification of the EGF receptor, a known oncogene. In normal cells, it is thought that PTEN downregulates phosphorylation events that promote cell growth. Its loss, therefore, allows for unregulated and unhindered proliferation.

Other Tumour-suppressor Genes

Another tumour-suppressor gene is the *p16-INK4A* gene, which negatively regulates cell cycle events. It is lost from chromosome 9 in a wide range of cancers (Kamb *et al*., 1994). Genes involved in the efficacy of DNA replication, *MLH1* and *MSH2*, are found deleted in 50% of hereditary non-polyposis colorectal cancers (Konishi *et al*., 1996). *DPC4* (deleted in pancreatic cancer) is lost from chromosome 18q in pancreatic cancer (Hahn *et al*., 1996). Still other tumour-suppressor genes include the Wilms' tumour-associated tumour suppressor *Wt1*, the human astrocytoma-associated tumour suppressors *NF1* and *NF2*, the von Hippel–Lindau syndrome tumour suppressor *VHL*, the papillary thyroid cancer tumour-suppressor gene *PTC* and tumour-suppressor genes associated with breast and renal cancer, *TSG101* and *TSC2*, respectively. Each of these genes encode protein products that negatively regulate the acquisition of a malignant phenotype by a normal cell.

Tumour-suppressor Genes, Normal Cellular Function and Carcinogenesis

In normal cells, products of tumour-suppressor genes have been shown to regulate negatively cell growth and proliferation. For example, the *Rb* gene product sequesters transcription factors that are required for normal cell cycle progression. The ability of Rb to function as a block to cell cycle progression is regulated by phosphorylation of the Rb protein on multiple serine residues (Harbour and Dean, 2000). In quiescence, Rb is hypo- or under-phosphorylated and binds members of the E2F transcription factor family. Upon growth factor stimulation of the cell, Rb becomes phosphorylated by cyclin-regulated kinases and releases E2F, which then induces gene transcription events necessary for cell division. In cancers, deletion or inactivation of Rb results in constitutively 'free' E2F, which in turn leads to unfettered gene transcription and oncogenic transformation. Similarly, as described above, p53 has been shown to sense DNA damage, cause cell cycle arrest, regulate transcription and stimulate apoptotic cell death pathways in normal cells (Marx, 1993). Loss of this function increases the chance of damaged DNA being transmitted to subsequent generations of cells. The exact role of BRCA1 is still unclear, but studies using mice that lack the *BRCA1* gene show that it is essential for cellular proliferation during early embryonic development (Zheng *et al*., 2000). BRCA1 may also regulate transcriptional events, since it is capable of acting as a coactivator of p53 and a corepressor of c-Myc. Recent studies also implicate a role for BRCA1 in chromatin remodelling, which is required for DNA transcriptional and replication events. PTEN regulates the phosphorylation status of phospholipids that are involved in regulating apoptotic pathways within the cell (Di Cristofano and Pandolfi, 2000). Taken together, these findings indicate that tumour-suppressor gene products act by negatively controlling cell growth in normal cells and that their loss contributes to the unregulated cell growth seen in tumour cells.

Mismatch Repair Genes

Critical regulators of genomic integrity, as exemplified by mismatch repair genes, also have been implicated as tumour-suppressor genes. The microsatellite instability genes described above, *MLH1* and *MSH2*, are important to the maintenance of genomic integrity by repairing mismatched base pairs that arise with a stable frequency during DNA replication (Kolodner and Marsischky, 1999). Mismatched base pairs are recognized and cleared by mismatch repair enzymes, and new bases are added in their place. Without such genes, repairs are not made and mutations are introduced into newly synthesized DNA. Alternatively, the stress of the mismatch structure may fragment the DNA. Both of these possibilities can lead to changes in the sequence of genes critical to cell growth or death. Although the alteration of mismatch repair genes may seem like a key event for all cancers, it has been determined that only 13% of gastric/colorectal cancers and less than 2% of other cancers have mutations in mismatch repair genes. Furthermore, the 2% occurrence is thought to reflect the normal rate of DNA mutation. These considerations therefore suggest that mismatch repair defects may be more specific for gastric cancers and not a general phenomenon associated with cancer development.

New Techniques for Identification of Tumour-suppressor Genes

RFLP

Knudson's original method of analysing karyotypes of tumour cells is still in use today for examining large, consistent chromosomal alterations, but new methods have evolved in the last 20 years. One such technique, restriction fragment length polymorphism (RFLP), utilizes bacterial restriction enzymes that cleave DNA at specific sites. DNA encoding a normal gene has a characteristic DNA fragment pattern, while tumour DNA shows an abnormal pattern. RFLP DNA fragments have been linked together to span an entire normal human genome. At a frequency of approximately every 10 million base pairs, a known gene has been mapped to specific RFLP fragments. This approach has yielded a crude map of the genome, which is more sensitive than karyotyping methods and allows one to map a loss of specific regions of a chromosome.

Comparative Genomic Hybridization (CGH)

Comparative genomic hybridization compares the ability of tumour RNA labelled with one fluorophore (i.e. green) and normal RNA labelled with another fluorophore (i.e. red) to hybridize to a chromosome spread from a given tissue type. The spread is analysed by fluorescence microscopy after hybridization of the RNAs. Losses (or gains) in a chromosome can be observed by the colour of the fluorophore hybridized to the region on the chromosome. For example, if a region of chromosome 13 is deleted in the tumour DNA, chromosome 13 will appear red since there would be no green-labelled (tumour-derived) RNA to hybridize with that region. Gains in the tumour DNA score green whereas equal expression in both normal and tumour cells scores yellow (a merging of red and green). Thus a complete genomic map of genetic changes that occur in a tumour cell can be obtained. Other more sensitive techniques, such as DNA microarrays, are currently being perfected to identify additional tumour-suppressor genes. It should be noted that in comparison with oncogenes, a very small number of tumour-suppressor genes have been discovered. The techniques that detect tumour-suppressor genes are somewhat insensitive, since even the most accurate screening approaches localize the region of loss only to a megabase or more.

Methylation

The techniques described above tend to identify tumour-suppressor genes that are grossly mutated in cancer. However, there are other mechanisms of tumour-suppressor gene inactivation, including point substitutions, small insertions and deletions. Almost half of all tumour-suppressor genes are also methylated in their promoter regions, preventing gene transcription (Baylin, 1997). Abnormally high levels of methylation appear in cancer cells that have a loss in the *p21/WAF1* gene. In normal cells, p21^{WAF1} protein negatively regulates the ability of DNA-methyltransferase to add a methyl group to CpG islands, thereby protecting these sites in the DNA from methylation. Inactivation or loss of p21^{WAF1} allows these sites to be methylated and transcriptionally silenced. Some tumour-suppressor genes shown to be methylated in tumours include *BRCA1*, *VHL*, and *p16INK4A*.

MOLECULAR MECHANISMS OF CANCER

Tumorigenesis *in vivo* is actually a multistep process requiring the alterations of two or more genes (Knudson, 1971). **Figure 2** depicts a single cell bearing a mutation or genetic 'hit' in gene A. This mutation is passed on to progeny cells, which, at a defined probability, sustain a second 'hit' in gene B. The figure depicts the mutation in gene A as a dominant 'oncogene-like' mutation and the mutations in both alleles of gene B as a recessive 'tumour-suppressor-like' mutation. Such alterations provide the initial steps in tumour formation. Every cell in the tumour carries the identical mutations that initiated tumour development.

The Two-hit Hypothesis of Knudson and Hereditary Cancers

One of the first concepts to arise regarding the molecular mechanism of tumours was suggested by Knudson and colleagues, who developed the two-hit hypothesis (Knudson, 1985). The assumptions of this hypothesis are threefold: malignant transformation of a single cell is sufficient to give rise to a tumour; any cell in a specific tissue is as likely to be transformed as any other of the same type; and once a malignant cell is generated, the mean time to tumour detection is generally constant. Once these assumptions are met, the model suggests that at least two events are necessary for carcinogenesis and that the cell with the first event must survive in the tissue long enough to sustain a second event.

Multistep Carcinogenesis Models

Land and Weinberg Model

At about the same time that Knudson proposed the two-hit hypothesis, Weinberg and Barrett independently suggested models of carcinogenesis based on the activation of a series of oncogenes. Weinberg suggested that the activation of two or more oncogenes is required for tumorigenesis and that the right combination must be activated in

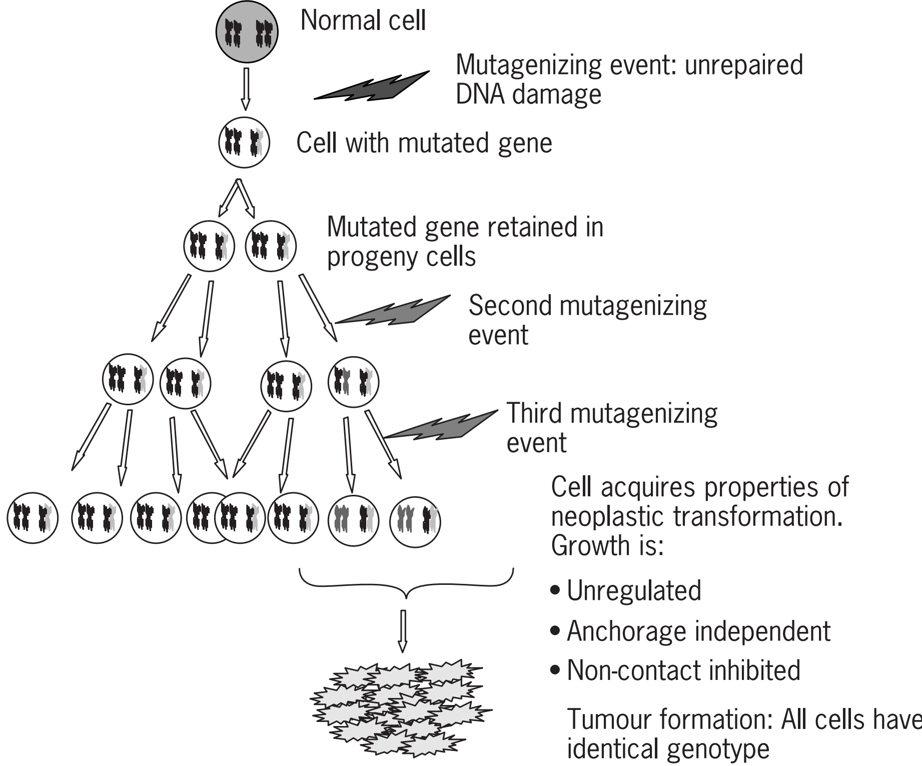


Figure 2 Acquisition of tumorigenic phenotype and clonality. Once a cell has acquired a mutation in an oncogene (depicted as the chromosome in light grey), that mutation is passed on to subsequent generations of progeny cells. These cells are still phenotypically normal, however. Cellular transformation occurs when a second and third mutation arise in a tumour-suppressor gene, e.g. in one of the previously mutated cells (depicted in dark grey). This cell now harbours three mutations in at least two different genes, and displays the hallmarks of neoplastic growth in culture.

the right context (Weinberg, 1983). Which oncogenes are activated is dependent on the signalling events each regulates. For example, Ras (a cytoplasmic oncoprotein) was shown to cooperate with Myc (a nuclear oncoprotein) to form tumours. Other combinations of cytoplasmic and nuclear oncoproteins also cause tumours to form, but one oncoprotein from each group must be activated.

Barrett's Model

Barrett's model further divides the process into tumour initiation vs tumour-promotion events (Boyd and Barrett, 1990). These investigators suggested that initiation is most often a mutational event, including mutations in a proto-oncogene, such as *ras*. Promotion, on the other hand, can be a mutational or an epigenetic change, and is defined as a series of '... qualitative, heritable changes in a subpopulation of initiated cells, resulting in malignancy or an increased potential to progress to malignancy.' In this model, morphological transformation (or initiation) occurs upon treatment with a mutagen or carcinogen. This event is followed by a loss or inactivation of a gene controlling cell senescence along with activation of immortalizing genes. With such changes, an immortal cell line is generated. The subsequent loss or inactivation of a tumour-suppressor gene or activation of a transforming oncogene then leads to the formation of a tumorigenic cell.

Vogelstein's Model

Studies by Vogelstein and colleagues led to a progression model in colon cancer which includes both the activation of oncogenes and the loss of tumour suppressors (Vogelstein and Kinzler, 1993). This model, dubbed the Vogelgram, is based on several observations. The first is that cancer cells contain 3–7 somatic mutations per cell. Second, benign tissue surrounding the malignant tissue frequently contains many of the same set of mutations found in the tumour but lacks at least one mutation that is found in tumour tissue. Third, certain genes have a high probability of mutating at each definable stage of colon cancer progression. Based on these and other genetic data, a model for colon cancer progression was formulated. **Figure 3** suggests that the loss of the tumour-suppressor gene *APC* occurs early in the process of transformation, converting colonic epithelial cells to a hyperproliferative state. Hypomethylation of DNA then occurs in the early ademona stage, followed by activation of the oncogene Ki-*ras* in carcinoma *in situ*. The tumour-suppressor genes *DCC* and *p53* are lost later in the disease,

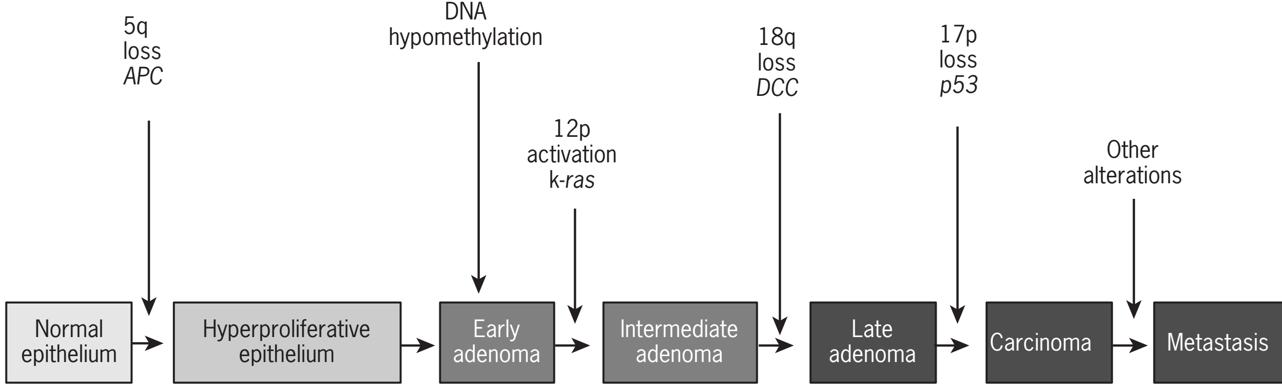


Figure 3 The Vogelstein model of multistep carcinogenesis. The progression of a normal colonic epithelium to metastatic colon cancer can be observed as it passes through several distinct stages. Chromosomal loss can be noted at different steps of progression. For example, chromosomal loss at 5q appears to occur prior to development of hyperproliferative epithelium whereas loss at 17p does not occur until the late adenoma transitions to a full carcinoma.

with the eventual development of a metastatic colon cancer.

CLINICAL CORRELATIONS

There are many ways in which mutations in cancer-promoting genes can occur. The predisposition to cancer can be inherited, as in patients with Li–Fraumeni syndrome, whose cells contain a germ-line mutation of *p53*, one of the cell cycle checkpoint regulators described above. Cells from patients with chronic myelogenous leukaemia often contain an abnormal chromosome resulting from a translocation between chromosomes 9 and 22, the so-called Philadelphia chromosome (Rowley, 1973). This abnormal fusion juxtaposes two genes, which code for the proteins BCR and the Abl tyrosine kinase, and results in aberrant activity and subcellular localization of the Abl protein. In breast cancer, *BRCA1* is mutated at specific sites in the gene. Such mutations are largely inherited.

In contrast to germ-line or inherited mutations, 'sporadic' cancers also can arise when a mutation occurs in a previously normal somatic cell. In this regard, environmental factors are thought to play major roles as mutagens or carcinogenic agents. For example, the relationship between tobacco smoke and lung cancer is well documented (Henderson *et al*., 1991). Anilines used in rubber tyre production are linked to the development of bladder cancer, while exposure to solar ultraviolet rays can cause melanoma (Case *et al*., 1993; Armstrong *et al*., 1997). Hundreds of chemical carcinogens that exist in food and products in daily use can either directly or through the production of secondary metabolites irreversibly alter a normal cell's DNA. So-called 'lifestyle' factors can also play a role. A link has been made between consumption of a diet high in animal fats and prostate cancer (Tzonou *et al*., 1999). In women, reproductive history and the resulting cumulative lifetime exposure to oestrogen correlate with an increased risk of breast cancer (Hankinson *et al*., 1995). How environmental factors trigger the activation and mutation of cancer-causing genes is, in many cases, still unclear. Subsequent chapters will detail what is known about a very complicated and intricate process.

REFERENCES

Armstrong, B. K., *et al*. (1997). Sun exposure and skin cancer. *Australasian Journal of Dermatology*, **38**, S1–S6.

Baylin, S. B. (1997). Tying it all together: epigenetics, genetics, cell cycle, and cancer. *Science*, **277**, 1948–1949.

Biscardi, J. S., *et al*. (1999). c-Src, receptor tyrosine kinases, and human cancer. *Advances in Cancer Research*, **76**, 61–119.

Bishop, J. M. (1982). Oncogenes. *Scientific American*, **246**, 80–92.

Bookstein, R., *et al*. (1989). Human retinoblastoma gene: long-range mapping and analysis of its deletion in a breast cancer cell line. *Molecular and Cellular Biology*, **9**, 1628–1634.

Boveri, T. (1929). The Origin of Malignant Tumors. Williams Wilkins, Baltimore, pp. 1–119.

Boyd, J. A. and Barrett, J. C. (1990). Genetic and cellular basis of multistep carcinogenesis. *Pharmacology and Therapeutics*, **46**, 469–486.

Case, R. A., *et al*. (1993). Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. Part I. The role of aniline, benzidine, alpha-naphthylamine, and beta-naphthylamine. 1954. *British Journal of Industrial Medicine*, **50**, 389–411.

Christofiori, G. and Semb, H. (1999). The role of the cell adhesion molecule E-cadherin as a tumor-supressor gene. *Trends in Biochemical Sciences*, **217**, 801–806.

Di Cristofano, A. and Pandolfi, P. P. (2000). The multiple roles of PTEN in tumor suppression. *Cell*, **100**, 387–390.

Folkman, J. (1992). The role of angiogenesis in tumor growth. *Seminars in Cancer Biology*, **3**, 65–71.

Hahn, S. A., *et al*. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*, **271**, 350–353.

Hankinson, S. E., *et al*. (1995). Reproductive factors and family history of breast cancer in relation to plasma estrogen and prolactin levels in postmenopausal women in the Nurses' Health Study (United States). *Cancer Causes and Control*, **6**, 217–224.

Harbour, J. W. and Dean, D. C. (2000). Rb function in cell-cycle regulation and apoptosis. *Nature Cell Biology*, **2**, E65–E67.

Heisterkamp, N., *et al*. (1985). Structural organization of the bcr gene and its role in the Ph′ translocation. *Nature*, **315**, 758–761.

Henderson, B. E., *et al*. (1991). Toward the primary prevention of cancer. *Science*, **254**, 1131–1138.

Irby, R. B., *et al*. (1999). Activating SRC mutation in a subset of advanced human colon cancers. *Nature Genetics*, **21**, 187–190.

Jacotot, E., *et al*. (2000). Apoptosis and cell cycle: distinct checkpoints with overlapping upstream control. *Pathologie et Biologie*, **48**, 271–279.

Kamb, A., *et al*. (1994). Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nature Genetics*, **8**, 23–26.

Knudson, A. G., Jr (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the USA*, **68**, 820–823.

Knudson, A. G., Jr (1985). Hereditary cancer, oncogenes, and antioncogenes. *Cancer Research*, **45**, 1437–1443.

Kolodner, R. D. and Marsischky, G. T. (1999). Eukaryotic DNA mismatch repair. *Current Opinions in Genetics and Development*, **9**, 89–96.

Konishi, M., *et al*. (1996). Molecular nature of colon tumors in hereditary nonpolyposis colon cancer, familial polyposis, and sporadic colon cancer. *Gastroenterology*, **111**, 307–317.

Kurschat, P. and Mauch, C. (2000). Mechanisms of metastasis. *Clinical and Experimental Dermatology*, **25**, 482–489.

Li, J., *et al*. (1997). PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*, **275**, 1943–1947.

Liu, W., *et al*. (1997). PTEN/MMAC1 mutations and EGFR amplification in glioblastomas. *Cancer Research*, **57**, 5254–5257.

Marx, J. (1993). How p53 suppresses cell growth. *Science*, **262**, 1644–1645.

Miyamoto, S., *et al*. (1998). Fibronectin and integrins in cell adhesion, signaling, and morphogenesis. *Annals of the New York Academy of Sciences*, **857**, 119–129.

Peddanna, N., *et al*. (1996) Genetics of colorectal cancer. *International Journal of Oncology*, **9**, 327–335.

Rous, P. (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *Journal of Experimental Medicine*, **13**, 397–411.

Rowley, J. D. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, **243**, 290–293.

Sager, R. (1989). Tumor suppressor genes: the puzzle and the promise. *Science*, **246**, 1406–1412.

Tzonou, A., *et al*. (1999). Diet and cancer of the prostate: a case-control study in Greece. *International Journal of Cancer*, **80**, 704–708.

Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. *Trends in Genetics*, **9**, 138–141.

Weinberg, R. A. (1983). Alteration of the genomes of tumor cells. *Cancer*, **51**, 1971–1975.

Weiss, R., *et al*. (eds) (1985). *RNA Tumor Viruses*, 2nd edn (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Wittinghofer, F. (1998). Ras signalling. Caught in the act of the switch-on. *Nature*, **394**, 317, 319–320.

Zheng, L., *et al*. (2000). Lessons learned from BRCA1 and BRCA2. *Oncogene*, **19**, 6159–6175.

FURTHER READING

Bishop, J. M. (1992). Oncogenes. *Scientific American*, **246**, 80–92.

Knudson, A. G. (1985). Hereditary cancer, oncogenes, and anti-oncogenesis. *Cancer Research*, **48**, 1437–1442.

Land, H., *et al*. (1983) Cellular oncogenes and multistep carcinogenesis. *Science*, **222**, 771–778.

Sager, R. (1989). Tumor suppressor genes: the puzzle and the promise. *Science*, **246**, 1406–1412.

Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. *Trends in Genetics*, **9**, 138–141.

Weiss, R., *et al*. (eds) (1985). *RNA Tumor Viruses*, 2nd edn (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Chapter 4

Inherited Predispositions to Cancer

Gareth Evans
St. Mary's Hospital, Manchester, UK

CONTENTS

- Introduction
- Retinoblastoma
- Genetic Syndromes
- Common Cancer Predisposition
- Conclusions

INTRODUCTION

There has been increasing evidence of familial predisposition to cancer since the classic model of hereditary retinoblastoma was outlined (Knudson, 1971). The notion that some cancer is hereditary has long been held by more than just a few diehard clinicians. The earliest reports of cancer families date back more than 180 years to a large cluster of breast cancer in the wife and family of a French physician named Broca and the cluster of gastric cancer in Napoleon's family. Despite the pioneering work of clinicians and researchers such as Henry Lynch and Mary-Claire King in the USA in the 1960s to 1980s, demonstrating the hereditary nature of at least a proportion of cancers such as those affecting the breast and colon, the hereditary element was not proven until the advent of molecular biology when abnormalities were demonstrated in cancer-predisposing genes. It is, therefore, only since 1987 that developments in molecular biology have proven the hereditary nature of a small proportion of certain common cancers. That cancer is now indisputably 'genetic' at the cellular level is beyond dispute. All tumours result from mutations or deletions of two types of gene (Eeles *et al*., 1996): the tumour suppressor gene, which needs to be inactivated to allow growth (like the brakes on a car), and the oncogene, which requires activation to promote growth (like the accelerator pedal of a car being stuck down). The great majority of these events are acquired whether through replication error (simple copying of DNA during cell division) or due to external agents (chemical mutagens, radiation, viruses). Occasionally, mutations in tumour suppressor genes can be inherited rather than acquired. Identifying the genes which cause hereditary disease has given an insight into many cancers. The role of cancer-predisposing genes in the causation of sporadic cancer is still the subject of much research, and we can still learn from the more obscure cancer-prone syndromes.

Broadly, the predisposition can be subdivided into rare genetic syndromes which have malignancy as a high-risk side effect and a larger group which cannot be easily identified clinically, and which have a strong family history of one or more common malignancies.

RETINOBLASTOMA

Retinoblastoma has been the model from which much of our current knowledge of tumour suppressor genes was fashioned. This early childhood eye malignancy was recognized as having a familial tendency in the nineteenth century. About 50% of cases are due to the inheritance of a gene defect in one copy of the retinoblastoma gene (*RB* on chromosome 13), and over 90% of individuals who carry a mutation will develop retinoblastoma, usually bilaterally. In 1971, Knudson (1971) proposed that the disorder was caused by mutational events in both copies of the gene. Those cases that inherited a mutated copy, only need one further mutation and are far more likely to develop the malignancy, which occurs at a younger age and is usually bilateral. The sporadic cases require two mutations ('hits') in a retinal cell as opposed to one in the familial case (**Figure 1**) and so bilateral tumours are extremely unlikely to occur and the unilateral tumours present later. This hypothesis, which has since been proven to be true, now bears the conceiver's name. Familial retinoblastoma may even be present in foetal life, as can be seen in **Figure 2 (see colour plate section)**. This case had a 13q deletion as a result of a maternal chromosomal translocation.

The discovery of retinoblastoma cases with cytogenetically visible constitutional deletions in the long arm of chromosome 13 (Francke and King, 1976; Knudson *et al*., 1976) concentrated research on that region. One of the genes deleted in these cases, esterase D, then acted as a marker for further studies. One study showed that although

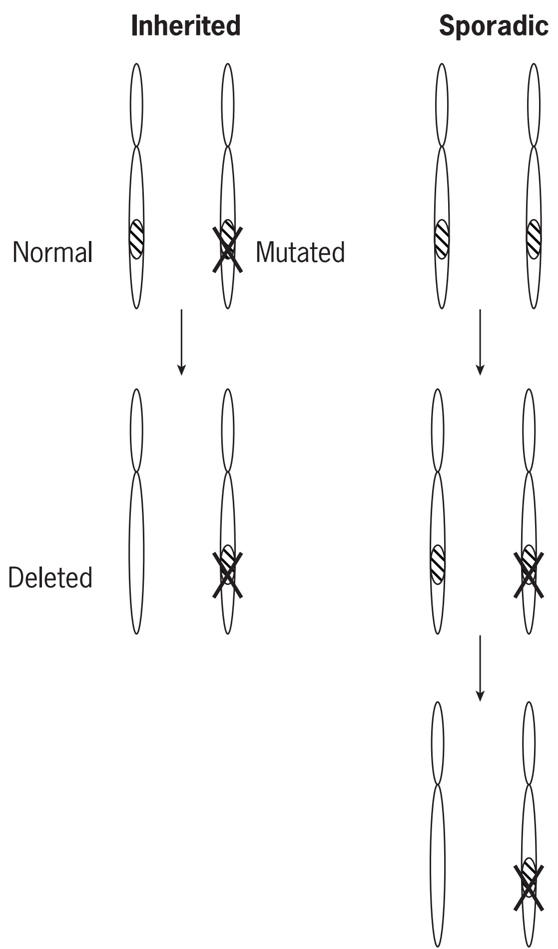


Figure 1 Ideogram of the 'two-hit' hypothesis. The first hit is usually a mutation (represented by a cross) which causes disruption of the protein product. The second hit is often loss of the whole gene by deletion of part or all of the chromosome on which the gene resides.

an individual was heterozygous for esterase D their tumour was hemizygous, suggesting loss of material by deletion or monosomy (Godbout *et al*., 1983). The introduction of restriction fragment length polymorphisms (RFLPs) led to further studies showing loss of constitutional heterozygosity. RFLPs rely on the differences in large portions of DNA between individuals and therefore between the two equivalent chromosomal regions in any one individual. This difference means that enzymes (endonucleases) which cut at specific gene codes will cut at different sites on the chromosome. There is therefore a good chance that a gene or genetic marker will end up on two different lengths of DNA when a particular endonuclease is used. If an individual is shown to have two lengths (heterozygous) on constitutional testing, but only one in their tumour, then loss of constitutional heterozygosity has occurred. These RFLPs were used to localize the gene further until it was eventually cloned (Friend *et al*., 1986). Since the isolation of the *RB* gene, many groups have tried to isolate the underlying defects that cause retinoblastoma and that make the gene important in cell regulation. These studies have confirmed that *RB* acts as a typical tumour suppressor gene with an initial mutational event consisting of small changes in nucleotides resulting in truncation of the resultant protein product. These are usually nonsense mutations leading to an early stop codon or small frameshift deletions or insertions with a similar downstream effect. Nonetheless, pathogenic missense mutations do occur although they are more difficult to prove. Simple cosegregation of a missense mutation in a small family is not enough, although the presence of an amino acid change in a functional domain with cross-species conservation does add credence. In the final estimation it is only with functional assays that a pathogenic effect can be proven. Initial studies showed a relatively low rate of mutation identification in the *RB* gene (Liu *et al*., 1995), but with a combination of strategies including a coding sequence analysis taking in intron/exon boundaries (for splicing mutations) and a deletion strategy such as fluorescent *in situ* hybridization (FISH) and Southern blotting, the great majority of aberrations of the *RB* gene can be identified (Lohmann *et al*., 1996).

The retinoblastoma gene also predisposes to osteosarcoma. An individual who has had an enucleation and irradiation for retinoblastoma is 500 times more likely to develop the bone tumour. As the gene is involved in many common carcinomas such as that of the breast, it is likely that survivors will be at risk of these tumours also.

GENETIC SYNDROMES

These are usually readily identifiable by a clinical phenotype or by laboratory tests. The syndromes may be autosomal dominant or recessive or X linked (**Tables 1** and **2**). Of these it is the dominant conditions which are of most interest as they are likely to represent the inheritance of a faulty copy of a tumour suppressor gene, which predisposes the individual to common cancers. Although the conditions are generally uncommon, tumour suppressor genes in general are likely to play a fundamental role in the genesis of tumours, which affect a third of all humans in their lifetime. The identification of those causing genetic syndromes is likely to lead to more specific treatment using gene therapy, as well as earlier identification, monitoring and, most hopeful of all, prevention of common cancers.

Familial Adenomatous Polyposis (FAP)

FAP is the model condition by which researchers have hoped to transpose knowledge of a rare genetic disease to a commonly occurring cancer. FAP is an autosomal dominant condition characterized by the development of hundreds to thousands of adenomatous polyps in the colon and rectum, usually by 30 years of age (**Figure 3; see colour**

Table 1 Examples of autosomal dominant syndromes predisposing to cancer and their chromosomal location

Name of disease	Location (chromosomal)	Protein
FAP	5q	APC
NF1	17q	Neurofibromin
NF2	22q	Merlin/ schwannomin
Von Hippel–Lindau	3p	pVHL
MEN1	11q	Menin
MEN2	10q	RET
Gorlin	9q	PTCH
Tuberous sclerosis (TSC1)	11q	Hamartin
(TSC2)	16q	Tuberin
Juvenile polyposis	18q and other(s)	pDPC/SMAD4
Peutz–Jeghers	19p and other(s)	pSTK11/LKB1
Cowden	10q	PTEN
Tylosis	17q	Not found

Table 2 Autosomal recessive and X-linked conditions predisposing to malignancy

Name of condition	Chromosomal location	Protein
Fanconi anaemia	8 loci	4 found
Bloom syndrome	15q	pBLM
Ataxia telangiectasia	11q	pATM
Xeroderma pigmentosa	7 types	2 types
Chediak Higashi	1q	pLYST
Albinism	11q	OCA1, OCA2
Bruton	Xq	BLk
Wiscott Aldrich	Xp	CD43

plate section). If untreated this leads to the almost inevitable development of a colorectal cancer by 60 years of age. The condition may be associated with osteomas and epidermal cysts and this subdivision was designated Gardner syndrome (Gardner, 1951). However, most FAP families show these extraintestinal features to some extent. FAP has, in common with many other conditions, been mapped to a chromosomal region as a result of the finding of a constitutional (present in every cell) chromosomal anomaly in a manifesting case (Herrera *et al.*, 1986). This patient had a small interstitial deletion on the long arm of chromosome 5 and manifested extracolonic features as well as multiple polyps. Following this discovery, Bodmer *et al.* (1987) localized the gene for FAP to 5q21–q22 by genetic linkage, using families mainly from the well-established St. Marks Polyposis Register in London. Of great interest was that the same region of chromosome 5 was implicated in sporadic colorectal cancer (Solomon *et al.*, 1987). This meant that the gene could be localized more accurately by using tumour material from sporadic cases as well as from FAP cases. If the tumour is analysed for loss of genetic material using probes mapped to the implicated region of chromosome 5, a deletion map can be drawn. When the normal or 'wild-type' allele is lost from an FAP patient's tumours, the assumption is that a second hit has occurred, removing the only functioning copy of a tumour suppressor gene. The nature of this loss is known as loss of constitutional heterozygosity (LOH).

Once the gene had been localized to a relatively small chromosomal region, several research groups embarked on a project to isolate contiguous sequences of genes spanning the area. Subclones were then used to identify the position of candidate genes, which were expressed in normal colonic mucosa. Two of these genes, *APC* and *MCC* (Kinzler *et al.*, 1991), were thought likely to be involved in tumorigenesis, because of the structure of the proteins for which they encoded. All that remained was to identify mutations in one of these genes, in the germ-line of patients with FAP. This was duly achieved when 10 *APC* germ line mutations were described (Nishisho *et al.*, 1991). These mutations were not only likely to disrupt the protein structure, but were also found only in the *APC* gene. The idea that Gardner syndrome was a separate entity was refuted since the mutations occurred in patients with or without extraintestinal manifestations. However, FAP was one of the first conditions in which a clear correlation between genotype (the genetic change in *APC*) and phenotype (the clinical picture) emerged. Patients with mutations in the early part of the gene (5′: exons 2–5) had a very mild clinical picture with late onset of polyps (Spirio *et al.*, 1993), whereas those with mutations from exon 9 through to codon 1450 of exon 15 had a classical disease course with nearly all patients manifesting the typical congenital retinal pigmentation. However, those with mutations beyond codon 1450 showed typical Gardner syndrome features (osteomas, cysts and desmoid disease) without retinal signs (Davies *et al.*, 1995). There are even families who with extreme mutations in exon 15 show little else other than desmoid disease (Scott *et al.*, 1996).

Currently it is possible to offer predictive genetic tests before symptoms to the majority of at-risk individuals. Although looking for germ-line mutations is laborious and not guaranteed to find the mutation, it is the most reliable. Nonetheless, testing using linkage analysis in families with more than one affected member is still very useful, especially when a germ-line mutation cannot be identified. This, combined with opthalmological screening for congenital hypertrophy of the retinal pigment epithelium (CHRPE), may reduce initial risks of 50% to well below 1% (Burn *et al.*, 1991).

Von Hippel–Lindau

Von Hippel–Lindau (VHL) is another dominantly inherited familial cancer syndrome. The most frequent complications

are benign tumours of blood vessels, particularly in the eye (retinal angiomas), and haemangioblastomas of the cerebellum. Other features include renal cell carcinoma, phaeochromocytoma and renal, pancreatic and epididymal cysts. The syndrome is very variable but most individuals present before 40 years of age (Maher *et al*., 1990). The first clue to the location of the gene for VHL was the finding of a reciprocal translocation involving chromosomes 3 and 8 in a family with hereditary renal carcinoma (Cohen *et al*., 1979). Later, Teyssier *et al*. (1986) were able to show deletions of the short arm of chromosome 3 in other renal cell carcinomas. Linkage in families with VHL was confirmed on 3p in 1988 (Seizinger *et al*., 1988). The gene was finally cloned in 1993 (Latif *et al*., 1993) and codes for a relatively small protein.

Type 1 Neurofibromatosis (NF1)

NF1 is more common than NF2, but the disease may be so mild that an affected individual may never present to their doctor. The main manifestations are in the skin, with the appearance of café au lait patches and cutaneous neurofibromas in the first and second decades, respectively. The most famous potential misdiagnosis of NF1 was Joseph Merrick, the 'elephant man,' who in reality probably had Proteus syndrome (Clark, 1994). One potential serious complication of NF1 is optic gliomas, which may occur in up to 15% of cases (Listernick *et al*., 1989). These are usually very low grade and asymptomatic and if they are not specifically sought levels of around 1.5% are found. Other CNS gliomas do occur but their frequency is probably well below 5%. Meningiomas and vestibular schwannomas do not occur in excess frequency in NF1 (McGaughran *et al*., 1999). Phaeochromocytoma and spinal neurofibromas may develop as well as rhabdomyosarcomas, but these are relatively rare. Malignant change in neurofibromas can result in a malignant peripheral nerve sheath tumour (MPNST) in about 10% of NF1 patients in their lifetime (McGaughran *et al*., 1999).

The *NF1* gene was cloned in 1990, although it took over a year fully to characterize the gene from the first discovery of deletions in the germ-line of some familial cases (Viskochil *et al*., 1990). It is a massive gene containing over 300 kilobases of DNA divided into 50 exons (Collins, 1991). The gene transcribes a 327 kDa GAP protein containing 2818 amino acids. The protein, which binds to the oncogenic protein Ras, is found in all tissues. It is expressed at the cellular level in the perinuclear vesicles and microtubules. As p120 GAP is expressed more in the neural crest this may explain why *NF1* specifically affects neural tissue. Although diagnosis is possible by looking for germ-line mutations, this approach has not found any particular hot spots of mutation, although extensive analysis using a number of approaches as for retinoblastoma does detect 95% of mutations (Messiaen *et al*., 2000). Predictive diagnosis therefore still depends mainly on linkage in existing families.

There is now good evidence that at least two variant conditions of NF1 are caused by mutations in the *NF1* gene. Watson syndrome was shown to be linked to the *NF1* locus (Allanson *et al*., 1991) and NF–Noonan syndrome to be due to mutations in the gene itself (Colley *et al*., 1996).

Type 2 Neurofibromatosis (NF2)

NF2 is an autosomal dominant genetic disease characterized by the development of bilateral vestibular schwannomas (acoustic neuromas) in the second and third decades. Only recently has it been formally separated from the more common NF1 (von Recklinghausen disease), after the National Institutes of Health Consensus Development Conference Statement on Neurofibromatosis (1987). The first probable reported case of NF2 was that of Wishart (1822). Bilateral vestibular schwannoma had been thought to be part of von Recklinghausen neurofibromatosis (NF1) after reports of similarities in cases to those with the peripheral form (Cushing, 1917). Several reports emphasized the paucity of skin findings in families with bilateral vestibular schwannoma (Gardner and Frazier, 1930), and suggested that bilateral vestibular schwannoma represented a separate central form of von Recklinghausen neurofibromatosis. It was not until the separate assignment of NF1 to chromosome 17 (Seizinger *et al*., 1987) and NF2 to chromosome 22 (Rouleau *et al*., 1987) that the diseases were finally shown to be two distinct disorders. NF2, although less common than the type 1 form (incidence 1 in 35 000 compared with 1 in 3000) (Evans *et al*., 1992a), is more likely to present clinically at some time. All cases will develop a CNS tumour by 55 years of age. Although most of these tumours are benign (meningiomas, schwannomas; see **Figure 4**), 6% will develop a malignant glioma or ependymoma (Evans *et al*., 1992b).

The clue to the location of the *NF2* gene was not a constitutional chromosomal anomaly, but rather cytogenetic abnormalities found on chromosome 22 in human meningiomas and later in vestibular schwannomas and other tumours from NF2 patients. This candidate region was then confirmed as the likely location for the *NF2* gene by linkage analysis in a large US family (Rouleau *et al*., 1987). The gene was isolated simultaneously by two groups (Rouleau *et al*., 1993; Trofatter *et al*., 1993), and genotype phenotype correlations have been identified (Evans *et al*., 1998a). Mutations which give rise to a truncated protein are associated with a severe, multitumour, early-onset disease course, whereas those that give a nearly normal protein product (missense mutations) or no product (large deletions) give mild disease. Another feature of NF2 that is likely to be an important factor in other tumour-prone disorders is mosaicism (Evans *et al*., 1998b). If a mutation occurs after conception, say at the

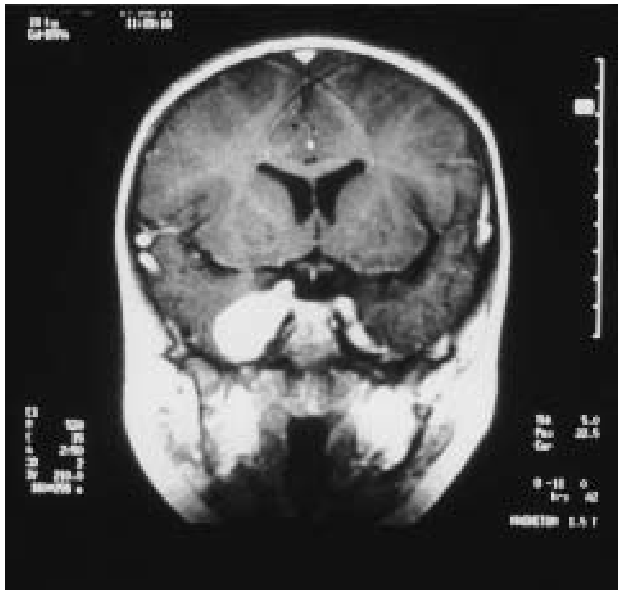


Figure 4 MRI scan of a 26-year-old man with type 2 neurofibromatosis. The scan shows bilateral enhancing tumours in the cerebello-pontine angles and meningiomas around the brain. The risk of developing bilateral tumours by chance is 1 in 2×10^6, yet 95% of individuals with mutations in the *NF2* gene develop bilateral vestibular schwannomas (acoustic neuromas).

eight-cell stage, roughly one eighth of all the cells will have an *NF2* mutation, which means that there are two different cell populations, one of which predisposes to tumours and could be transmitted to any offspring.

Gorlin Syndrome

Gorlin or naevoid basal cell carcinoma syndrome is another autosomal dominant condition which predisposes to malignancy. The condition is characterized by the development of multiple jaw keratocysts in the second decade and basal cell carcinomas in the third decade onwards. Gene mutation carriers also have a recognizable appearance or morphology. They have macrocephaly with bossing of the forehead and the face is usually covered with white milia. The facial features are often coarse and the shoulders slope downwards. Most individuals have a skeletal anomaly such as a bifid rib or wedge-shaped vertebra and ectopic calcification, particularly in the falx, is almost certain by 20 years of age.

Individuals with Gorlin syndrome are also at risk of developing the childhood brain malignancy medulloblastoma, which occurs in 5% of cases (Cowan *et al.*, 1997) and cardiac and ovarian fibromas (Gorlin, 1987; Evans *et al.*, 1993). Malignant transformation has been described in the ovarian fibromas (Strong, 1977), but they usually remain benign, although they can reach a large size and are often calcified (**Figure 5**). The clue to the location

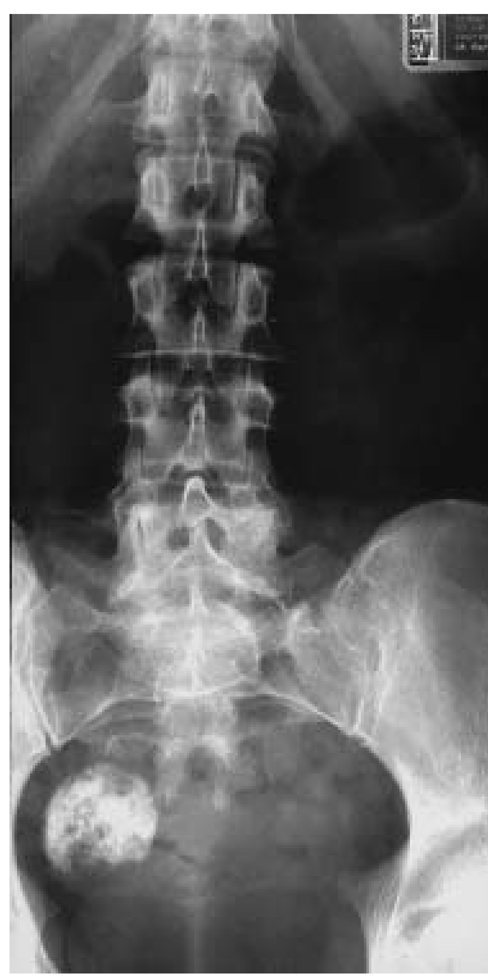


Figure 5 Large calcified ovarian fibroma on abdominal X-ray in a patient with Gorlin syndrome.

of the Gorlin gene again came from tumour deletion studies. Gailani *et al.* (1991) found that 40% of basal cell carcinomas that they studied had deletions of the proximal region on the long arm of chromosome 9. The condition has now been shown to be linked to that region using affected families and there is no locus heterogeneity (Farndon *et al.*, 1992). The gene itself was subsequently identified as a homologue of the drosophila gene *PTCH* (Hahn *et al.*, 1996). As the mean age at onset of medulloblastoma in Gorlin patients is 2 years compared with over 7 years in the general population, and there is loss of the normal copy of the gene in tumours (Cowan *et al.*, 1997), this confirms *PATCHED* as a tumour suppressor in both medulloblastoma and basal cell carcinoma. Basal cell carcinomas occur at great frequency in the periphery of the radiation field 5–10 years after irradiation in Gorlin syndrome (**Figure 6; see colour plate section**).

Tuberous Sclerosis

Tuberous sclerosis is a condition in which hamartomas are a primary feature. These may occur in the brain (the

'tubers' of the name), or in the kidney, heart and elsewhere. Patients have a number of external features, which make early or presymptomatic identification possible. Depigmented patches of skin or 'ash leaf patches' can be seen with a Wood's light and a characteristic facial skin eruption known as adenoma sebaceum is often present. Subungual (under the finger nails) fibromas are another feature peculiar to tuberous sclerosis. Tumours may occur in the heart during foetal life (rhabdomyomas) and the malignancy to which the condition particularly predisposes is the brain tumour glioma. Unusually for a condition which is likely to be caused by one gene, two separate genetic loci have eventually been implicated. In 1987, Fryer *et al.* (1987) showed linkage to 9q in several families. However, many reports following this had contradictory results. This has now been shown to be due to locus heterogeneity, for although linkage to 9q has been confirmed in some families, the other major locus was actually cloned first in 1993 on chromosome 16 (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). The later cloning of *TSC1* was partly due to the fewer individuals affected with the disease at this locus, although the linkage came first as actual families are more numerous (DeVries and Bolton, 2000).

Multiple Endocrine Neoplasias

The multiple endocrine neoplasias are further conditions which predispose to benign tumours and at least one malignancy. In MEN1 the organs affected are the parathyroid glands, pituitary and pancreas. The most serious problem is with islet cell tumours of the pancreas, which secrete gastrin. These cause the Zollinger–Ellison syndrome of which MEN1 makes up a large proportion of cases. The gastrin-secreting tumours may become malignant, seeding to the liver and other organs. However, many cases do not manifest the condition overtly even late in life. The serum calcium level is raised in 90–97% of cases, but laborious testing and repeated screening may be necessary. The availability of genetic tests has, therefore, greatly simplified screening of at-risk individuals. The location of the *MEN1* gene was confirmed at 11q13 by linkage analysis in families (Larsson *et al.*, 1988). This and other studies have also shown loss of 11q alleles in the MEN1 tumours. The *MEN1* gene was eventually cloned in 1997 (Chandrasekharappa *et al.*, 1997) and the protein product was termed menin. Presymptomatic testing is now available by mutation testing of an affected individual, or by linkage analysis in families with more than one affected individual.

MEN2a or Sipple syndrome is an autosomal dominant disease with high penetrance and variable expression. The association of medullary carcinoma of the thyroid and phaeochromocytoma are the hallmarks of the condition. Parathyroid tumours are less commonly found. MEN2 makes up 25% of all cases of medullary thyroid carcinoma, with nearly all MEN2 cases developing this tumour some time in life. The tumours in MEN2 are often bilateral and are preceded by C cell hyperplasia. Phaeochromocytoma occurs in 50% of individuals and is often multifocal. Screening of at-risk cases involves serum calcitonin estimation and monitoring of blood pressure, and has been greatly enhanced by the development of genetic testing, which removes over 50% of individuals from screening programmes. The test has targeted those in which thyroid cancer can be prevented by early or prophylactic thyroidectomy.

The *MEN2a* gene was localized to chromosome 10 by linkage analysis using RFLPs (Simpson *et al.*, 1987), and later localized to 10q21.1 by *in situ* hybridization. Although researchers concentrated on trying to identify a tumour suppressor gene, it was eventually found that MEN2 was due to activating mutations in an oncogene called *RET* (Mulligan *et al.*, 1993), although MEN2b differs from MEN2a in that the primary feature is the development of mucosal neuromas especially of the tongue. Medullary thyroid cancer is also a major feature and phaeochromocytoma also occurs but, in contrast to MEN1 and MEN2a, there is no parathyroid disease. Both conditions are caused by activating mutations in *RET*, although most of MEN2b is caused by a single mutation and MEN2a by five different substitutions at cysteine residues.

Other Dominant Syndromes

Tylosis, juvenile polyposis, Peutz–Jeghers syndrome, multiple exostosis and multiple lipomatosis are other dominantly inherited disorders which may predispose to malignancy.

Wilms' Tumour

Wilms' tumour, like retinoblastoma, has also been known for some time to have a hereditary element. However, the genetic basis is far more complex and the familial element much smaller. The first step to identifying a gene came with the discovery of a cytogenetically visible deletion in chromosome 11 in families with autosomal dominant aniridia who appeared to be predisposed to Wilms' tumour (Riccardi *et al.*, 1978). Deletions in this area (11p13) also lead to genital and renal anomalies and mutations within the *WT1* gene itself lead to abnormal genital development (Pelletier *et al.*, 1991). At first it was thought that this locus would be similar to retinoblastoma, but *WT1* has now been shown to be one of at least three genes involved in Wilms' tumour development. In 1989, Koufos *et al.* (1989) demonstrated tight linkage to 11p15.5 in a family with

Beckwith–Wiedemann syndrome (a mainly sporadic growth disorder with neonatal hyperinsulinism and features such as exompholos) and Wilms'. Beckwith–Wiedemann syndrome is now known to be due to complex mechanisms involving a number of genes including *CDKN1C* and *IGF2* where there is either loss of maternal copy or gain of paternal copy in an imprinted area (Lam *et al.*, 1999). In addition, Grundy *et al.* (1988) excluded both 11p13 and 11p15 in linkage analysis of a large family with dominant Wilms' tumour. A third and fourth locus has now been confirmed in families manifesting primarily Wilms' tumour alone, but a further locus probably exists.

Autosomal Recessive and X Linked Conditions

A list of these conditions and the chromosomal locations of the predisposing genes can be found in **Table 2**. These are less likely to present to the clinician as they are generally less common and mainly predispose to haematological malignancy.

COMMON CANCER PREDISPOSITION

Recent years have seen an enormous improvement in our understanding of the mechanisms of carcinogenesis. Most cancers require a number of genetic changes in a cell before an invasive tumour results. Few are likely to be caused purely by the loss of two copies of a single tumour suppressor gene as in retinoblastoma and the number of changes probably varies between four and 10. A combination of loss of function of tumour suppressor genes and activation of oncogenes is usually involved. The particular combination and order may alter both the histological and invasive nature of the cancer. There is now evidence that a minority of people who develop common cancers have inherited a faulty gene which puts them at high risk of malignancy, but this is not recognized as a syndrome apart from in the family history. Adenocarcinomas are more likely than carcinomas of squamous epithelium to have a strong hereditary component with 4–10% of all breast, ovarian and colon cancer resulting from an inherited gene defect. The discovery of germ-line (inherited) mutations in the *TP53* gene on the short arm of chromosome 17 in families with a peculiar combination of early and multiple tumours was the first proven example of this. Otherwise, predisposition can be relatively site specific with genes being isolated in recent years for melanoma, prostate cancer, pancreatic cancer and paraganglioma, but this chapter will focus on the three common cancers which have attracted the greatest attention.

Li–Fraumeni Syndrome (LFS)

This cancer predisposition syndrome was first outlined in 1969 by Li and Fraumeni (1969). They reported four families with autosomal dominant predisposition to soft tissue sarcoma, breast cancer and other tumours in children and adults. Many reports have followed, either describing further families or reporting an increased risk of cancers in first degree relatives of cases with soft tissue sarcoma. In 1988, Li *et al.* (1988) analysed 24 kindreds with an aggregation of tumours typical of the syndrome. They showed a predominance of soft tissue sarcoma, osteosarcoma and breast cancer, with an excess of adrenocortical carcinoma, brain tumours and leukaemia. Williams and Strong (1985) applied a segregation analysis to test the hypothesis that the disease was due to an autosomal dominant gene. They not only confirmed this, but also were able to predict that 50% of gene carriers would develop an invasive cancer by 30 years of age and 90% by 70 years. Although the syndrome (also known as SBLA) is rare, its importance lies in the unusual range of cancer predisposition and that it is caused by the first of the major predisposition genes to be identified.

Linkage studies using markers on many chromosomes was not an easy approach owing to the paucity of suitable families, because the condition is so lethal that often no living affected members are available. The genetic fitness of cases is also likely to be reduced owing to development of malignancies in childhood and the faulty gene may therefore die out in families, making fewer available for analysis. Also, until an individual has developed a syndromal cancer, their genetic status is unknown. Penetrance is high but not complete, so one cannot be completely sure of the status of an unaffected individual even in their sixties. The lack of an identifiable phenotype present in many of the conditions described earlier is also a major hurdle. Therefore, Malkin *et al.* (1990) used a candidate gene approach. They argued that the condition was probably due to a mutation in a tumour was an unlikely choice as retinoblastoma had not been observed in any LFS families. They chose the *TP53* gene, which was the second to be recognized, but the first to be cloned. This gene had been implicated in at least half the typical cancers featuring in LFS by analysis of the tumours by mutation and deletion studies. Malkin *et al.* examined the *TP53* gene in normal somatic cells of affected and unaffected members from five families. This was achieved by amplifying the genomic region encoding exons 5–8, which contain most of the conserved domains and are frequently mutated or deleted in tumours. They then sequenced the region using multiple primers. Affected members in all five families showed mutations in this region, with two occurring at codon 248, which is a hot spot for tumour mutations. All were in the conserved region IV in which no polymorphisms had been found in the germ-line before. They were also able to show loss of

the ‘wild-type’ allele in a tumour in one affected family member.

The involvement of *TP53* was then confirmed in a further family by Shrivastava *et al*. (1990). However, the initial impression that the process would be simple does not appear to be the case. A Manchester study (Santibanez-Koref *et al*., 1991) found that only two out of eight families had *TP53* mutations in exons 5–8, although they did confirm the hotspot at codon 248. Toguchida *et al*. (1991) found eight germ line mutations in 196 osteosarcoma patients but none in 200 controls. In contrast to previous studies, a family history was not present in some of the cases and the mutations were not all in the conserved regions of *TP53*. The absence of family history was not surprising, as one would predict a high new mutation rate in LFS to maintain the frequency of the condition in the population. Subsequent work has shown that with complete gene sequencing it is possible to identify mutations in over 70% of classical families. There is also evidence of genotype phenotype correlations, with much higher penetrance for mutations in the DNA core-binding domain (Birch *et al*., 1998). Recent evidence has also shown that up to 80% of childhood adrenocortical tumours having germ-line mutations in *TP53* (Varley *et al*., 1999). This makes this childhood tumour the most hereditary of all malignancies including retinoblastoma. While it is still likely that nearly all LFS families have *TP53* involvement, a mutation in the *hCHK2* gene has now been identified in at least one family (Bell *et al*., 1999). A typical family with a *TP53* mutation at codon 191 is shown in **Figure 7**. The early age of the breast cancers is particularly noticeable, with one woman having bilateral disease as well as other primary tumours. Clearly, predictive tests in this and other families are now possible. However, until a mutation is found in an affected family member, reassurance of at-risk members is not possible even if the conserved domains are screened.

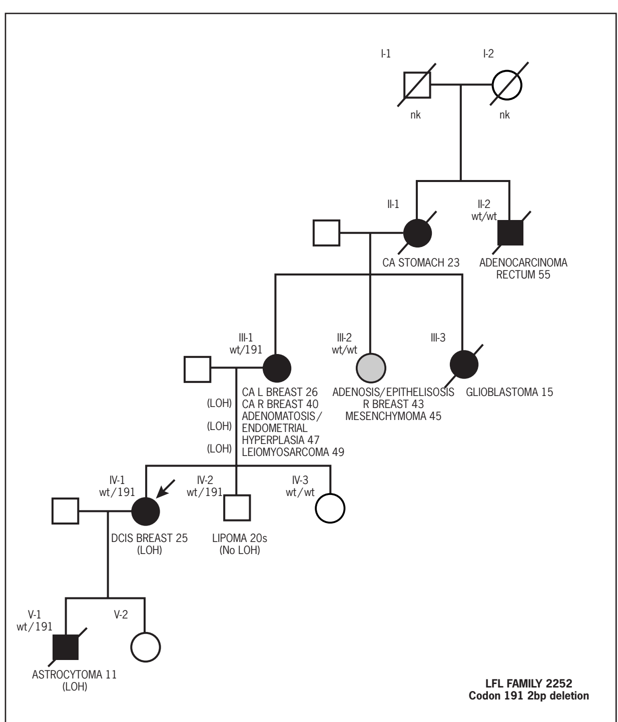


Figure 7 A family tree showing a dominant inheritance pattern of a TP53 codon 191 two base pair deletion in a family with LFS.

Breast Cancer

Breast cancer is now known to occur as part of a high penetrance predisposition such as in LFS, and in BRCA1/2 families, but may also be caused by mutations in genes such as *ATM* and *PTEN* which confer a risk of <50%. Breast cancer has long been known to have a familial tendency, as discussed earlier, and there is a profusion of supporting literature. Evidence from meticulous epidemiological studies shows that 4–5% of breast cancer is due to a dominant cancer gene with high penetrance and a population frequency in the USA of 0.003 (Claus *et al*., 1990). Studies in the UK have confirmed this population frequency and gave useful data on which risk estimation can be based. Important factors which point towards a possible familial predisposition are the number of relatives, particularly first degree, who have been affected, the age at which they developed the disease (early onset more significant) and whether bilateral or associated with other tumours.

The search for the gene or genes responsible for dominantly inherited breast cancer was dogged by some of the problems found in LFS. Ascribing status is difficult in unaffected cases even late in life and many of the known affected cases have died. Obtaining samples in some cases may depend on the use of stored paraffin block material, which may be unavailable and relies on having polymerase chain reaction technology and suitable probes, which are of course the norm nowadays. Additionally, as breast cancer is so common, affecting one in 11–12 women in their lifetime in the UK (HMSO, 1998), chance aggregations are likely to occur and non-gene mutation carriers in dominant families may be affected. The other problem, which could only be found by trial and error, was that of locus heterogeneity. Many chromosomal locations had been implicated by cytogenetic and LOH studies on tumour material. Chromosomal regions known to show involvement in more than 20% of tumours by molecular studies are 1p, 1q, 3p, 11p, 13q, 14q, 15q, 17p, 17q and 18q. Many of these regions were already known to contain tumour suppressor genes, e.g. *RB* on 13q, *TP53* on 17p and *DCC* on 18q, and these genes are likely to be involved in a multistage process towards malignancy. In a major breakthrough, Hall *et al*. (1990) were able to

show linkage in some breast cancer families to 17q12–21. They looked at over 20 families from whom they had collected samples over many years, including many cases that had subsequently died. Nevertheless, they still had to use a PCR probe to work with paraffin block tumour samples in some cases. When all families were included in the linkage analysis the region on 17q was excluded. However, when the families were stratified in terms of their average age at onset, the first seven families showed a significant linkage to 17q12 (lod score approaching 6). They argued that a large proportion of early-onset breast cancer families (<46 years) were caused by a mutated gene on 17q. Without the work of Hall *et al*., and subsequently by the Breast Cancer Linkage Consortium, it could have been many years before research was focused on this region. The problem of genetic locus heterogeneity was only overcome by a combination of meticulous collection of samples, innovative ideas and luck. Another possible gene that was implicated at about this time was the oestrogen receptor gene on chromosome 6. However, this has not since been confirmed. Following the discovery of linkage of breast cancer to 17q (Hall *et al*., 1990), Narod *et al*. (1991) undertook linkage on five families with breast/ovarian aggregation. They found that three of the families were linked to a locus at 17q12–q23 and their additive lod scores reached statistical significance. Subsequent work by the Breast Cancer Linkage Consortium showed that 80% of breast/ovarian families with four or more affected patients were linked the 17q locus (Easton *et al*., 1993). The following year heralded the identification of the first major breast cancer predisposing gene *BRCA1* (Miki *et al*., 1994). Surprisingly, *BRCA1* does not appear to be involved as a significant acquired mutation (somatic) in non-hereditary breast cancer. In the same month that the *BRCA1* gene was identified, the location of a second gene dubbed *BRCA2* was announced on chromosome 13. A year later *BRCA2* was cloned and again there was little evidence of involvement in sporadic disease (Wooster *et al*., 1995). It is now clear that although mutations in *BRCA1* and *BRCA2* account for the majority of high-risk breast cancer families (85%) and nearly all breast/ovarian families, in smaller aggregations they account for <50% of the hereditary element (Ford *et al*., 1998). While there is no doubt that *BRCA1/2* are highly penetrant genes, initial estimates of the lifetime risk of 85% (Easton *et al*., 1993) appear slightly high. Population studies do detect *BRCA1* and *BRCA2* mutations in blood samples from apparently sporadic breast cancer patients (Peto *et al*., 1999). Furthermore, founder mutations in the Jewish and Icelandic populations where *BRCA1/2* mutation frequencies can be as high as 2–2.5%, are associated with lifetime risks of breast cancer of 40–60% (Struewing *et al*., 1997). Outside populations with significant founder effects the frequencies of *BRCA1/2* mutations combined is probably no higher than 0.2%.

Having identified the most important high penetrance genes, the search is on for lower penetrance genes. Aggregation of breast cancer has been shown to occur in ataxia telangiectasia heterozygotes (Swift *et al*., 1987, 1991), who are the carriers of the recessive gene which causes a disease which predisposes especially to haematological malignancy in childhood. A mother of an affected child is at 3–5-fold risk of breast cancer, which would fulfil a dominant gene model with 25–40% penetrance and a population frequency of about 0.01. Since the isolation of the *ATM* gene (Savitsky *et al*., 1995) there have been conflicting studies as to whether this gene is a significant cause of breast cancer. Breast cancer is also thought to occur in 30% of women with Cowden's disease (a condition predisposing to multiple hamartomas), but since the discovery of the underlying gene defects in the *PTEN* gene, no studies have found the gene to be involved in familial aggregations of breast cancer.

Colon Cancer

It has been estimated that about 8% of colorectal cancer is due to the inheritance of a dominant predisposing gene (Solomon, 1990). Only a small proportion of this subset is due to FAP and there are perhaps 10 times as many individuals born with a gene for so-called hereditary nonpolyposis colorectal cancer (HNPCC). This latter inherited form of colorectal cancer has been outlined by Lynch and has in the past often been further subdivided into Lynch syndrome I and II. The type I form was considered to be site specific and to particularly predispose to proximal tumours, which has a major bearing on screening (Lynch *et al*., 1988). Type II was considered to predispose gene carriers to endometrial, ovarian, pancreatic, upper urinary tract and stomach cancers as well as colorectal and multiple tumours (Lynch *et al*., 1985). The peak age for these cancers to occur is the fifth decade with proximal colorectal tumours in two-thirds of cases manifesting this complication. The cancers are probably preceded by polyp development, but the mucosa is not lined with hundreds of them as in FAP.

It had been assumed that the gene or genes predisposing to HNPCC were tumour suppressor genes. In 1988, Vogelstein *et al*. (1988) proposed the classic model of progression to cancer in which several genes were involved starting with loss of tumour suppressor genes. Activation of oncogenes at a later stage such as the *ras* genes is also important, but it had not been considered that it was the rate of mutation that was the key factor in HNPCC. The most important tumour suppressor in colorectal cancer, the *APC* gene on chromosome 5 (that causes FAP), had been cloned in 1991 (Nishisho *et al*., 1991). Clues from tumour studies led to the isolation of several other tumour suppressor genes which are important in colorectal cancer development. The earliest of these was the *TP53* gene, whose

position on chromosome 17 was first implicated in 1981. The *DCC* gene on chromosome 18 was shown to be deleted in some colorectal cancers in 1990 and the *MCC* gene on chromosome 5 in 1991. None of these genes was found to be mutated in the germ-line of familial cases of colorectal cancer. The only report of positive linkage had been that of several Lynch type II families to the Kidd blood group on chromosome 18. The major breakthrough came from yeast genetics in which genes involved in repair of DNA suddenly became major candidates for human disease. For some time it had been noticed that tumours in HNPCC and sporadic patients showed instability in the DNA, which was manifested as a different size of microsatellite repeat in the tumour DNA compared with the blood. The discovery that a human version of a yeast DNA mismatch repair gene *MSH2* mapped to recently linked locus on chromosome 2 quickly led to the identification of the first HNPCC gene (Fishel *et al*., 1993). This then allowed a candidate gene approach, which proved that another mismatch repair gene, *MLH1*, was also an important contributor to HNPCC (Bronner *et al*., 1994). Since that time, two further genes, *PMS1* and *PMS2*, have been implicated in a tiny proportion of families and more recently *MSH6* has been found to cause HNPCC and also families with endometrial cancer. Although the mismatch repair genes are inactivated in both copies like a tumour suppressor gene, the mechanism to cancer development is different. Although it had been thought that as much as 13% of colorectal cancer could be due to HNPCC, it is now clear that the mismatch repair genes and in particular *MLH1* become inactivated somatically sometimes by methylation. More realistic estimates of HNPCC are therefore that it accounts for 1–2% of colorectal cancer. This means there are significant unidentified genes yet to be discovered.

Ovarian Cancer

Ovarian cancer, like breast and colon cancer, has had many reports of familial aggregation dating back at least to 1950. Increased risk of ovarian malignancy may be inherited as part of several genetic conditions. Gorlin syndrome (Strong, 1977) Peutz–Jeghers syndrome and XY females are all at heightened risk. In addition to this, ovarian cancer is part of HNPCC and breast/ovarian aggregation (now known to be due to *BRCA1* and *BRCA2*). There have been several reports of familial site-specific ovarian cancer (Fraumeni *et al*., 1975), but many contain cases of breast and other malignancies. The association of breast and ovarian cancer in both family reports and epidemiological studies of breast (Ridolfi *et al*., 1977) and ovarian cancer (Schildkraut *et al*., 1988) had suggested the presence of an autosomal dominant predisposing gene.

Tumour studies are less numerous than in either the breast or colon, but loss of constitutional heterozygosity has been found on chromosome regions: 3p, 6q, 11p, 13q, 17p, 17q and Xp. Following the isolation of *BRCA1* and *BRCA2*, it became clear that two of these loci (17q and 13q) were significant for hereditary disease and that apparently site-specific ovarian cancer was mainly caused by *BRCA1* (Steichen-Gersdorf *et al*., 1994). Whether this predisposition is mainly to ovary or breast may depend on the position of the mutation in each gene (Gayther *et al*., 1995). It is now thought unlikely that there is a significant other ovarian cancer gene (Ford *et al*., 1998). Predictive tests are now possible in many families once the underlying mutation has been identified. Current evidence suggests that about 40–60% of women at risk in these families will opt for testing and >50% will opt for prophylactic surgery for the ovaries and or breasts (Meijers-Heijboer *et al*., 2000).

CONCLUSIONS

The last 10 years has seen an enormous advance in our understanding of cancer and its familial elements. A great deal of this knowledge derives from the study of rare cancer-predisposing syndromes. This research is not esoteric because the cloning of these genes will benefit not only the small proportion of people who suffer from these conditions but also those who suffer from the common cancers occurring in these syndromes. Gene therapy, which could be directed at replacing the function of a deleted tumour suppressor gene, may be available in the early part of the twenty-first century. From the diseases in the first section this could be applied to cancer of the colon, skin, kidney and thyroid as well as virtually all common brain tumours.

The cloning of further tumour suppressor genes for breast, colon and ovarian cancer will have major implications in the treatment of these common cancers. The possibility of preventive treatment in high-risk families is also a real hope.

Currently, predictive genetic tests are available at some specialist genetic centers for FAP, NF1, NF2, von Hippel–Lindau disease, MEN1, MEN2 and tuberous sclerosis, HNPCC, BRCA1 and BRCA2 and can be combined with clinical screening protocols (**Table 3**). These may depend on a suitable family structure (blood samples for linkage are needed on the extended family including two affected cases) as even when the genes are cloned no guarantee can be made of identifying the underlying mutation (see NF1/BRCA1/2/HNPCC). Faster and more sensitive methods of gene mutation identification and sequencing will open the way for more readily available mutation studies in these cloned genes. Even while this chapter is being published further discoveries will be made and much of what is written here will be superseded in 5 years or so. The preceding sections should, however, give the reader a good grasp of the current state of the art and the discoveries that have brought it about.

Table 3 Chromosomal location and implications of various dominant cancer syndromes

Disease	Location	Tumours	Probable earliest tumour (years)	Risk in lifetime (%)	Start of screening (years)
FAP	5q	Adenomas	1st	100	10–16
		Bowel cancer	?4, 7	99	
NF1	17q	Neurofibroma	1st	100	Birth
		Glioma, sarcoma	1st	10	
NF2	22q	Schwannomas	1st	100	Birth
		Meningiomas	1st	60	
		Gliomas		10	
vHL	3p	Haemangioblastoma	1–2	90	5
		Renal carcinoma	20	70	15
MEN1	11q	Parathyroid, insulinoma, gastrinoma	5	95	5
MEN2a	10q	Medullary thyroid cancer, parathyroid, phaeochromocytoma	3	80	3–4
MEN2b	10q	As in MEN2a, except parathyroid	1	100	Birth
Gorlin	9q	Basal cell carcinoma	5	90	Birth
		Medulloblastoma	1	5	
Cowden	10q	Breast	30	30	35
		Thyroid			
LFS	17p	Sarcoma (bone/soft tissue), adrenal, breast cancer, gliomas.	1st	95	1st
BRCA1	17q	Breast, ovary, colon, prostate carcinoma	>16	80–90	30
BRCA2	13q	Breast, ovary, colon, prostate carcinoma	>16	80–90	30
		Male breast		10	
HNPCC	2p,3p 2q,7p	Colorectum, ovary, endometrium, ureter, gastric, pancreas	>16	80	25

REFERENCES

Allanson, J. E., *et al*. (1991). Watson syndrome: is it a subtype of type 1 neurofibromatosis. *Journal of Medical Genetics*, **28**, 752–756.

Bell, D. W., *et al*. (1999). Heterozygous germ line hCHK2 mutations in Li–Fraumeni syndrome. *Science*, **286**, 2528–2531.

Birch, J. M., *et al*. (1998). Cancer phenotype correlates with constitutional TP53 genotype in families with Li–Fraumeni syndrome. *Oncogene*, **17**, 1061–1068.

Bodmer, N. F., *et al*., (1987). Localisation of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, **328**, 614–616.

Bronner, C. E., *et al*. (1994). Mutation in the DNA mismatch repair gene homolog hMLH1 is associated with hereditary non-polyposis colorectal cancer. *Nature*, **368**, 258–261.

Burn, J., *et al*. (1991). The UK Northern Region genetic register for familial adenomatous polyposis coli: use of age of onset, congenital hypertrophy of the retinal pigment epithelium, and DNA markers in risk calculation. *Journal of Medical Genetics*, **28**, 289–296.

Chandrasekharappa, S. C., *et al*. (1997). Positional cloning of the gene for multiple endocrine neoplasia type 1. *Science*, **276**, 404–407.

Clark, R. D. (1994). Proteus syndrome. In: Huson, S. M. and Hughes, R. A. C. (eds). *The Neurofibromatoses* 402–413 (Chapman and Hall, London).

Claus, E. B., *et al*. (1990). Age of onset as an indicator of familial risk of breast cancer. *American Journal Epidemiology*, **131**, 961–972.

Cohen, A. J., *et al*. (1979). Hereditary renal cell carcinoma associated with a chromosomal translocation. *New England Journal of Medicine*, **301**, 592–595.

Colley, A., *et al*. (1996). Neurofibromatosis/Noonan phenotype: a variable feature of Type 1 neurofibromatosis. *Clinical Genetics*, **49**, 59–64.

Cowan, R., *et al*. (1997). The gene for the Naevoid Basal Cell Carcinoma (Gorlin) syndrome acts as a tumour suppressor gene in medulloblastoma. *British Journal of Cancer*, **76**, 141–145.

Cushing, H. (1917). *Tumours of the Nervus Acusticus and the Syndrome of the Cerebello-pontile Angle*. (W.B. Saunders, Philadelphia).

Davies, D. R., *et al*. (1995). Severe Gardner's syndrome in families with mutations restricted to a specific region of the APC gene. *American Journal of Human Genetics*, **57**, 1151–1158.

DeVries, P. J. and Bolton, P. F. (2000). Genotype–phenotype correlations in tuberous sclerosis. *Journal of Medical Genetics*, **37**, E3.

Easton, D., *et al*. (1993). Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. *American Journal of Human Genetics*, **53**, 305–313.

Eeles, R. A., *et al*. (1996). *Genetic Predisposition to Cancer*. (Chapman and Hall, London).

European Chromosome 16 Tuberous Sclerosis Consortium (1993). Identification and characterisation of the tuberous sclerosis gene on chromosome 16. *Cell*, **75**, 805–808.

Evans, D. G. R., *et al*. (1992a). A genetic study of type 2 neurofibromatosis in the United Kingdom: I prevalence, mutation rate, fitness and confirmation of maternal transmission effect on severity. *Journal of Medical Genetics*, **29**, 841–846.

Evans, D. G. R., *et al*. (1992b). A clinical study of type 2 neurofibromatosis. *Quarterly Journal of Medicine*, **84**, 603–618.

Evans, D. G. R., *et al*. (1993). Complications of the Naevoid Basal Cell Carcinoma Syndrome: results of a population based study. *Journal of Medical Genetics*, **30**, 460–464.

Evans, D. G. R., *et al*. (1998a). Genotype–phenotype correlations in type 2 neurofibromatosis (NF2): Evidence for more severe disease with truncating mutations. *Journal of Medical Genetics*, **35**, 450–455.

Evans, D. G. R., *et al*. (1998b). Somatic mosaicism: a common mechanism for sporadic disease in tumour prone syndromes? Lessons from type 2 neurofibromatosis. *American Journal of Human Genetics*, **63**, 727–736.

Farndon, P., *et al*. (1992). Localisation of the gene for Gorlin (Naevoid basal cell carcinoma) syndrome on chromosome 9. *Lancet*, **i**, 581–582.

Fishel, F. S., *et al*. (1993). A mutator gene homolog MSH2 and its association with hereditary non-polyposis colorectal cancer. *Cell*, **260**, 1027–1038.

Ford, D., *et al*. (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *American Journal of Human Genetics*, **62**, 676–689.

Francke, U. and King, F. (1976). Sporadic bilateral retinoblastoma and $13q^{-}$ chromosomal deletion. *Medical Pediatrics and Oncology*, **2**, 379–380.

Fraumeni, J. F., *et al*. (1975). Six families prone to ovarian cancer. *Cancer*, **36**, 364–369.

Friend, S. H., *et al*. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*, **323**, 643–646.

Fryer, A. E., *et al*. (1987). Evidence that the gene for tuberous sclerosis is on chromosome 9. *Lancet*, **i**, 659–661.

Gailani, M., *et al*. (1991). Evidence for a tumour suppressor gene on chromosome 9 in basal cell carcinomas of the skin. *American Journal of Human Genetics*, **49**, Supplement, 454.

Gardner, E. J. (1951). A genetic and clinical study of intestinal polyposis, a predisposing factor for carcinoma of the colon and rectum. *American Journal of Human Genetics*, **3**, 167–176.

Gardner, W. J. and Frazier, C. H. (1930). Bilateral acoustic neurofibromas: a clinical study and field survey of a family of five generations with bilateral deafness in thirty eight members. *Archives of Neurology Psychiatry*, **23**, 266–302.

Gayther, S. A., *et al*. (1995). Germline mutations of the BRCA1 gene in families: evidence for a genotype/phenotype correlation. *Nature Genetics*, **11**, 428–433.

Godbout, R., *et al*. (1983). Somatic inactivation of genes on chromosome 13 is a common event in retinoblastoma. *Nature*, **304**, 451–453.

Gorlin, R. J. (1987). Naevoid basal cell carcinoma syndrome. *Medicine*, **66**, 98–113.

Grundy, P., *et al*. (1988). Familial predisposition to Wilms tumour does not map to the short arm of chromosome 11. *Nature*, **336**, 374–376.

Hahn, H., *et al*. (1996). Mutations of the human homolog of *Drosophila* patched in the naevoid basal cell carcinoma syndrome. *Cell*, **85**, 841–851.

Hall, J. M., *et al*. (1990). Linkage of early onset familial breast cancer to 17q21. *Science*, **250**, 1684–1689.

Herrera, L., *et al*. (1986). Gardner syndrome in a man with an interstitial deletion of 5q. *American Journal of Medical Genetics*, **25**, 473–476.

HMSO (1998). *Cancer Statistics Registration England and Wales 1992*. (London, HM Stationery Office).

Knudson, A. G. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the USA*, **68**, 820–823.

Koufos, A., *et al*. (1989). Familial Wiedemann–Beckwith syndrome and a second Wilms tumour locus both, map to 11p15.5. *American Journal of Human Genetics*, **44**, 711–719.

Lam, W. W., *et al*. (1999). Analysis of CDKN1C (p57KIP2) mutations in familial and sporadic Beckwith–Wiedemann syndrome (BWS) provides a novel genotype–phenotype correlation. *Journal of Medical Genetics*, **36**, 518–523.

Larsson, C., *et al*. (1988). Multiple endocrine neoplasia type 1 gene maps to chromosome 11 and is lost in insulinoma. *Nature*, **332**, 85–87.

Latif, F., *et al*. (1993). Identification of the von Hippel–Lindau disease tumour suppressor gene. *Science*, **260**, 1317–1320.

Li, F. P. and Fraumeni, J. F., Jr (1969). Soft tissue sarcomas, breast cancer, and other neoplasms: a familial syndrome? *Annals of Internal Medicine*, **71**, 747–752.

Li, F. P., *et al*. (1988). A cancer family syndrome in 24 kindreds. *Cancer Research*, **48**, 5358–5362.

Listernick, R., *et al*. (1989). Optic gliomas in children with neurofibromatosis type 1. *Journal of Pediatrics*, **114**, 788–792.

Liu, Z., *et al*. (1995). Germline mutations in the RB1 gene in patients with hereditary retinoblastoma. *Genes Chromosomes Cancer*, **14**, 277–284.

Lohmann, D. R., *et al*. (1996). Spectrum of RB1 germ-line mutations in patients with hereditary retinoblastoma. *American Journal of Human Genetics*, **58**, 940–949.

Lynch, H. T., *et al*. (1985). Hereditary nonpolyposis colorectal cancer (Lynch syndromes I and II): Clinical description of resource. *Cancer*, **56**, 934–938.

Lynch, H. T., *et al*. (1988). Differential diagnosis of hereditary nonpolyposis colorectal cancer (Lynch syndrome I and

Lynch syndrome II). *Diseases of the Colon and Rectum*, **31**, 372–377.

Maher, E. R., *et al*. (1990). Clinical features and natural history of von Hippel–Lindau disease. *Quarterly Journal of Medicine*, **77**, 1151–1163.

Malkin, D., *et al*. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas and other neoplasms. *Science*, **250**, 1233–1238.

McGaughran, J. M., *et al*. (1999). A clinical study of type 1 neurofibromatosis in North West England. *Journal of Medical Genetics*, **36**, 197–203

Meijers-Heijboer, E. J., *et al*. (2000). Presymtomatic DNA testing and prophylactic surgery in families with a BRCA1 or BRCA2 mutation. *Lancet*, **355**, 2015–2020.

Messiaen, L. M., *et al*. (2000). Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Human Mutation*, **15**, 541–545.

Miki, Y., *et al*. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 120–122.

Mulligan, L. M., *et al*. (1993). Germline mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2a. *Nature*, **363**, 458–460.

Narod, S. A., *et al*. (1991). A familial breast–ovarian cancer locus on chromosome 17q12–q23. *Lancet*, **ii**, 82–83.

National Institutes of Health Consensus Development Conference Statement on Neurofibromatosis. (1987). *Archives of Neurology*, **45**, 575–579.

Nishisho, I., *et al*. (1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science*, **253**, 665–669.

Pelletier, J., *et al*. (1991). WT1 mutations contribute to abnormal genital system development and hereditary Wilms tumour. *Nature*, **353**, 431–434.

Peto, J., *et al*. (1999). Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *Journal of National Cancer Institute*, **91**, 943–949.

Ridolfi, R. L, *et al*. (1977). Medullary carcinoma of the breast. *Cancer*, **40**, 1365–1385.

Riccardi, V. M., *et al*. (1978). Chromosomal imbalance in aniridia–Wilms tumour association: 11p interstitial deletion. *Pediatrics*, **61**, 604–610.

Rouleau, G., *et al*. (1987). Genetic linkage analysis of bilateral acoustic neurofibromatosis to a DNA marker on chromosome 22. *Nature*, **329**, 246–248.

Rouleau, G. A., *et al*. (1993). Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2. *Nature*, **363**, 515–521.

Santibanez-Koref, M. F., *et al*. (1991). P53 Germline mutations in 2 out of 8 families with the Li Fraumeni syndrome. *Lancet*, **ii**, 1490–1491.

Savitsky, K., *et al*. (1995). A single ataxia telangectasia gene with a product similar to PI3-kinase. *Science*, **268**, 1749–1753.

Schildkraut, J. M. and Thompson, W. D. (1988). Relationship of epithelial ovarian cancer to other malignancies within families. *Genetics and Epidemiology*, **5**, 355–367.

Scott, R., *et al*. (1996). Familial infiltrative fibromatosis (Desmoid tumours) caused by a recurrent 3′ APC mutation. *Human Molecular Genetics*, **5**, 1921–1924.

Seizinger, B. R., *et al*. (1987). Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell*, **49**, 589–594.

Seizinger, B. R., *et al*. (1988). Von Hippel–Lindau disease maps to the region of chromosome 3 associated with renal cell carcinoma. *Nature*, **332**, 268–269.

Shrivastava, S., *et al*. (1990). Germ-line transmission of a mutated p53 gene in a cancer prone family with Li–Fraumeni syndrome. *Nature*, **348**, 747–749.

Simpson, N. E., *et al*. (1987). Assignment of multiple endocrine neoplasia type 2A to chromosome 10 by linkage. *Nature*, **328**, 528–530.

Solomon, E. (1990). Colorectal cancer genes. *Nature*, **343**, 412–414.

Solomon, E., *et al*. (1987). Chromosome 5 allele loss in human colorectal carcinomas. *Nature*, **328**, 616–619.

Spirio, L., *et al*. (1993). Alleles of the APC gene: an attenuated form of familial polyposis. *Cell*, **75**, 951–957.

Steichen-Gersdorf, E., *et al*. (1994). Familial site specific ovarian cancer is linked to BRCA1 on 17q12–21. *American Journal of Human Genetics*, **55**, 870–875

Strong, L. C. (1977). Genetic and environmental interactions. *Cancer*, **40**, 1861–1866.

Struewing, J. P., *et al*. (1997). The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *New England Journal of Medicine*, **336**, 1401–1408.

Swift, M., *et al*. (1987). Breast and other cancers in families with ataxia telangiectasia. *New England Journal of Medicine*, **316**, 1289–1294.

Swift, M., *et al*. (1991). Incidence of cancer in 161 families affected by ataxia telangiectasia. *New England Journal of Medicine*, **325**, 1831–1836.

Teyssier, J. R., *et al*. (1986). Recurrent deletion of the short arm of chromosome 3 in human renal cell carcinomas: shift of the c-raf1 locus. *Journal of the National Cancer Institute*, **77**, 1187–1191.

Toguchida, J., *et al*. (1991). A survey of germ-line and somatic p53 mutations in patients with bone and soft tissue sarcomas. *American Journal of Human Genetics*, **49**, Supplement, 458.

Trofatter, J. A., *et al*. (1993). A novel Moesin-, Ezrin-, Radixin-like gene is a candidate for the neurofibromatosis 2 tumour suppressor. *Cell*, **72**, 1–20.

Varley, J. M., *et al*. (1999). Analysis of a panel of patients with childhood adrenocortical tumours for germline TP53 mutations. *American Journal of Human Genetics*, **65**, 995–1006.

Viskochil, D., *et al*. (1990). Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell*, **62**, 187–192.

Vogelstein, B., *et al*. (1988). Genetic alterations during colorectal tumour development. *New England Journal of Medicine*, **319**, 525–532.

Williams, W. R. and Strong, L. C. (1985). Genetic epidemiology of soft tissue sarcoma in children. In: Muller H. R. and

Weber, W. (eds), *Familial Cancer, First International Research Conference*. 151–153 (Karger, Basle).
Wishart, J. H. (1822). Case of tumours in the skull, dura mater, and brain. *Edinburgh Medical and Surgical Journal*, **18**, 393–397.
Wooster, R., *et al*. (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12–13. *Science*, **265**, 2088–2090.
Wooster, R., *et al*. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378**, 789–792.

FURTHER READING

Bishop, D. T. and Thomas, H. J. W. (1990). The genetics of colorectal cancer. *Cancer Surveys*, **9**, 585–604.
Claus, E. B., *et al*. (1994). Autosomal dominant inheritance of early onset breast cancer. *Cancer*, **73**, 643–651.
Devilee, P. J. and Cornelisse, C. J. (1990). Genetics of breast cancer. *Cancer Surveys*, **9**, 605–630.
Eeles, R., *et al*. (eds) (1996). *Genetic Predisposition to Cancer*. (Chapman and Hall, London).
Evans, D. G. R. (1995). Practical implications of the new cancer genetics. In: Peckham, M., *et al*. (eds), *Oxford Textbook of Oncology*. (Oxford University Press, Oxford).
Hodgson, S. and Foulkes W. D. (1998). *Inherited Susceptibility to Cancer*. 3–19. (Cambridge University Press, Cambridge).
Huson, S. M., *et al*. (1989). A genetic study of von Recklinghausen neurofibromatosis in south east Wales. II. Guidelines for genetic counselling. *Journal of Medical Genetics*, **26**, 712–721.
Lemoine N. R., *et al*. (1994). *Cancer: A Molecular Approach*. (Blackwell Scientific, Oxford).
Varley, J., *et al*. (1997). Li–Fraumeni syndrome – A molecular and clinical review. *British Journal of Cancer*, **76**, 1–14.

Chapter 5

Human DNA Tumour Viruses

Beverly E. Griffin and Wilson Caparrós-Wanderley
Imperial College School of Medicine at St. Mary's, London, UK

CONTENTS

- Introduction
- Papillomaviruses (PVs)
- Epstein–Barr Virus (EBV)
- Kaposi Sarcoma-associated Herpesvirus (KSHV)
- Hepatitis B Virus (HBV)
- Other DNA Viruses

INTRODUCTION

Almost all vertebrate species, whether man, monkey or marsupial, have their own cohort of viruses. These are usually species specific, although they may share sequence homologies, genes and gene functions with similar viruses from other species. In practice, this has meant that a detailed study of a virus from one species may have profound impact on predicting the properties of a similar virus in another species, particularly where the viral sequences are available for comparison. However, this is not always the case, as best illustrated for two small, highly related DNA tumour viruses, the primate virus, simian virus 40 (SV40), and the mouse virus, polyomavirus (Py). For SV40, all activities for regulating cell growth, and transformation to a tumorigenic phenotype, are carried out by one viral protein, the large T-antigen (LT). However, with Py, differential splicing of the RNA, made from a single region of the viral genome, gives rise to distinct messages and two proteins – the so-called large and middle T-antigens (LT and MT) – where the function for altering cell growth is carried by one protein, LT, and that for inducing cellular transformation by the other, MT. Thus, cross-species speculations about related viruses need to be made with care. Similarly, the general notions about host range specificity of a virus cannot be deemed absolute, since mutations can occur that result in host alteration. One of the most interesting cases in point, in the DNA virus field, comes from the identification in Africa of a pox virus that appears recently to have crossed the species barrier from monkey to man. The current interest in 'emerging viruses' focuses on the origins of new species and evolutionary and environmental factors that contribute to their birth (Morse, 1993).

This chapter deals with a subset of DNA viruses, those associated with tumour formation, and is restricted to discussing in detail only those viruses where a good case can be made for an association with human cancer. In animal models, such as those employing SV40 and Py, the experimental evidence showing tumour causality in appropriate models is unambiguous. For human disease, the situation is by definition more complicated and causality or association depends on drawing together studies from many fields, including not only virology, but also epidemiology, oncology and molecular biology. Notably, with the four viruses that are discussed in detail, that is, papilloma viruses, two herpesviruses (Epstein–Barr and Kaposi sarcoma-related virus) and hepatitis B virus, all have the ability to persist in infected cells and evade host immune systems. For none of these viruses has a very good *in vitro* lytic system been identified, impeding progress in answering crucial questions about them. Even without this, much progress has been made, mainly owing to the judicious use of molecular methods, and model systems, where identified.

Over the last few years, the International Agency for Research on Cancer (IARC), Lyon, as part of their programme for evaluating carcinogenic risks to humans, has considered four of the human DNA viruses most likely to play a causal, predisposing, or auxiliary role in the development of cancer in humans. These are various members of the papillomavirus (PV) family (related to SV40 and mouse Py, and in former times considered to belong to the same papovavirus family), hepatitis B virus (HBV), and the two members, Epstein–Barr virus (EBV) and Kaposi sarcoma-associated virus (KSHV), of the herpesvirus family. The general conclusion reached by a consortium of individuals contributing to reports on three of these viruses (HPV, HBV and EBV) is that 'there is sufficient evidence' for their carcinogenicity, as it relates to defined forms of cancer (IARC, 1994, 1995, 1997). For KSHV, the most recently identified of these viruses, the conclusion reached in 1998 is that 'KSHV is probably carcinogenic to humans.' The ultimate proof of viral causality of malignancy will be concomitant eradication of the virus and of the disease, and is a target for the future. In

the case of HBV, where there are effective antiviral vaccines, this could be realized in the foreseeable future.

PAPILLOMAVIRUSES (PVs)

General Definition and Classification

Papillomaviruses (PVs) are a family of DNA viruses that cause hyperproliferative lesions of the mucosal and cutaneous epithelia (papillomas, warts and condylomas) in a wide variety of higher vertebrates, including humans. Most of these lesions are benign, self-limiting and regress with time, but some of them tend to progress towards malignancy and invasive carcinoma (e.g. carcinoma of the uterine cervix).

All PVs belong to the subfamily Papillomavirus, which constitutes one of the two members of the family Papovaviridae. The other member of this family, the subfamily Polyomavirus, is discussed later in this chapter. PVs and Polyomaviruses were initially grouped together because they share properties of small-sized, nonenveloped virions, icosahedral capsids, superhelical double-stranded DNA genomes, and use the nucleus as site of multiplication. Subsequent research has shown that, despite these similarities, the two genera are not evolutionarily related. They have different genomic organisations, their DNAs do not hybridize and there is no immunological cross-reactivity between them. Furthermore, in contrast to Polyomaviruses, PVs multiply only in differentiating epithelium and cannot be propagated *in vitro* (Howley, 1996).

PVs are highly species specific, hence their classification is based on their host range and DNA relatedness. Each virus is first named after its natural host followed by a number, and sometimes a letter, which indicates, respectively, its type and subtype (e.g. bovine (B)PV-4, Human (H)PV-6b, etc.) (**Table 1**). Classification of different isolates from one species into types and subtypes is based, at present, on their degree of sequence homology. On the basis of the site of infection, HPVs have also been classified into two main groups: cutaneous and mucosal. Each group can, in turn, be subdivided into 'high-' or 'low'-risk types according to the probability of malignant progression associated with the type of lesions they cause. Both this approach and the sequence homology method give rise to equivalent phylogenetic trees (**Figure 1**) (Shah, 1990).

Table 1 PV-associated pathological conditions

Species	Pathology	Virus Type[a]
Deer	Cutaneous fibropapillomas	DPV
Cattle	Alimentary tract carcinoma	**BPV-4**
	Cutaneous fibropapillomas	BPV-2
Cottontail rabbit	Skin carcinomas	**CRPV**
Humans	Skin warts	HPV-1, -2, -3, -7 and -10
	Epidermodysplasia verruciformis	HPV-**5**, **-8**, -17 and -20
	Anogenital warts (condylomas):	
	Exophytic condylomas	HPV-6 and -11
	Flat condylomas	HPV-**16**, **-18**, **-31**, **-33**, -42, **-43**, etc.
	Respiratory tract papillomas	HPV-6 and -11
	Conjunctival papillomatosis	HPV-6 and -11
	Focal epithelial hyperplasia (FEH)	HPV-13 and -32

[a]Virus types predominantly recovered from malignant lesions are indicated in bold.

Virion Structure

Nonenveloped icosahedral PVs replicate in the nucleus of squamous epithelial cells. PV particles are about 50 nm in diameter and encapsulate a single copy of the circular 8 kbp double-stranded DNA genome in the form of a chromatin-like complex with cellular histones. They have a density in caesium chloride of 1.34–1.36 g L^{-1} and, owing to the lack of lipids, are resistant to ether and other solvents (Pfister and Fuchs, 1994).

The viral capsid is composed of 72 capsomers with a star-shaped morphology and displaying a cylindrical channel along their axis (**Figure 2**). All capsomers are pentamers of the L1 protein, a 55-kDa protein which represents about 80% of the total capsid protein. L1 protein is required for virus attachment to the cell surface receptor and constitutes the basic structural component of the capsid. The remaining 20% of the capsid is composed of a 70-kDa protein known as L2. The exact function(s) of this protein is still unclear, but it may be involved in the efficient self-assembly of the viral capsid and attraction and/or proper positioning of the viral genome during viral assembly (Howley, 1996).

Genomic Organisation

The Papillomavirus genome is divided into an 'early' region (about 4.5 kbp in size), a 'late' region (about 2.5 kbp) and a long regulatory region (LCR) (about 1 kbp). There are two open reading frames (ORFs) in the late region (L1 and L2) and up to eight ORFs (E1 to E8) in the early region. There are no ORFs in the LCR, but this region contains the viral origin of replication and control elements for transcription and replication. In contrast to Polyomaviruses, all ORFs in the Papillomavirus

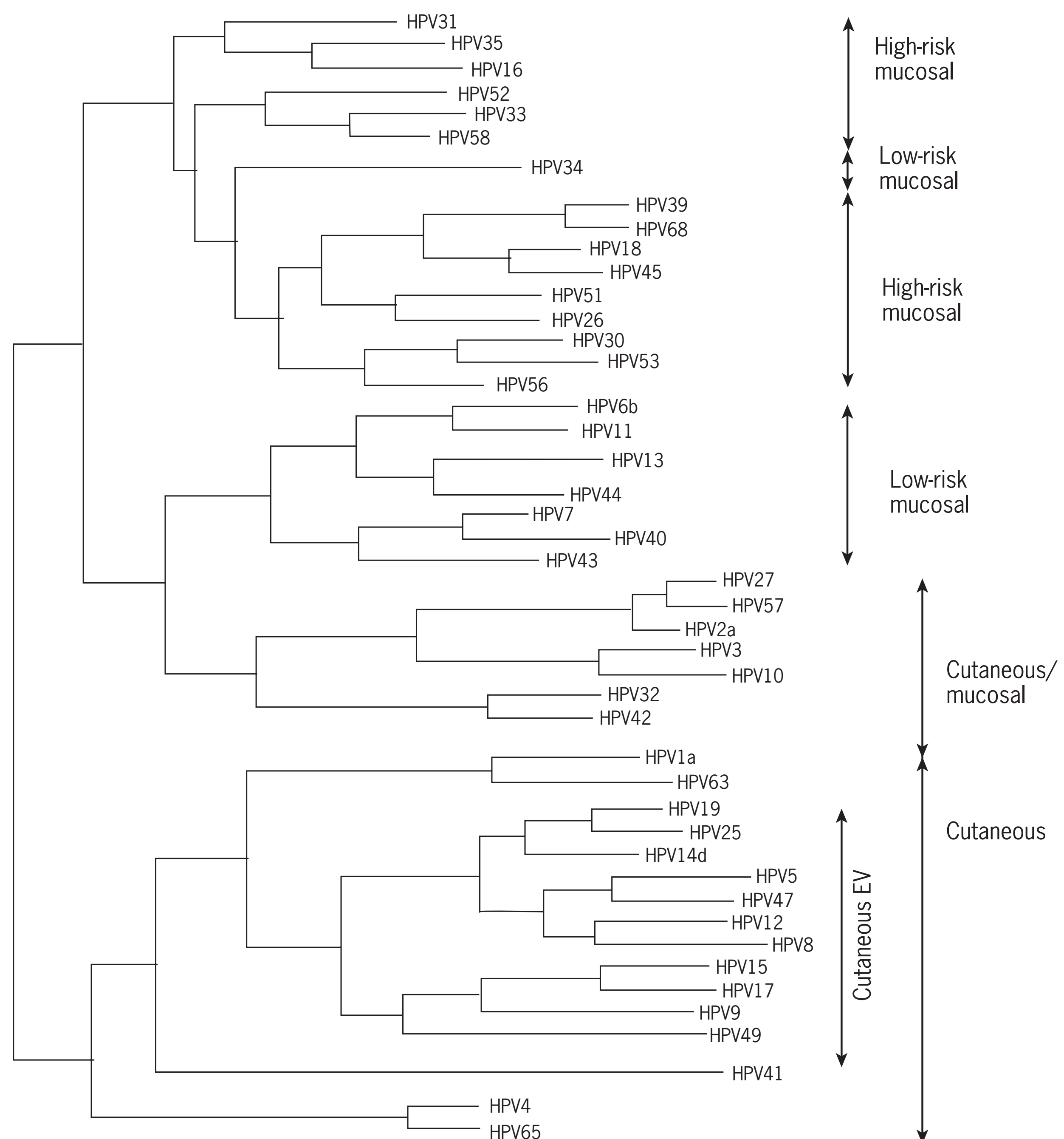


Figure 1 Phylogenetic tree of HPVs, based on DNA sequence homology of a 384-bp fragment of the E6 ORF. The clinical classification of HPV types, according to the site of infection, is indicated on the right. (Adapted from Van Ranst *et al.*, 1992.)

genome are located on one DNA strand (**Figure 3**) (Shah, 1990).

The properties of the proteins encoded by the two late ORFs, L1 and L2, have been described in the previous section.

The **E1** proteins, molecular weights (MW) 68–85 kDa, are essential for viral replication and in this role they are similar to the SV40 large T-antigen. They are phosphoproteins with DNA-dependent ATPase and ATP-dependent helicase activities. The 5′ portion of the E1 ORF can sometimes be translated as a smaller protein involved in modulation of viral replication (Chow and Broker, 1994).

The **E2** ORF codes for a family of proteins of which only the full-length member, MW 43–48 kDa, can support viral replication. Full-length E2 is also a transcriptional activator whilst truncated forms of E2 (also known as E2C

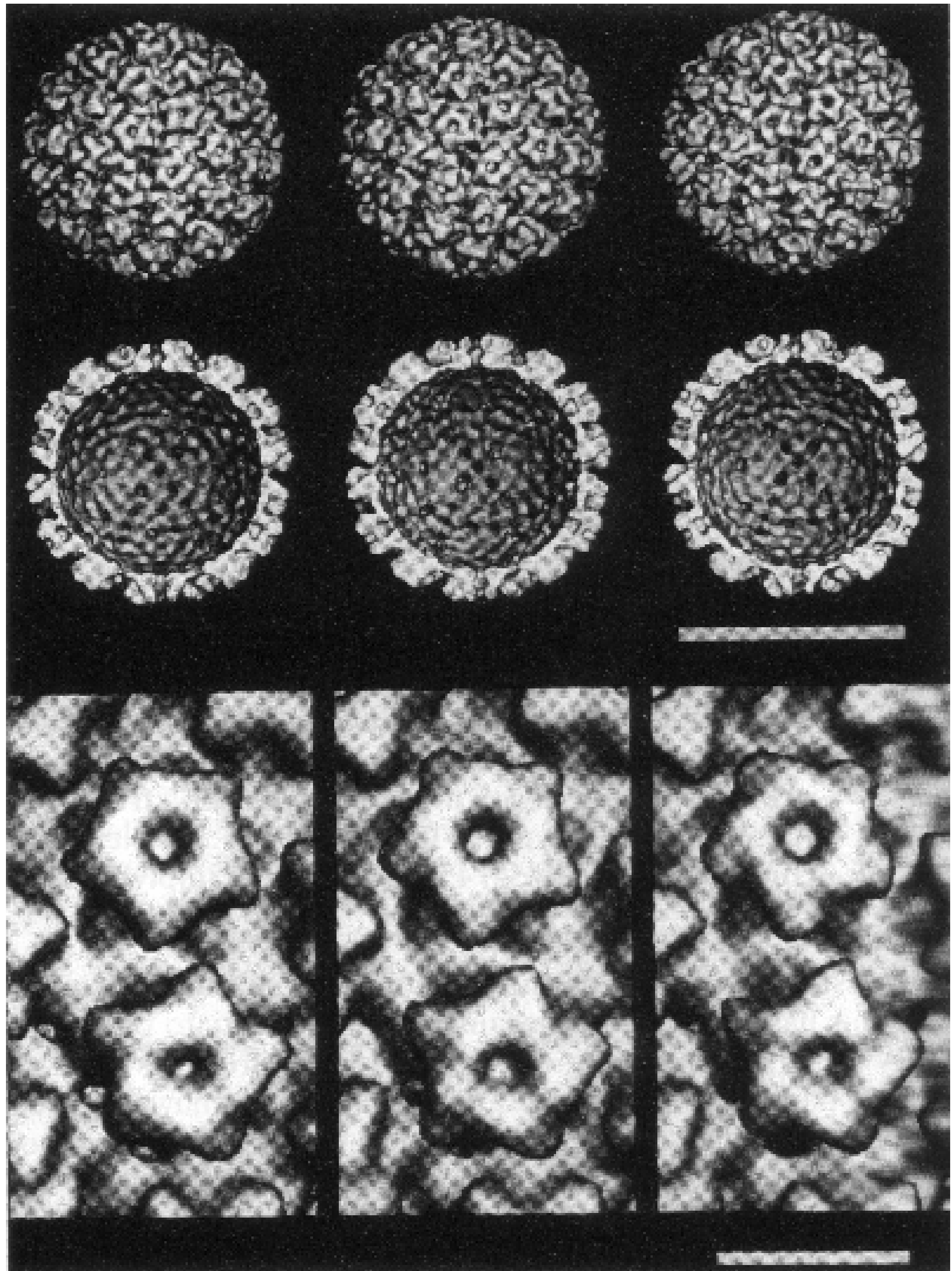


Figure 2 Structure of HPV. Surface-shaded representations of reconstructions of HPV-1 from warts (left columns), L1 capsids (middle columns) and L1L2 capsids (right columns). Outside view of capsids (top row), inside view (middle row) and a close-up view of pentavalent and hexavalent capsomeres (bottom row). Internal density to a radius of 20 nm was computationally removed to show internal features of the capsid. No differences are apparent. Bars = 50 nm (top and middle rows) and 10 nm (bottom row). (From Hagensee *et al.*, 1994, *Journal of Virology*, **68**, 4503–4505.)

and E2M), derived from different promoter usage and alternative mRNA splicing, are transcriptional repressors (Arrand, 1994; Chow and Broker, 1994).

The **E3** ORF is only present in some PVs and its function is not yet known.

The **E4** protein is expressed from a spliced mRNA (E1^E4) as a doublet of MW 17 and 16 kDa. E4 appears to be involved in the disruption of the cytokeratin network during the late phase of the viral life cycle. In HPV1 induced warts, E4 accumulates abundantly (up to 30% of

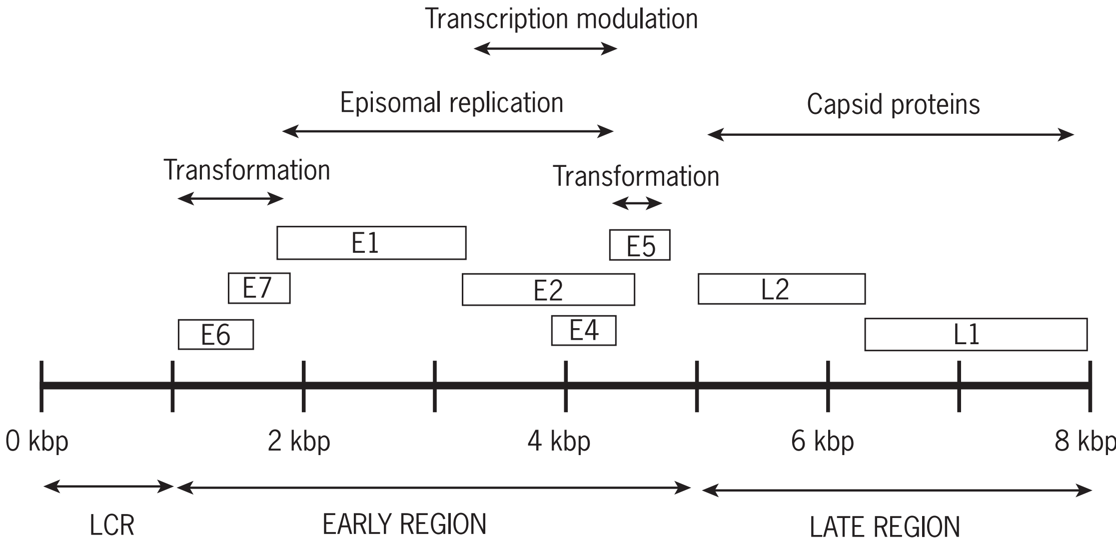


Figure 3 Organisation of a PV consensus genome (in kilobase pairs). The open boxes represent ORFs labelled E1–E6, or L1 and L2, according to their position in the 'Early' or 'Late' region of the genome. Locations of gene functions are listed above the genome. (Adapted from Pfister and Fuchs, 1994.)

the total protein mass) with the capsid proteins, but this does not occur in lesions caused by other PVs (Arrand, 1994).

The **E5** proteins are small (44–90 amino acids) and extremely hydrophobic polypeptides, which are present within intracellular membrane compartments, including the Golgi apparatus. In bovine (B)PVs, E5 is the major transforming protein and appears to stimulate mitogenesis by interfering with growth factor receptor signal transduction pathways (Stöppler *et al*., 1994).

The **E6** proteins, MW 16–18 kDa, have transforming and transcriptional transactivating activities and are localized in the nuclear matrix and in non-nuclear membrane fractions. E6 and E7 proteins are the major transforming proteins of HPVs. Both proteins appear to have arisen from duplication events involving a Cys-X-X-Cys motif. E6, similarly to adenovirus E1B and SV40 LT, associates with the tumour suppressor p53, resulting in its ubiquitin-dependent degradation. In HPVs, this degradation is mediated only by the E6 proteins of 'high-risk' but not 'low-risk' types, suggesting an important role for this process in the development of malignancy (Stöppler *et al*., 1994).

The **E7** proteins, MW 14–21 kDa, have transforming and transcriptional transactivating activities and are localized in the cytoplasm and the nucleolus. E7, similarly to the adenovirus E1A and SV40 LT, binds to the retinoblastoma tumour-suppressor gene product Rb-p105. This prevents Rb-p105 from interacting with the E2F transcription factor, thereby allowing initiation of the S-phase of the cell cycle. The E7 proteins of 'high-risk' HPVs are more effective in binding Rb-p105 than that of 'low-risk' types. E7 also has the ability to bind cyclins and cyclin-dependent kinases (cdks) and hence further disrupt the regulation of the cell cycle (Stöppler *et al*., 1994).

The **E8** ORF is only present in some PVs. In BPV4, the E8 protein is a small hydrophobic polypeptide localized in the cell membrane. E8 contributes to cell transformation by conferring anchorage-independent growth.

There are marked differences in the state and functional activity of PV genomes in benign tumours and in cancers of different species (Shah, 1990; Chow and Broker, 1994; Howley, 1996). In benign lesions the viral genome replicates as multicopy extrachromosomal plasmids. In carcinomas, however, the situation can be completely different. In cattle, the BPV type 4 genome is detected in alimentary tract papillomas but is apparently lost in carcinomas. In cottontail rabbit carcinomas the viral genome is present in the episomal form. In contrast, in some human genital tract cancers the viral genome, generally accompanied by deletions and mutations, is integrated into the cellular DNA. Integration with respect to the cellular DNA is not site-specific, but there is some specificity with respect to the site in the circular viral genome where the break for integration occurs. Viral genomes of HPV 16 and HPV 18 are found to be almost always interrupted in the E1–E2 region, producing a break that disrupts transcription of the E2 ORF, but not transcription of the E6, E7 and part of the E1 ORFs.

Transcription and Replication of the Viral Genome

The subdivision of the viral genome into 'early' and 'late' regions is based on the close association existing between viral replication and squamous epithelial differentiation. Upon entry into the basal stem cells of the epithelium, viral early gene expression is activated at a very low level,

leading to the temporary amplification and establishment of the DNA plasmids. Early gene expression in the basal cell layer also stimulates cell growth and, as the epithelial cells move upwards in tissues progressing through their differentiation programme, this pattern of expression is maintained. Finally, in the uppermost layers of the epithelium, viral replication and late gene expression are activated in the now fully differentiated epithelial cells, resulting in the production of infectious viral progeny (Chow and Broker, 1994; Howley, 1996).

Several factors contribute to the transcriptional complexity of PV genomes (Arrand, 1994; Pfister and Fuchs, 1994; Howley, 1996): first, the presence of multiple promoters (for example, the LCR of BPV-1 contains at least seven promoters); second, complex and multiple splice patterns which, associated with the activity of the different promoters and the use of different polyadenylation signals, give rise to an extensive variety of viral mRNAs, many of which are polycistronic; furthermore, some ORFs (e.g. E2 and E6) appear to be represented to different extents in different messages; finally, complex control by proteins and factors produced by the virus (E2, E6 and E7) or the host at both the intracellular (i.e. retinoic acid, NF-IL6, Oct-1, etc.) and extracellular (glucocorticoid hormones, TGF-β1 and -β2, EGF, etc.) levels occurs.

Detection of HPV Infection

Although detection and diagnosis of HPV-associated lesions can normally be achieved by colposcopy, histology and cytology, these methods are unable to identify specific HPV type(s) present in lesions. Serological responses against almost all HPV-derived antigens have been detected in infected individuals. However, the diagnostic utility of these serological responses is questionable because (1) they appear to be, for most antigens, non-type-specific, and (2) they persist for longer than the actual infection. Detection of HPV DNA is, therefore, the only reliable diagnostic tool available to establish current infection by specific HPV types. This approach involves the detection of the viral genome either directly (by Southern blot hybridization) or by polymerase chain reaction (PCR) protocols (Shah and Howley, 1996).

Pathogenesis of Infections

In contrast to some animal PV infections in which fibroblastic proliferation is prominent, the pathological effect of HPV infection is confined to the epithelium. All the layers of the normal epithelium are represented in the lesion, accompanied by certain characteristic histological features. The increased division rate of the basal cell layer leads to an irregularly thickened prickle cell layer, with abnormal mitoses also observed in all suprabasal layers. The granular layer contains foci of cells showing koilocytosis (cytoplasmic vacuolization) and nuclear changes (enlargement, hyperchromasia, degeneration and pyknosis). Koilocytosis is also a feature of the cornified layer of non-nucleated dead cells, which may also display hyperkeratosis (Shah, 1990).

HPV infection can be acquired in a variety of ways such as abrasions of the skin, sexual intercourse and passage through an infected birth canal, and results in a variety of clinical conditions (**Table 1**) (Shah and Howley, 1996). Most of these lesions have benign prognoses, but they may be associated with high levels of morbidity. For example, exophytic anogenital warts, one of the most common sexually transmitted diseases, usually cause itching, burning and pain, and have a significant negative effect on the psychosexual wellbeing of the individual. Other lesions, however, may undergo malignant transformation. In the case of flat anogenital warts, lesions in the uterine cervix may progress towards invasive carcinoma and 50% of the diagnosed population will ultimately die of the disease, accounting for about 15% of cancer-related deaths worldwide. The progression of benign papillomas to invasive cancers has certain characteristics which are shared in different species (Shah, 1990; Shah and Howley, 1996). First, only some of the virus types that infect a species have oncogenic potential. In epidermodysplasia verruciformis (EV), more than 20 different HPV types are recovered from the macular plaques characteristic of this disease. However, only two types, HPV5 and HPV8, predominate in the carcinomas that arise from these lesions. Similarly, over a dozen HPV types infect the human genital tract, but a majority of genital tract carcinomas are associated with only a few viral types, the so-called 'high-risk' types (predominantly HPV 16 and 18). Second, there is a long period between the initial infection and the development of invasive cancers. In humans this period may be between 5 and 40 years. Finally, cofactors are often involved in malignant progression. For example, carcinomas in EV patients arise preferentially in lesions that are exposed to sunlight.

Immunology of Infection

Viral infections are controlled by a combination of non-antigen-specific and antigen-specific immune responses. Most viruses induce these immune responses by causing lytic cell death which, in turn, causes inflammation and stimulates the production of cytokines. PV infection, in contrast, is non-lytic and, consequently little or no local inflammation is induced. This situation probably reflects the reduced ability of PVs to invoke effective immune responses that are capable of eliminating established lesions (Frazer, 1996). Nonetheless, there is evidence of involvement of the immune system in the control of PV infections.

Humoural (antibody) immune responses directed against almost all PV proteins have been detected in

infected individuals. However, only antibodies directed against L1 or L2 have been found to be neutralizing and to protect against infection. Antibodies against the E6 and E7 proteins of high risk HPVs, although not effective at the prophylactic level, are commonly associated with carcinoma of the uterine cervix. Antibodies against the E2 and E4 proteins have also been associated with carcinoma of the cervix, but not universally (Frazer, 1996; Shah and Howley, 1996).

The persistence of PV-induced lesions suggests that the development of an effective cellular immune response against PVs following infection is neither immediate nor universal. Nonetheless, several observations suggest that the host's cell-mediated immune response is responsible for limiting the growth and promoting the regression of PV-induced lesions. First, there is a high prevalence of PV-induced lesions and malignant tumours in hosts with suppressed cellular immunity. Second, the regression of anogenital and skin warts in humans is associated with a pronounced local infiltration of mononuclear cells (activated T lymphocytes, macrophages and, to lesser extents, NK cells and B lymphocytes) invading the epidermis and destroying the neoplastic tissue. On this basis, the cellular immune response in spontaneously regressing warts appears to be consistent with a delayed type hypersensitivity (DTH) reaction to foreign antigen. Nonetheless, the presence of specific cytotoxic CD8+ T lymphocytes (CTLs), which are normally involved in the resolution of viral infections, has been notoriously difficult to demonstrate in HPV infections. Indeed, only a handful of studies (i.e. Tarpey *et al*., 1994; Nakagawa *et al*., 1997) have shown E6- or E7- specific CTLs in humans. Finally, vaccination with PV antigens has been found to induce a specific T cell proliferative or CTL responses against L1, L2, E2, E6 and E7 in animals and humans (Frazer, 1996; Shah and Howley, 1996).

EPSTEIN–BARR VIRUS (EBV)

History, Definition and Classification

A relevant point to note about EBV (or HHV-4) is that it is almost ubiquitous in the adult human population, with the great majority of individuals carrying antibodies to the virus. Infection of B lymphocytes by EBV is mediated through interaction of a viral envelope glycoprotein with the receptor for the C3d complement component, CD21 (CR2), although, notably, EBV can infect epithelial cells that lack this receptor. The average age of seroconversion to this virus differs markedly in various parts of the world, usually being considerably later in the more socio-economic privileged parts of the world than in crowded, or poorer, populations. Viral antibody prevalence in terms of age is given in **Figure 4**. EBV has a particularly interesting international history. It was first observed by Epstein and colleagues in London in the 1960s on electron microscopic examination of a cell line (EB), established with extracts from an African tumour, called Burkitt's lymphoma (BL). From its physical appearance, the virus was defined as a member of the herpesvirus family. BL itself had been identified about 10 years earlier during travels by the Irish-born surgeon, Denis Burkitt in sub-Saharan Africa. Carvings from earlier periods showed this tumour of B lymphocytes to be a disease long prevalent in certain parts of Africa. Later, a genuinely serendipitous finding showed the virus to be the causative agent of infectious mononucleosis. This came about when a laboratory technician in Philadelphia, with no antibodies to EBV, developed mononucleosis and on subsequent testing was found to be EBV-antibody positive. A causal effect for EBV in mononucleosis was thus confirmed (reviewed by Griffin, 1998). An association between EBV and a tumour of epithelial cells, nasopharyngeal carcinoma (NPC), was discovered when sera from these patients were included in a general antibody screening programme in New York, also in the 1960s.

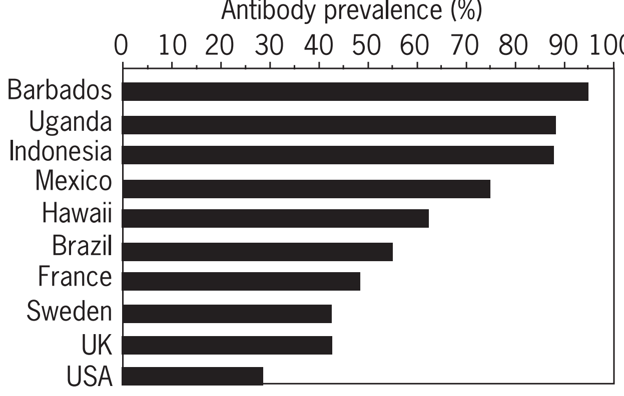


Figure 4 EBV antibody prevalence, age 4-6 years, in different parts of the world. (From IARC, 1997, p. 83.)

The natural reservoir of EBV, whether in the B lymphocyte or epithelial or other cell population, is still a matter of controversy. EBV is known to exist in circulating lymphocytes in the body (about 1 in $10^5 - 10^6$ in a normal individual). In culture, the virus is capable of extending the lifetime of B lymphocytes for an unlimited time period, a phenomenon called immortalization. EBV is sub-classified as γ-herpesvirus, having a restricted host range with its site of latency residing in lymphocytes, compared for example with α-herpesviruses, such as the simplex viruses, which have broad host ranges, and are latent in sensory ganglia. Herpesvirus classifications are given in **Table 2**. Although many other mammalian herpesviruses, notably those from Old World primates, belong to the γ-herpesvirus subfamily, only one other human herpesvirus identified to date, that of KSHV (HHV-8), belongs to this subclass. A full list of

Table 2 Biological characteristics of herpesvirus subfamilies

Characteristic	Alpha	Beta	Gamma
Genus	Simplexvirus Varicella-Zoster virus	Cytomegalovirus Muromegalovirus	Lymphocryptovirus Rhadinovirus
Host range	Broad	Restricted	Restricted
Prevalent genomic organisation*	D, E	Variable	B, C
Productive cycle	Short	Long	Long
Spread in culture	Efficient	Moderate	Poor
Site of latency	Sensory ganglia	Lymphoreticular tissues	Lymphocytes
Proliferation of latently infected cells	No	No	Yes

(From IARC, 1997, p. 36.)
*See **Figure 5**.

herpesviruses and their taxonomies is given elsewhere (IARC, 1997).

Unlike the Papillomaviruses, there is no well-defined classification of different strains of EBV. The viral genome is composed of double-stranded DNA, with sizes that range from about 175 to >200 kbp pairs in cells from different sources. In lytically infected (virus producer) cells, the viral DNA is linear and carries small repetitive sequences at each end. In latently infected (non-virus-producing) cells, it is circular, having undergone recombination via its repetitive terminal sequences. With regard to its structure, EBV differs from other human herpesviruses, as illustrated in **Figure 5**, and in organisation more resembles that of its host cell DNA than do the other viruses. Its size variation is not a property that has been used in classification since the viral genome is composed of unique sequences interspersed with repetitive elements, and size is largely dictated by copy numbers of the repeats. The largest of these, called IR1, or *Bam*HI W after the restriction enzyme that cleaves it intact from the DNA (**Figure 6**), is >3000 bp in size. An obvious classification sought, but not found, has been one that would allow for association of specific viral types with different EBV-associated malignancies, two of which are noted above. To date, pathology-specific strains of virus have not been identified. Rather, restricted viral gene expression, in part controlled by the host cell, may play a role in the genesis of a pathological lesion. Genetic polymorphisms, designated 1 and 2, have, however, been identified which differ in sequences of some viral nuclear antigen (EBNA) genes, and to some extent in their biological properties, and their global localization. Unlike the better-defined distinctions of herpes-simplex viruses 1 and 2, which are separately classified, the functional differences between EBV 1 and 2 are not sufficiently distinct to allow for unique classification. Indeed, the polymorphisms may merely reflect 'hotspots' for mutational recombination events in the genome. Such a hypothesis is not totally fanciful. At least one viral isolate, Jijoye (from a primary African BL), on propagation in culture, has given rise to a novel isolate, P3HR-1, with a deletion that maps within one of the key latent EBV genes (that for EBNA-2, see below) affected by the polymorphism in EBV 1 and 2.

Virion Structure

Whereas smaller DNA viruses, such as SV40, Polyoma, Papilloma and Adenoviruses, are nonenveloped, all the herpesviruses have an outer envelope and within this, a capsid that contains the viral DNA. By electron microscopy (EM), in composition and appearance EBV resembles a typical herpesvirus with a toroid-shaped protein core wrapped with genomic DNA, as shown in **Figure 7**. Its nucleocapsid is composed of 162 capsomeres and its outer envelope is made up of glycoprotein (gp) spikes, many of which are composed of a 220/350-kDa protein, the principal target of a virus-neutralizing antibody response. Size variation of this protein reflects the number of glycosylated amino acid residues it contains. To date, gp220/350 is still the prime candidate for producing an EBV vaccine that might prevent, or delay, infection *in vivo*. The high lipid content of the envelope results in relative instability of EB virions at room temperature, and their rapid inactivation by lipid solvents, such as ether and chloroform, or by detergents. This is another difference between the herpesviruses and the small DNA viruses, the latter being generally stable under these conditions. Between the nucleocapsid and the envelope is a region called the tegument, which is frequently distributed asymmetrically, and by EM shows no distinctive features. The composition of the tegument in EBV has been much less carefully studied than in some other herpesviruses, notably herpessimplex viruses.

Genomic Organisation and Key Viral Latent Functions

EBV was the first herpesvirus to have its complete DNA sequenced, as presented simplistically in **Figure 6**, determined (Baer *et al*., 1984). In its overall structure, with unique sequences interspersed with repetitive elements, the viral genome appears to be a mini-version of its human host, with one notable exception, that is, every repetitive region (IR1–IR4 and TR) includes ORFs, occurs within a gene and also encodes a protein. There is no 'junk' DNA, the role often assigned to repetitive sequences in cellular

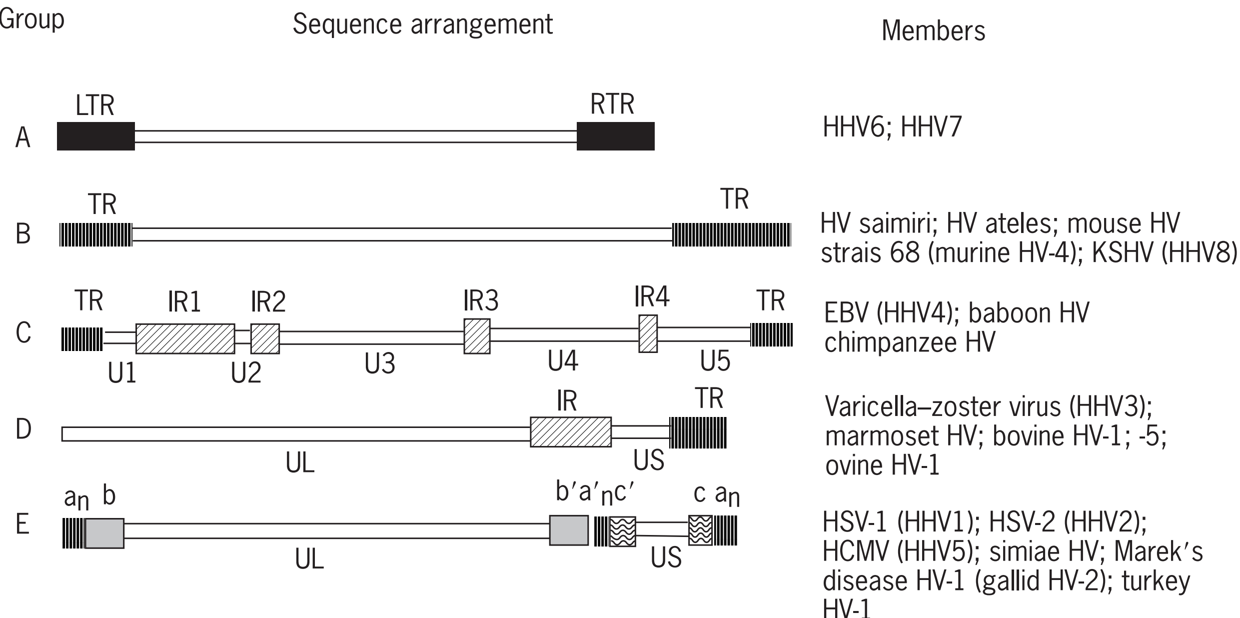


Figure 5 Architecture of herpesvirus genomes, showing unique (U) and repetitive (R) regions. Viruses have been grouped in categories (A–E) and their designations are given on the right. According to the nomenclature used by different groups, LTR = left terminal repeat; RTR = right terminal repeat; TR = terminal repeat; IR = internal repeat; UL = long terminal repeat; US = short teminal repeat. Symbols used in group E viruses represent sequence arrangements within repeats. (Adapted from IARC, 1997, p. 35.)

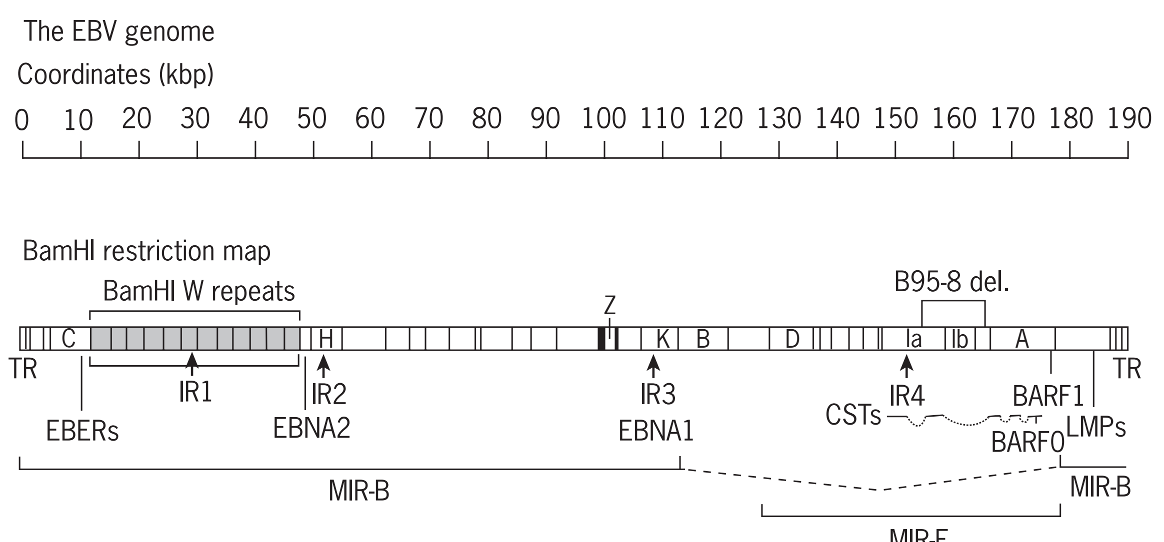


Figure 6 Physical map of the EBV genome, and location of some key gene and repetitive (IR) elements (see **Figure 5**). Coordinates for a 'typical' genome are given and genes allocated to the *Bam*HI restriction enzyme fragment in which they are located (see **Figure 10**). The CSTs encompass several restriction fragments. Its transcript is shown. MIR-B contains the minimum viral sequence required for immortalization of B lymphocytes and MIR-E the minimum for epithelial cells. (From Griffin and Xue, 1998, *Annals of Medicine*, **30**, 249–259.)

DNA, in EBV. Whereas genomes of the smaller tumour viruses, depending upon the stage in the cell cycle in which they are expressed, are classically divided into 'early' and 'late' ORFs, herpesviruses are divided into 'immediate early' (before DNA replication is initiated), 'early' and 'late' genes. An alternative classification divides their genes into 'latent' and 'lytic' functions. The latter, probably simplistic, classification is useful for discussion purposes

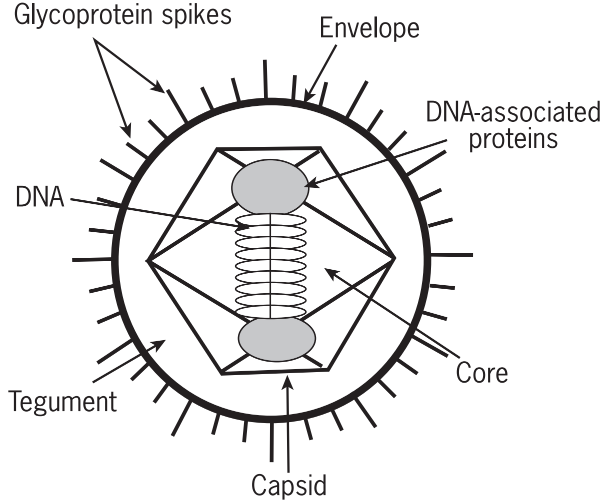


Figure 7 Schematic structure of a herpesvirus. (From IARC, 1997, p. 34.)

Table 3 Nomenclature of latent EBV gene products

Adopted terminology[a]	Alternative nomenclature[a]	
EBNA-1	EBNA-1	EBNA-1
EBNA-2	EBNA-2	EBNA-2
EBNA-3A	EBNA-3	EBNA-3A
EBNA-3B	EBNA-4	EBNA-3B
EBNA-3C	EBNA-6	EBNA-3C
EBNA-LP	EBNA-5	EBNA-4
LMP-1		
LMP-2A	TP-1	
LMP-2B	TP-2	
EBER-1		
EBER-2		

[a]EBNA, EBV nuclear antigen; LMP, latent membrane protein; EBER, EBV-encoded RNA; TP, terminal protein.
(From IARC, 1997, p. 53.)

when dealing with a large and complicated genome. Out of the 100 or so genes encoded by EBV, many of which are still poorly characterized, latent functions, as derived from data mainly drawn from EBV gene expression in B lymphocytes, consist of a small number of species only. These include six discrete EBV nuclear antigens (EBNAs), three discrete membrane antigens (LMPs) and two small RNAs (EBERs), of as yet unknown function. Interestingly, among other viruses studied in detail, only adenovirus encodes similar small RNAs (VA I and II) that structurally resemble EBERS, and although themselves not fully functionally characterized, are thought to modulate translation of viral proteins. For EBV, the EBERs are mainly localized in the nucleus and thus they may play alternative roles. Because of their very high levels of expression, EBERs have proved useful for detecting the presence of EBV in cells although, notably, they are apparently not expressed in all cells. For example, they are not found in a nonmalignant pathology associated with immunosuppression, oral hairy leucoplakia (OHL), where infected cells are frequently undergoing lytic replication. The nomenclatures used for the latent antigens in the EBV field are given in **Table 3**. The rest of the viral genes have been categorically designated as lytic, or lytically related. This distinct dichotomy into latent and lytic gene expression may be reassessed with time, since many EBV-associated tumours have recently been shown to express genes now designated as immediate early (or lytically related). Some of these, which may play initiating roles in the viral lytic cycle, may have other roles in tumours (discussed by Griffin and Xue, 1998). Alternatively, as proposed for KSHV (Ganem, 1998), a small amount of replication and thereby re-infection may be relevant to, and essential for, tumour growth.

These latent functions, on the assumption that most or all may play roles in the alteration of cell growth induced by EBV, at least for B lymphocytes in culture, are briefly defined as follows (IARC, 1997).

EBNA-1: a DNA-binding protein identified in all EBV-infected cells and responsible for EBV genome replication in latently infected cells. EBNA-1 is not recognized by the host cellular immune system, probably as a consequence of the glycine–alanine-rich repetitive (IR3) sequence within the protein. In transgenic mice, it is tumorigenic. This antigen and its pivotal function in EBV latency has recently been reviewed (Leight and Sugden, 2000)

EBNA-2: a transactivator both of other viral and cellular functions, and a key protein in B cell immortalization in culture. It is not generally expressed in EBV-associated tumours, although this does not rule out an early role in tumour induction. It is expressed in post-transplant lymphoproliferative disorders and in infectious mononucleosis.

EBNA-LP: appears to be important for the stimulation of B cell growth in culture and, like EBNA-2, be a contributing factor in post-transplant lymphoproliferative disorders and infectious mononucleosis. EBNA-2 and EBNA-LP are the first two proteins to be identified following cellular infection with the virus.

EBNA-3A, 3B and 3C: often considered together because they are derived from adjacent regions of the viral genome. EBNA-3A and 3C, but possibly not 3B, are involved in growth stimulation of B cells, but all three may have regulatory roles in the transcriptional control of other key viral functions. EBNA-3C has been compared in its properties to HPV E7 and adenovirus E1A proteins, both associated with cell growth alterations induced by their respective viruses.

LMP-1: often found expressed in EBV-associated tumours. In *in vitro* assays using heterologous promoters like SV40 LT, it is capable of inducing tumorigenic transformation of rodent fibroblasts in culture. It alters cytokeratin expression and inhibits cell differentiation.

This transmembrane antigen may recruit signalling antibodies and is absolutely required for both the initiation and maintenance of B cell growth in culture. In transgenic mice, LMP-1 produces a pathological response in keratinocytes, which has not been fully characterized.

LMP-2A and 2B: map across the terminal junctions of the viral DNA and therefore can only be expressed in latently infected cells, where the genome is circular. They do not appear to be directly involved in the *in vitro* growth stimulation of B cells, but may be important for the maintenance of latency. LMP-2A is a phosphoprotein, stably phosphorylated on tyrosine, and thus may have other unidentified functions.

Two other genes, *BARF1* and *BARF0*: more recently identified, both of which may play key roles, particularly in epithelial cell growth regulation. Their importance to B lymphocyte growth stimulation *in vivo* is less clear. The ***BARF1*** gene, like LMP-1, is fully competent for inducing tumorigenic cellular transformation of rodent cells, and even B lymphocytes in culture, when expressed under a strong, heterologous promoter. It has some homology with the human intercellular cell adhesion molecule 1 (ICAM-1). In limited studies carried out to date, BARF1 has been found expressed in most EBV-associated nasopharyngeal carcinomas (NPCs) examined. Its activities remain to be fully characterized. The second gene comes from ***Bam*HI I/A transcript**, also called complementary strand transcripts (**CSTs**), or ***BARF0*** gene. Primary CSTs extend over about 25 kbp of the viral genome (**Figure 6**) and spliced variants of it make up the major transcripts in NPCs. They were first identified in 1989 (Hitt *et al*., 1989) as a family of processed, multiply spliced polyadenylated RNAs and were subsequenty designated as 'complementary' in recognition of the fact that they were generated from the DNA strand with opposite polarity to that specifying numerous previously known viral genes. Each of the ORFs in the polycistronic CSTs, created by splicing events, overlap genes on the opposite strand, most of which are associated with lytic replication, which has led to the speculation that they may be involved in the maintenance of viral latency. CSTs are expressed also in BLs and other EBV-associated tumours, but at lower levels. They are often designated as latent functions as a consequence of their ubiquitous expression in tumours, but have also been found in lytically infected cells. A protein first described as a product of *BARF0*, the largest and terminal (3′ end of the gene, with its termination codon in the polyadenylation signal of the message) of the CST ORFs, was later identified in uninfected cells, casting doubt on its authenticity. BARF1 and CST expression and function(s) in EBV infected cells are key targets for future research.

The locations of some of the genes described on the physical map of EBV are given in **Figure 6**, and their designations and functions, where known, are summarized in **Table 4**. A unique working nomenclature has been established for EBV genes, where B stands for the *Bam*HI restriction DNA fragment containing a particular gene, a letter represents fragment size relative to the other *Bam*HI products (A being the largest and g the smallest in the sequenced B95-8 EBV genome; Baer *et al*., 1984), R (right) or L (left) denotes its direction (and polarity) on the conventional physical map of the genome, and a number denotes which reading frame is represented within a particular fragment. Thus, ***BARF1***, above, is the first rightwardly expressed ORF in the *Bam*HI A fragment. ***BARF0*** was not predicted by the DNA sequence, so it carries an aberrant designation. The differential expression of these genes in various EBV-associated tumours, or in lymphoblastoid cell lines (LCLs) generated by infecting B lymphocytes with the virus, have now led to subclassifications of viral latency, as simplistically illustrated for EBNAs and LMPs in **Figure 8**, and given in detail in **Table 5**.

The EBV genome also includes two other genes with interesting homologies to human genes: ***BCRF1*** and ***BHRF1,*** IL-10 and Bcl-2 homologs, respectively. Their roles in the virus have not been defined.

Cellular Immortalization *In Vitro*

Following a procedure first described for the small DNA tumour viruses, the minimal region of the EBV genome required for growth stimulation of cells in culture has been determined, using transfection protocols and fragments of the viral genome. The results of studies carried out on B lymphocytes (B) and epithelial (E) cells, representing the main tumour cell types associated with EBV tumours, are shown in **Figure 6**. Notably, there is no overlap between the minimal immortalizing regions (MIR-B and MIR-E) in these cells, supporting the argument that cell-type-specific functions may exist within the viral genome. For MIR-E, the data are consistent with findings on the transcription of EBV in NPCs, as determined by analysis of a comprehensive cDNA library made from the tumour (Hitt *et al*., 1989).

Pathogenesis

EBV is the causal agent for infectious mononucleosis, usually a self-limiting B cell proliferative disease, mainly a problem for economically privileged parts of the world where seroconversion and the development of antibodies to the virus occur late (**Figure 4**). With the hereditary immunodeficiency disorder, X-linked lymphoproliferative disease (XLP), or Duncan syndrome, fortunately rare, infection with EBV is usually fatal. Children that survive are at high risk of developing fatal lymphomas. With the so-called endemic form of Burkitt's lymphoma (BL), an acute problem for sub-Saharan Africa where it is the most prevalent cancer of children, there is a nearly 100% association with EBV. Again, in nasopharyngeal carcinoma (NPC), a head and neck tumour of poorly differentiated epithelial cells found with high frequency among the southern

Table 4 EBV gene products and proposed functions[a]

Open reading frame	Common name	Proposed function
Latent genes		
BKRF1	*EBNA-1*	Plasmid maintenance
BYRF1	*EBNA-2*	*trans*-Activation, transformation
BERF1	*EBNA-3A*	*trans*-Activation, transformation
BERF2	*EBNA-3B*	Unknown
BERF3/4	*EBNA-3C*	*trans*-Activation, transformation
BWRF1	*EBNA-LP*	*trans*-Activation, transformation
BNLF1	*LMP-1*	Transformation
BNRF1	*LMP-2A/2B*	Maintenance of latency
BARF0		*Unknown*
Immediate early genes		
BZLF1	*ZEBRA*	*trans*-Activation, initiation of lytic cycle
BRLF1		*trans*-Activation, initiation of lytic cycle
BI′LF4		*trans*-Activation, initiation of lytic cycle
Early genes		
BMRF1		*trans*-Activation
BARF1		Limited homology to *ICAM-1*
BALF2		*DNA binding*
BALF5		*DNA polymerase*
BORF2		Ribonucleotide reductase subunit
BaRF1		Ribonucleotide reductase subunit
BXLF1		Thymidine kinase
BGLF5		Alkaline exonuclease
BSLF1		Primase
BBLF4		Helicase
BKRF3		Uracil DNA glycosylase
Late genes		
BLLF1	*gp350*	Major envelope glycoprotein
BXLF2	*gp85 (gH)*	Virus–host envelope fusion
BKRF2	*gp25 (gL)*	Virus–host envelope fusion
BZLF2	*gp42*	Virus–host envelope fusion, binds MHC class II
BALF4	*gp110 (gB)*	Unknown
BDLF3	*gp100–150*	Unknown
BILF2	*gp55–80*	Unknown
BCRF1		Viral interleukin-10
BHRF1		Viral *bcl-2* analogue

[a]See ***Table 3***. ZEBRA, EBV replication activation; gp, glycoprotein; MHC, major histocompatibility complex; BARFO, major ORF in CSTs, function unknown. (Adapted from IARC, 1997, p. 50.)

Chinese and in some other parts of Asia, the viral association is 100%. These associations (see Introduction), largely based on clinical, epidemiological and serological approaches, have now been known for nearly half a century. What still is not known, however, is the precise contribution of EBV to these diseases, whether causal or merely contributory. If contributory only, in no case has the corresponding co-factor(s) been definitively identified, although there are candidates such as malaria for BL and smoked, salted fish consumption for NPC. What is firmly established, however, is the fact that the geographical, racial and age incidence of individuals that develop these EBV-related malignancies are totally distinct (**Figure 9**).

During the last 20 years, following the cloning and sequencing of the viral genome (Baer *et al*., 1984), which has allowed for the development of alternative, sensitive assays for identifying EBV and its gene projects, other tumours have been associated to varying degrees with the presence of this virus. These include a variety of tumours of different histopathological types, including subsets of lymphoepitheliomas, Hodgkin disease, stomach and breast cancers and T cell lymphomas. Notably, in none of the cases does the frequency of association approach that seen for endemic BL and NPC. However, the Working Group set up to explore the risk of EBV to humans, (IARC, 1997) concluded that there is sufficient evidence for the

carcinogenicity of this virus, in the causation of BL and other non-Hodgkin lymphomas, immunosuppression-related lymphomas, Hodgkin's disease (HD), sinonasal angiocentric T cell lymphoma and NPC, to allow them to conclude that EBV is a human carcinogen. Subsequent to this document, more information on the expression of EBV in breast cancer has been published, and the viral genome has also been identified on two occasions in carcinomas of the liver, previously a preserve of the hepatitis viruses. The future will undoubtedly bring more 'associations' for this ubiquitous virus, and hopefully, if suitable animal models are identified, notions about its actual role in disease. The sole argument that this virus alone could be sufficient for inducing malignancies under appropriate circumstance comes from the fact that many of the polyclonal lymphomas that develop as a consequence of immunosuppression (natural or induced) have a high frequency of association with EBV.

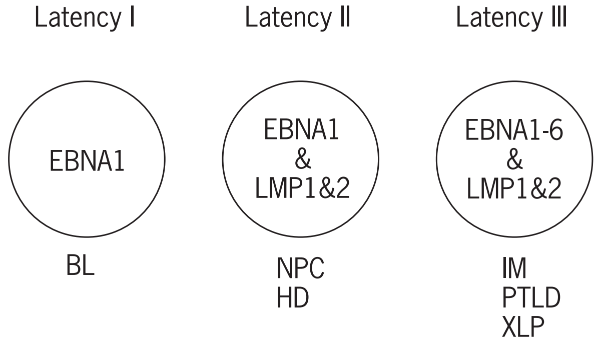


Figure 8 Patterns of latency gene expression in categories designated Latency I–III. BL = Burkitt's lymphoma; NPC = nasopharyngeal carcinoma; HD = Hodgkin's disease; IM = infectious mononucleosis; PTLD = post-transplant lymphoproliferative disease; XLP = X-linked lymphoproliferative disease (from Khanna *et al.*, 1999).

Immunological Considerations

One of the dominant characteristics about EBV is its adaptation to allow for persistence in its host(s), and gene expression, even in the presence of a functional immune system. EBV co-replicates with host DNA, and EBNA-1, required for latent replication, is tolerated, not eliminated, although there are epitopes for class I and class II HLAs in the viral antigen (Khanna *et al.*, 1999). The dominant feature in this protein that allows for its tolerance appears to be the repetitive (IR3) sequence that it harbours. In some cases of BL, where anti-EBNA-1 may be the sole antibody produced, this would allow for viral persistence. In situations where other antigens are expressed, for example in infectious mononucleosis or other EBV-associated malignancies, memory/activated T-cells appear to be important in limiting cell expansion and in targeting productively infected cells that express lytically related antigens. Immunological data suggest that vaccines designed to control primary EBV infection, a desirable objective in view of its carcinogenic role in humans, may profit by

Table 5 Patterns of latent EBV gene expression[a]

Type of latency	Gene product	Co-stimulatory molecules	Examples
IA	*EBERs*, *EBNAl*, *CSTs*		Burkitt's lymphoma
IB	*EBERs*, *EBNA1*, *CSTs* *LMP2A*		Gastric carcinoma
II	*EBERs*, *EBNA1*, *CSTs*, *LMP1*, *2A*, *2B*, *BARF1*	CD30 CD23 CD40 B7.1 LFA-1, -3 1CAM-1	Hodgkin disease Nasopharyngeal carcinoma T cell lymphoma
III	*EBERs*, *EBNA 1–4*, *6* *LP*, *LMP-1*, *2A*, *2B*	CD30 CD23 CD40 CD44 B7.1 LFA-1, -3 1CAM-1	Post-transplant lympho proliferative disorder Infectious mononucleosis
Other	*EBERs*, *EBNA1*, *2*		Smooth-muscle tumours

[a]See **Tables 3** and **4** and **Figure 8**.
(From Griffin, 2000, *Mutation Research*, **462**, 395–405.)

Figure 9 Comparative sites of highest frequencies of nasopharyngeal carcinoma (NPC) and Burkitt's lymphoma (BL), showing their disparate global locations. Black regions are sites of greatest frequencies and grey regions those of intermediate frequencies (for NPCs).

including dominant determinants of antigens associated with the viral life cycle (Khanna *et al*., 1999). This may be particularly relevant since animal models show that the development of neutralizing antibodies does not always correlate with protection from EBV infection. To this end, many of the dominant epitopes, including those found in latent and some lytically related proteins, have been mapped (**Figure 10**). Some of these might prove of value in the development of cytotoxic T lymphocyte (CTL) epitope-based vaccines, the aim of which would be to reduce morbidity and possibly clear infection. Since evidence suggests that many individuals having EBV-associated tumours retain detectable levels of EBV-specific T cells, needed for surveillance, this may be a reasonable approach. Even for BL, the tumour in which viral gene expression appears most tightly regulated, subpopulations of cells expressing lytically related antigens have been identified in some individuals (Labrecque *et al.*, 1999), making them also candidates for immunotherapeutic approaches. The recognition of the important contribution of EBV to diseases of humans has greatly stimulated efforts over the past few years to control this virus.

KAPOSI SARCOMA-ASSOCIATED HERPESVIRUS (KSHV)

History

In 1872, M. Kaposi, a Hungarian dermatologist, described a pigmented angiosarcoma, now called 'classic' or sporadic Kaposi sarcoma (KS), that mainly affected skin on the lower limbs, and was most prominent in elderly men of Mediterranean and eastern European origin. KS was also an African problem. In the 1960s and 1970s, the frequency and distribution of KS altered, and in many cases could be related to transplant therapies in other parts of the world. Whereas modest increases in KS were being reported in various countries prior to the onset of the syndromes now covered under the generic name AIDS, its frequency and epidemiology were drastically influenced by the spread of this virus. Over the past decade or so, although the histopathological presentations of all types of KS are identical, this malignancy has been generally subclassified into classic (sporadic), endemic (African), epidemic (AIDS related) and immunosuppression-associated (transplant) types, to reflect its origin. From being a comparatively rare form of cancer, KS is now fairly common in certain parts of the world. Exactly how common, however, is a controversial topic. The epidemiology of this cancer, and particularly the fact that in the early days it was the most common tumour in AIDS patients, with 15–20% of them developing KS, suggested that this disease might have an infectious aetiology (IARC, 1997). Thus, an active search to find such an agent was initiated.

The history of the discovery of KSHV is different from that of EBV, the human virus it most resembles, and owes much to the development of molecular biological methodologies. One of these in particular, called representational difference analysis (RDA), was used by a group in the USA, working with the husband and wife team Moore and Chang (Chang *et al*., 1994), in their search for a KS infectious agent. RDA consists of generating genomic representative entities from diseased and normal tissues, preferably from the same individual, using PCR amplification. These are stably associated with priming PCR sequences and hybridized to an excess of representative, nonligated amplified sequence, with no attached primers, from normal tissue. Following this procedure, only unique sequences found in the diseased tissues will contain priming sequences on both strands, which allows them to be substrates for subsequent PCR reactions. Repeating such a process enriches the sample for unique sequences.

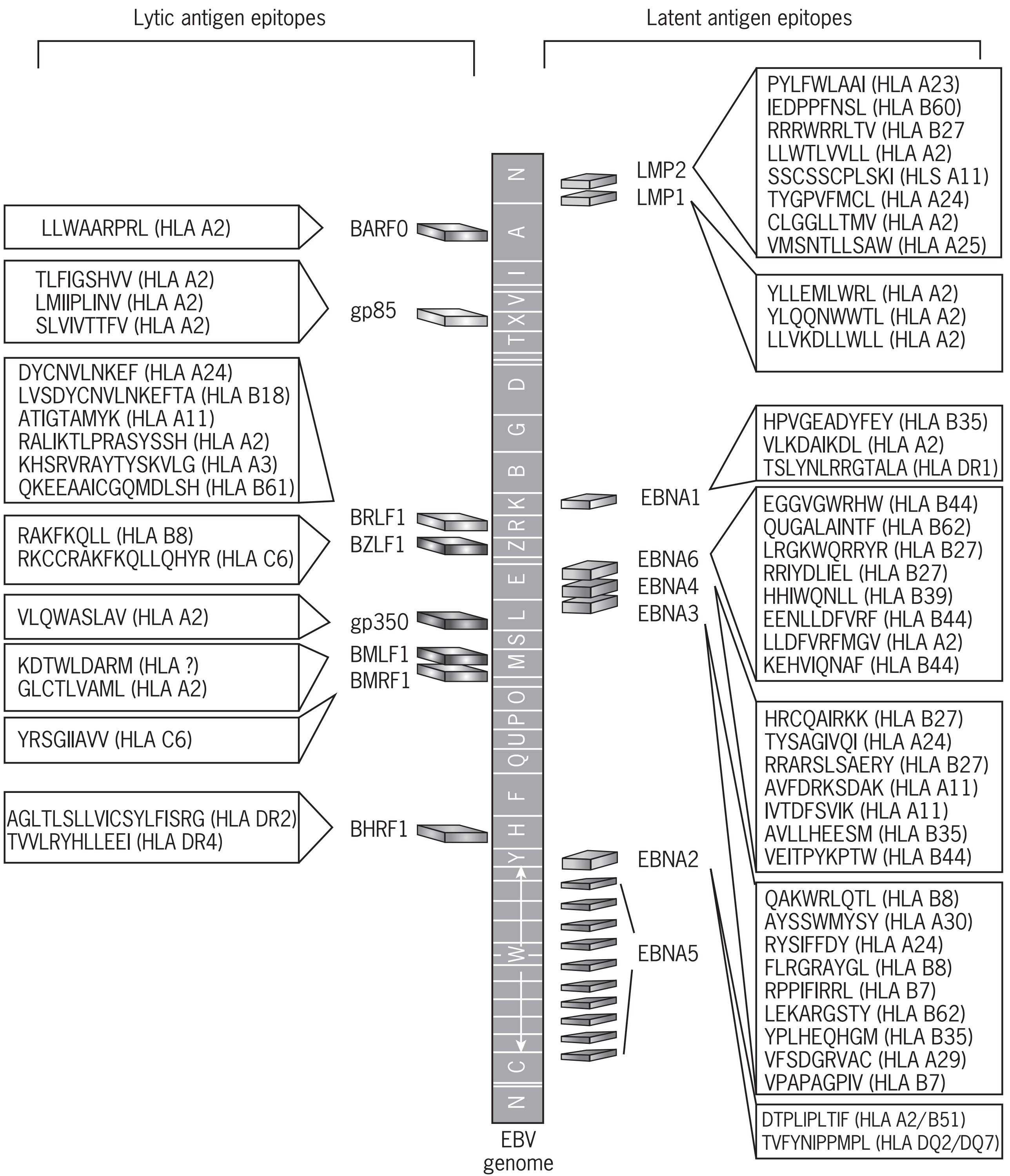


Figure 10 Schematic distribution of HLA class I and class II restricted cytotoxic T lymphocyte epitopes with EBV latent and lytic antigens, at the peptide level. BARF0, the largest ORF in the CST transcripts is given here as a lytic function, although it is also expressed during the latent cycle. (Adapted from Khanna *et al.*, 1999.)

These can then be purified and their sequences determined. By RDA, using tissues from AIDS-associated KS, Chang *et al*. (1994) identified sequences that were homologous with, but distinct from, other members of the γ-herpesvirus family, most notably EBV and the oncogenic primate virus, herpesvirus saimiri. They correctly concluded, as was subsequently shown, that this work was consistent with the presence of a new human

herpesvirus in KS lesions. Interestingly, the homologies they identified were with EBV late viral genes (in *BDLF1* and *BcLF1*, see EBV section).

Virion and Genome Structure

KSHV, or HHV8, has morphological features typical of herpesviruses (**Figure 7**), consisting of 100–150 nm particles surrounded by a lipid envelope, with an internal electron-rich central core. Its DNA was assessed by pulsed-field electrophoresis as 160–170 kb, consistent with that of other γ-herpesviruses, but more uniform than that observed with EBV. Both circular and linear forms of KSHV have been identified. Larger genomes reported to exist in some tumour-derived cell lines have been shown to result from DNA duplications, possibly associated with propagation in culture. In the same year, 1996, as the physical studies were reported, the complete sequence of the viral genome was published and an open reading frame map generated. This aptly illustrates the rapidity with which this field was being and has progressed. The sequence study (Russo, 1996) showed that the structure of KSHV was essentially similar to that of herpesvirus saimiri (HVS) (**Figure 5**). For KSHV, the genome has a single 140.5-kb long unique region, containing about 80 ORFs, flanked on either side by variable-length terminal repeats, about 800 bp in size. Within the genome, there were small repeat regions, some but not all of which appeared to be within ORFs, but overall there was little organisational similarity to EBV. In addition to numerous homologies with HVS, the sequence of KSHV also showed homologies with EBV genes, mainly those coding for late viral antigens where similarities that ranged from 44% to greater than 70% were observed. With EBV late genes, *BDLF1* and *BcLF1*, identified in the initial studies of Chang *et al.* (1994), the degree of homology at the DNA level was about 75%. Interestingly, the important viral DNA polymerases of these two viruses have 72% homology, although whether the enzymes themselves can be functionally interchanged is not known.

Although homologies between KSHV and EBV immediate early functions were observed, there were no homologues to EBV latent genes found in the ORFs of KSHV. Different isolates of KSHV appear to have highly conserved genomes. The phylogenetic tree of KSHV, based on aligned amino acid sequences as they relate to other herpesviruses, is shown in **Figure 11**. A close relationship with HVS, from squirrel monkeys (not apparently oncogenic in its natural host but tumorigenic to other nonhuman primates), and with equine herpesvirus 2 (with a more distant relationship with EBV) is seen (IARC, 1997).

Putative Key Tumour Genes

Studies on KSHV genes, to designate them as latent, immediate early, early and late genes, and identify those

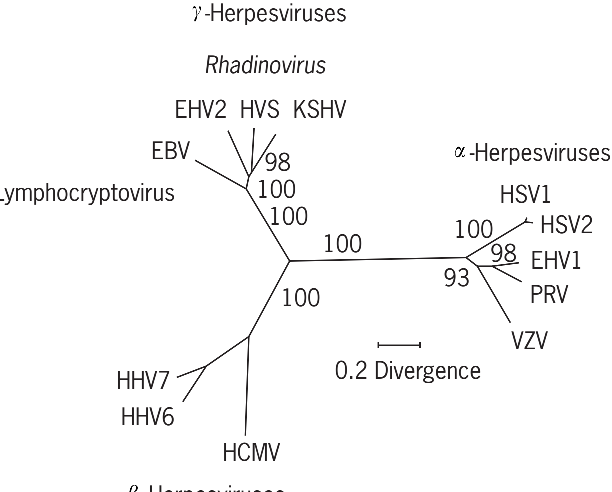


Figure 11 Phylogenetic tree of KSHV (HHV8) in relation to other herpesviruses. The comparison shows KSHV to be most closely related to the γ-herpesviruses, EBV, equine herpesvirus 2 (EHV2) and herpesvirus saimiri (HVS), its nearest relative (see **Figure 5**). (From IARC, 1997, p. 385.)

that may play key roles in the oncogenic activity of this virus, have been initiated. Here, as with other herpesviruses, latent transcripts are defined as constitutively expressed mRNAs which are not susceptible to chemical induction (e.g. with agents such as the promoter-stimulating phorbol ester (TPA) or n-butyrate, which affects chromatin structure) but are susceptible to inhibition by cycloheximide, an inhibitor of protein synthesis. Immediate early genes, on the other hand, are those whose transcripts are inducible, but resistant to the action of cycloheximide. Early gene expression is blocked by cycloheximide, but not by phosphonoacetic acid (PAA), an inhibitor of the virus-encoded DNA polymerase, whereas late lytic cycle gene expression is not blocked by the latter (Sun *et al.*, 1999). The same definitions are used when considering EBV genes. In the case of KSHV, many apparently nonlatent genes have been found expressed in the virally associated tumours (see below), and the same is becoming apparent for EBV. Their roles in these settings are still undefined. Lines derived from primary effusion lymphoma (PEL) cells infected with KSHV have proved useful in identifying some of the genes in KSHV that act as possible tumour-inducing agents. Several of these (ORF 71–73, see below) classified as latent genes since their transcription products are constitutively expressed, are clustered in the viral genome. Notably, a completely different gene designation system from that for EBV has been adopted for KSHV based on gene numbers from the sequence.

LANA (latency associated antigen (ORF 73): a large (226/234-kDa) protein that reacts with sera from AIDS patients, characterized by a typical speckled nuclear pattern. Antibodies to LANA have been postulated to have

prognostic value for the likelihood of an individual developing KS. They do not cross-react with EBV-specific antigens.

v-FLIP (ORF 71): so named because of its homology with a cellular anti-apoptotic factor, c-FLIP, which regulates the apoptosis triggered by Fas and other members of this tumour necrosis factor receptor family. In KSHV, these genes are overlapping and probes for v-FLIP also recognize LANA, but not vice versa (Sun *et al*., 1999). Notably, HVS also expresses a v-FLIP, which appears to be a late function. Interestingly, probes for v-FLIP also recognize a lytic cycle transcript.

ORF K72: tentatively identified as a latent function, with about 30% amino acid homology to the human cellular cyclin, D2, and expressed in persistently infected cells alongside LANA and v-FLIP. In culture, K72 (alternatively, KSHV v-cyclin), phosphorylates the retinoblastoma protein, Rb. K72 is a small ORF (60 amino acids), and may also be transcribed during the lytic cycle thus, like v-FLIP, possibly playing more than one role related to its associated malignancies, in the virus.

In addition to these three apparently latent genes, numerous viral genes associated with later stages of the viral life cycle have also been identified in the sarcomas. One of the immediate early KHVS genes is structurally and functionally related to an EBV-transactivator gene. In functional assays in culture, this viral gene has been found competent to initiate reactivation of a cascade of genes associated with the virus lytic cycle. KSHV also encodes a number of homologues of proinflammatory cytokines, such as IL-6 and macrophage inflammatory protein (MIP), as well as *bcl-2*, a homologue of another anti-apoptotic gene (also found in EBV, in *BHRF1*) and v-GCR, a G protein-coupled receptor. By definition, several of these appear to act as late genes, but to be expressed in tumours.

At this stage in the understanding of KSHV and its role in malignancy, temporal expression of tumour-associated genes thus seems to differ in large part from that of EBV, where the dominant components are latent genes. Whether latent gene expression in tumours may also dominate in the case with KSHV remains to be seen. Data suggesting that KSHV may express proteins related to the membrane-associated oncogenes of EBV, LMP1 and 2A have been obtained (Glenn *et al*., 1999). For KSHV, on the other hand, it has been proposed that tumour growth may be enhanced by viral chemokines or cytokines expressed by adjacent infected cells that have undergone a switch from latency to a lytic cycle type, giving to such genes an important role in malignancy (Ganem, 1998). Notably, a similar situation exists for some EBV-associated BLs (see above), and may account for the proliferation of this tumour with its remarkable doubling times of 28–60 h.

Many of the studies on gene expression in KSHV are currently being made on PEL cell lines, since growth of cells in culture from KS often results in apparent loss of the viral genome. Whether PEL lines are suitable models for KSHV expression in tumour settings remains to be seen. Like lymphoblastoid cell lines as a model for the role of EBV in BL, they may be imperfect, but nonetheless of value for studying the function of the viral genes.

Pathogenesis

Attempts to detect KSHV in peripheral blood mononuclear cells (PBMCs) from healthy individuals in countries with a low prevalence of KS, even by very sensitive PCR approaches, has not been generally successful. In KS-risk countries, variable associations of the virus with PBMCs have been reported. On the assumption that KSHV may be sexually transmitted, studies on semen specimens have also been carried out, with controversial results. However, sexual behaviour does seem to be a risk factor in transmission of this virus and in the development of KS, the risk running parallel to that of HIV infection. The notion of a KS-associated infectious agent, independent of HIV infection, appears to be real, in that the virus has been identified in all four epidemiological forms of KS, with no significant differences in detection rates (IARC, 1997). In addition to an association with KS, another neoplastic condition, primary effusion lymphoma, a rare, distinct type of non-Hodgkin lymphoma, has also been associated with KSHV infection. The cells in this malignancy are usually large and irregularly shaped, with abundant cytoplasm, and prominent nucleoli and mitotic features, the latter properties also found in BL. Notably, both KSHV and EBV can often be identified in these tumours. In AIDS patients, this is a fulminant lymphoproliferation and the median survival time of the individual is short. Other B and T cell lymphomas, explored for its presence, have not revealed any KSHV. On the other hand, there is evidence for a role for this virus in Castelman disease, at least in AIDS patients. This is a rare, usually polyclonal, non-neoplastic disorder of unknown aetiology, first reported in 1956.

Several scenarios could account for the association of KSHV with these tumours, particularly with KS. First, the virus may be the aetiological agent, and one or more of the functions noted above, or other as yet undiscovered viral gene products, may play a critical role in disease. On the other hand, as often suggested for EBV, the virus may be a contributory factor to the malignancy, e.g. by stimulating cytokines which enhance cell growth. Alternatively, the virus may be a mere passenger with the capacity to infect the cell types now associated with tumours that harbour it. Again, as with EBV, it is difficult to distinguish among these possible scenarios. With both viruses, the fact that viral homologues exist in other primates, which in model studies are shown to produce tumours, can be viewed as supportive evidence for these herpesviruses as acting both risk factors and tumour-inducing agents, under the appropriate conditions.

HEPATITIS B VIRUS (HBV)

History

Jaundice (from *jaune*) is a disease of the liver that has been known for centuries. Its most notable characteristic is an orange–yellow discoloration of the skin and conjunctivae, caused by deposition of elevated levels of bilirubin produced from damaged hepatocytes. Viral hepatitis, a general term for infections of the liver, can be caused by a number of hepatitis viruses, only one of which, HBV, is classified as a DNA virus. HBV is an unusual DNA virus in that it has, as a component of its life cycle, an RNA reverse transcriptase activity, a function normally associated with RNA retroviruses. The mechanism of HBV replication is unique for a DNA virus, in that it involves an RNA intermediate. HBV has a striking tropism for hepatocytes, but also can be detected in PBMCs.

Although viral hepatitis is a major public health problem, the identification of its viral association(s) and of HBV as one of the infectious agents of the disease was long in coming. Prior to its ultimate discovery, epidemiological differences among the diseases had suggested the possible existence of more than one infectious agent. In the early 1960s, Blumberg, looking for inherited polymorphic traits in blood from different parts of the world, identified an antigen, subsequently designated as 'Australia antigen', in sera from an Australian aborigine which reacted specifically with an antibody found in serum from an American haemophilia patient. This antigen proved to be geographically restricted, being relatively rare in American and Western European individuals, but more common in African or Asian populations, and in patients with certain distinct pathologies, including leukaemia, leprosy and Down syndrome. The association of Australia antigen, now known as hepatitis B surface antigen (HBsAg), with viral hepatitis was made several years later. This seminal finding led others to undertake studies aimed at identifying the infectious agent. In 1970, Dane and colleagues first identified, by EM, virus-like particles in the sera of patients with Australia antigen-associated hepatitis (Dane *et al*., 1970). At this time, it was estimated that of the 3.5 billion people in the world, as many as 175 million might be carriers of HBV. In spite of the millions of viral carriers, the cancer now associated with HBV, hepatocellular carcinoma (HCC), is relatively uncommon, but nonetheless correlates in frequency with those parts of the world with the highest percentages of carriers of the virus, and possibly reflects a progressive disease, initiated by hepatocyte infection and proceeding through the development of chronic hepatitis, to tumour formation.

HBV is the prototype for a family of viruses, now called the Hepadnaviridae, found in woodchucks, ground and tree squirrels and Peking ducks, all sharing distinctive morphologies and genomic characteristics. Studies on these viruses lend support for a causal relationship between HBV and some liver cancers. In particular, the duck hepatitis virus has been associated with the development of hepatoma in its host. The Working Party dealing with the association between HBV and cancer (IARC, 1994) concluded that chronic infection with hepatitis B is carcinogenic to humans. It reached the same conclusion with regard to the RNA virus, hepatitis C (HCV). In making this judgment, several criteria for causality were used, one being that a strong association, as found with the hepatitis viruses, is a better indicator of causality than a weak association.

Virion Structure and the Virus Life Cycle

The HB infectious virion is a 42-nm double-shelled spherical particle (originally called the Dane particle) that consists of an outer envelope composed of HBsAg and an inner core, or nucleocapsid, with its own antigens, hepatitis B core (HBcAg) and e (HBeAg), antigens, together with HBsAg, acting as markers for the presence of intact virions and infectivity. Infectious virions also contain a small (3.2-kb) circular, partially single-stranded DNA, and an endogenous DNA polymerase that can produce a fully double-stranded genome. Electron microsopic (EM) analyses show that in patient's sera, however, the concentrations of incomplete HBsAg structures may exceed those of complete virions. One of these forms is a small (20–25 nm) spherical particle and the other(s) is a tubular or filamentous form varying in size from 20×20–200 nm. Structures of these particles are schematically illustrated in **Figure 12**. Interestingly, similar shaped tubular (or filamentous particles) are also seen in early EM pictures of the small DNA papovaviruses, where they have been postulated to represent precursor forms of the mature, virion spherical particles, although this has not been proved. Interestingly, the major capsid protein, VP1, of papova viruses, expressed on its own *in vitro*, spontaneously reassembles to form viral-like capsid particles, composed of icosohedral (major product) and tubular (minor product) structures.

The site of primary replication of HBV is the liver, although it has been postulated, based on abnormalities observed in patients with acute hepatitis, that haematopoietic stem cells many also support viral replication. HBV infection probably involves viral attachment to specific receptors, although these have not been identified, nor have mechanisms for attachment and penetration been elucidated. Recent data suggest that HBV nucleocapsids do not enter the nucleus. Rather, they are arrested at the membrane and release the partially single-stranded genomic DNA into the nucleus where it is converted to covalently closed-circular (CCC) double-stranded DNA, which in turn serves as the viral transcription template.

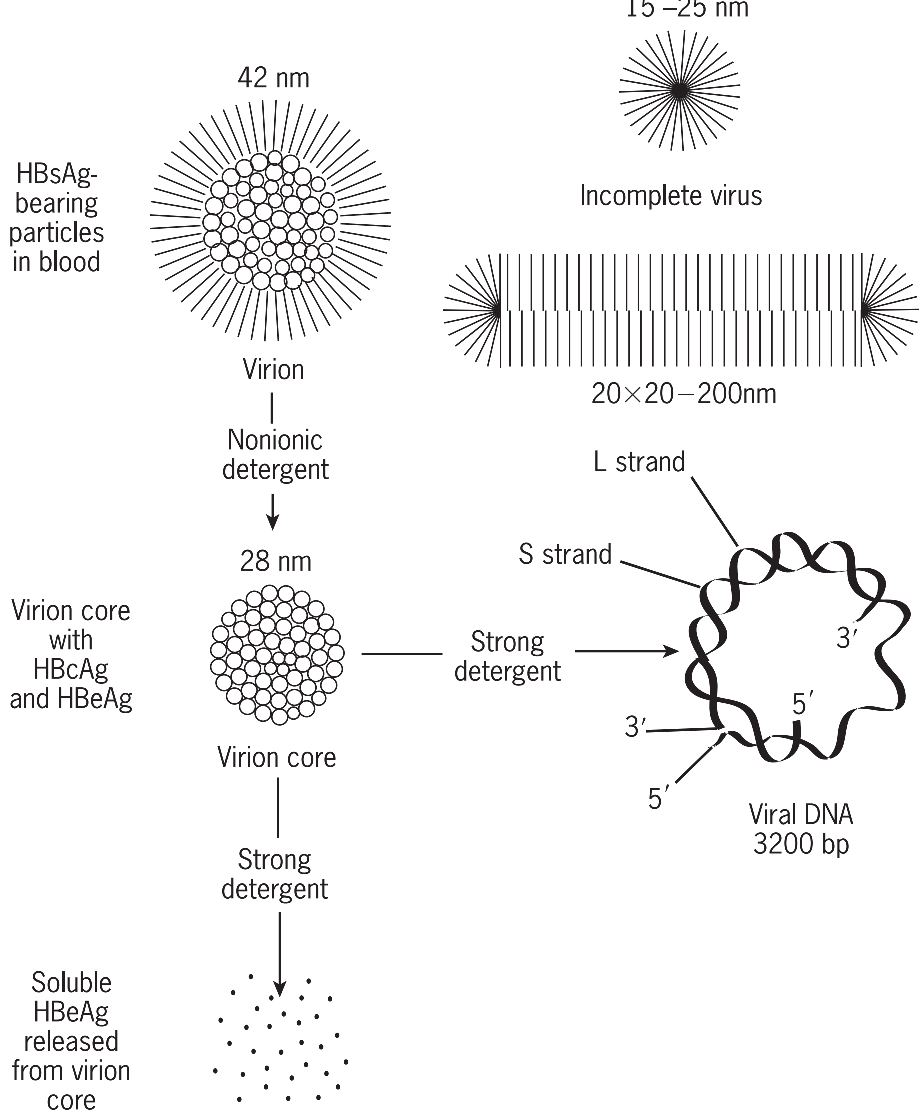


Figure 12 Diagram of the different forms adopted by hepatitis B virus (HBV).

Genome Structure, Replication and Antigens

The viral genome is organized into four transciption units, each with its own independent promoter. They share, however, a common polyadenylation site. Transcription yields four extensively overlapping viral RNAs, with sizes of 3.5, 2.4, 2.1 and 0.7 kb. These are exported into the cytoplasm where viral proteins are translated, and viral particle assembly and genome replication occurs (Chisari, 2000). HBV replication involves reverse transcription of the RNA pregenome (pRNA) to produce minus-strand DNA, the template for plus-strand DNA. Replication results in an encapsidated double-stranded open circular genome that is either recycled back to the the nucleus (to amplify the pool of CCC-DNA) or becomes enveloped by the viral protein and proceeds to complete the life cycle, as illustrated in **Figure 13**. The transcriptional products themselves are complex.

The 3.5-kb transcript: specifies viral genes reversibly transcribed as a first step in genome replication. The POL protein has numerous activities, acting as reverse transcriptase, DNA polymerase and RNase H, all essential functions for viral replication. The core protein (HBcAg) can form homodimers that self-assemble into capsid particles. In the cytoplasm, these also contain pregenomic viral RNA and POL, and the whole particle acts as the site for viral replication. The precore protein has a sequence that directs it to the endoplasmic reticulum (ER), where it undergoes limited proteolysis to produce the e antigen (HBeAg) which is secreted into the plasma membrane. Its role in the viral life cycle, in spite of its obvious importance, is still poorly understood.

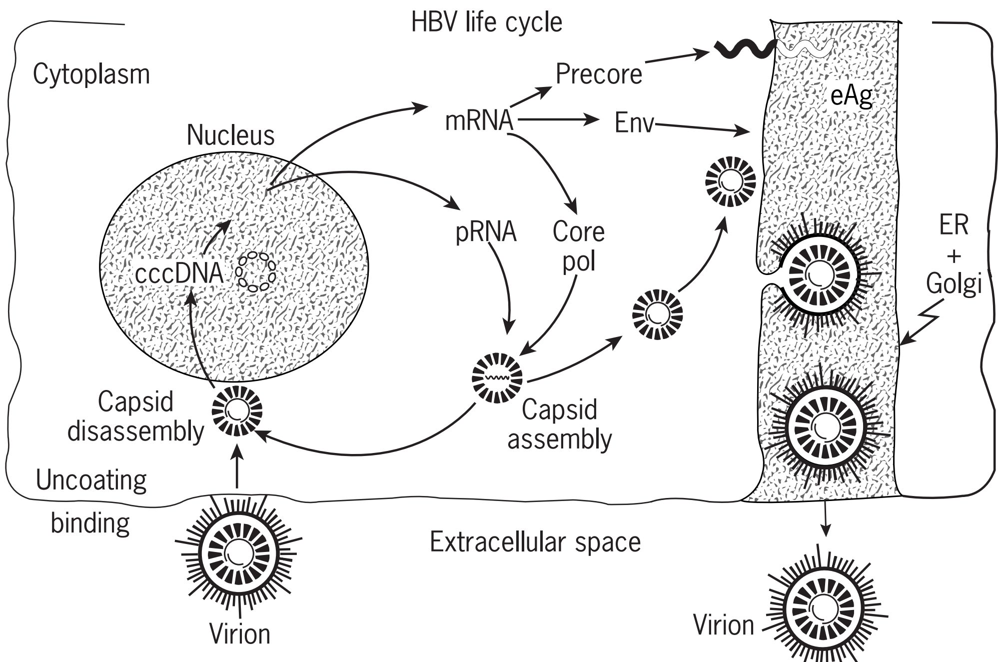


Figure 13 The HBV life cycle. Entry of the virus is still poorly defined. While the RNA containing capsid is maturing into a DNA-containing capsid, it migrates bidirectionally within the cytoplasm. One pathway terminates at the endopasmic reticulum (ER), where it interacts with envelope proteins which trigger an internal budding reaction, resulting in the formation of virions that are transported out of the cell by the default secretory pathway. The second pathway transports the maturing capsid to the nucleus to amplify the pool of covalently closed circular (CCC) DNA. (Adapted from Chisari, 2000.)

The 2.4- and 2.1-kb transcripts: produce the large, middle and small envelope proteins, which share common carboxy termini. The small, but major, transcript encodes the hepatitis B surface antigen (HBsAg). The middle and large transcripts which encode preS-2 and preS-1 antigens, respectively have amino acid extensions. The envelope proteins are cotranslationally inserted into ER membranes, where they aggregate, bud into the lumen and then are secreted by the cell as 22 nm subviral particles (**Figures 12** and **13**) or, if they have enveloped nucleocapsids before budding, become 42 nm infectious virions. The filamentous particles are generated when the large envelope protein is overexpressed. These are not usually secreted, but rather give an histologically distinct appearance (like 'ground glass') to cells, and hypersensitize them to the cytopathic effects of interferon-γ.

The 0.7-kb transcript: encodes the transactivator X protein. In the woodchuck model, the X protein is required to initiate infection. By virtue of its ability to transactivate expression of other genes, X is generally considered to be an important contributor to the pathologies induced by HBV. It was earlier thought to have properties associated with cell transformation and, on overexpression in transgenic mice, it can induce a high baseline incidence of HCC. However, in spite of considerable interest in X, its precise role *in vivo* still seems to be far from thoroughly characterized and its function as an oncogene has been called into question.

A transcriptional map of HBV, showing antigen locations, is given in **Figure 14**.

Pathogenesis

The natural history, clinical manifestations and geographical variation (as illustrated in **Figure 15**) of HBV infection, are highly variable. Chronic infection with HBV as related to endemicity, geography and mode and time of infection is given in **Table 6** and a brief resumé of the pathological consequences of infection in **Figure 16**. Serological patterns that accompany acute and chronic infections are given in **Figure 17**.

In spite of the wealth of knowledge that has been generated on this virus, the mechanism(s) by which HBV induces cellular transformation remains largely obscure. No viral oncogenes have been identified, and most or all of the viral antigens appear to be primarily involved in one or other aspect of the virus life cycle. In the woodchuck model, the X protein (see above) has been associated with infection. One promising lead on pathogenesis lies in the fact that HBV has a high mutation rate, which, although 100–1000 times lower than that observed with RNA

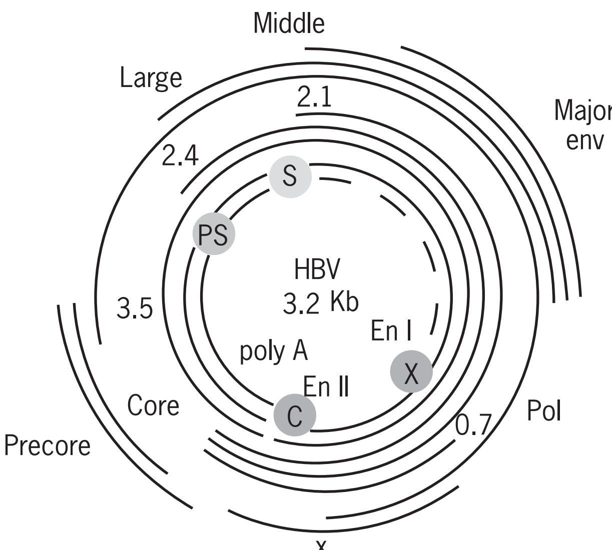


Figure 14 Transcriptional map of HBV, with the partially double-stranded 3.2-kb open circular genome present in virions shown in the centre. The core (C) pre-S (PS), HBs (S) and HBx (X) promoters are given inside round icons. (Adapted from Chisari, 2000.)

viruses, is nonetheless many times greater than normally seen with DNA viruses (IARC, 1994). However, although many HBV gene mutants have been identified and investigated, none to date has been proved beyond reasonable doubt to confer oncogenicity on the gene in question. Alternatively, in hepatocytes, HBV integrates into the host chromosome, giving rise to the possibility of insertional mutagenesis. Viral genome insertions in HCC have not, however, proved specific, nor have they been instrumental in pointing to insertional mutagenesis as explaining a viral role in HCC. The ability of X protein to affect expression of other genes allows for the possibility of either enhancement of expression of cellular genes associated with transformation, or down-regulation of tumour-suppressor genes, but has led to no definitive answers. Although approaches which consist of both direct and indirect influences on aberrant cell growth have been explored, and have generated interesting but frequently conflicting, data, the molecular route(s) by which hepadnaviruses predispose their host to malignancy remains an open question. There will probably be no simple answers, and one or more of the pathways that have been explored to date may play a role in diseases associated with this virus.

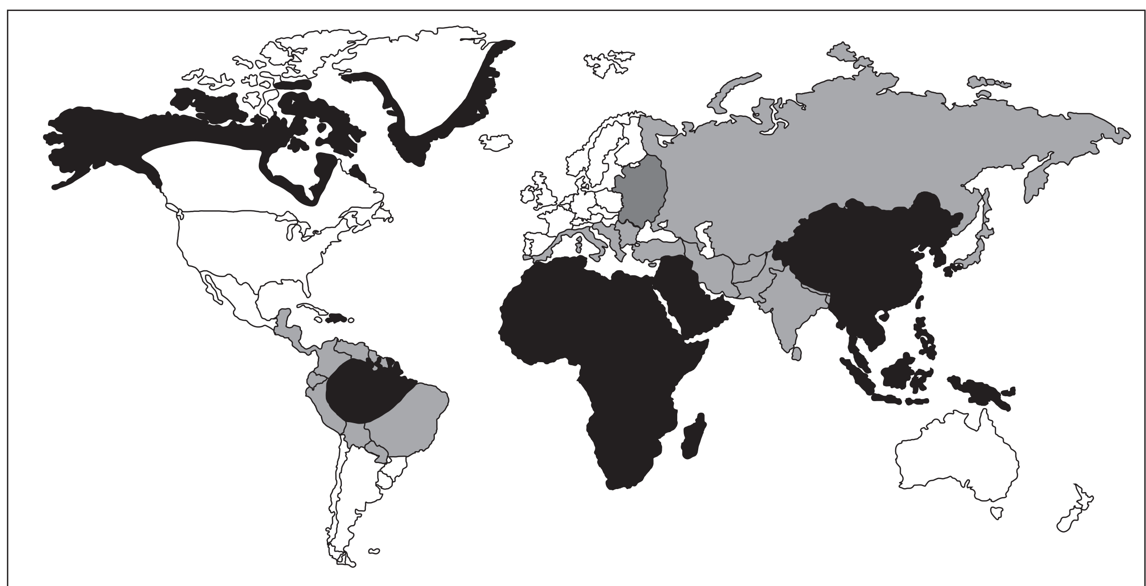


Figure 15 Geographical distribution of HBV. (From IARC, 1994, p. 56.) Black areas: high; grey areas: intermediate; white areas: low. (See **Table 6**.)

Table 6 Chronic infection with HBV : geography, mode and time

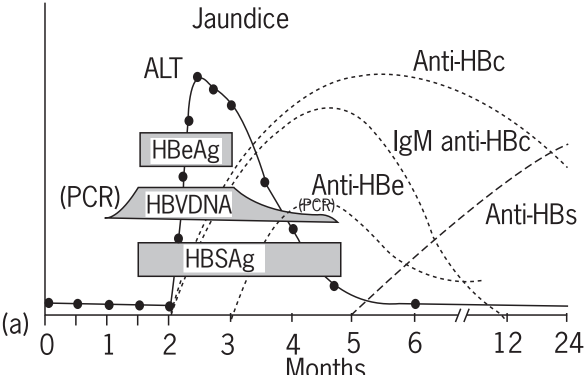
Endemicity	Geographical area	Predominant time of infection
High, ≥8%	China, Southeast Asia, Pacific Basin, sub-Saharan Africa, Amazon Basin	Perinatal, childhood
Intermediate, 2–7%	East, central and southern Europe, Middle East, South Asia, Japan	Perinatal, childhood, adulthood
Low, <2%	North America, western Europe, Australia, southern Latin America	Adulthood

(Taken from IARC, 1994, p. 59.)

Hepatitis B virus

- Hepatotropic, noncytopathic, 3.2-kb circular DNA
- Acute, chronic hepatitis, hepatocellular carcinoma (HCC)
- Over 2 billion people infected
- Over 350 million people chronically infected
- 100-fold increased risk (40% lifetime risk) of HCC
- Over 1 million deaths each year

Figure 16 HBV, disease association and frequencies. (Adapted from Chisari, 2000.)



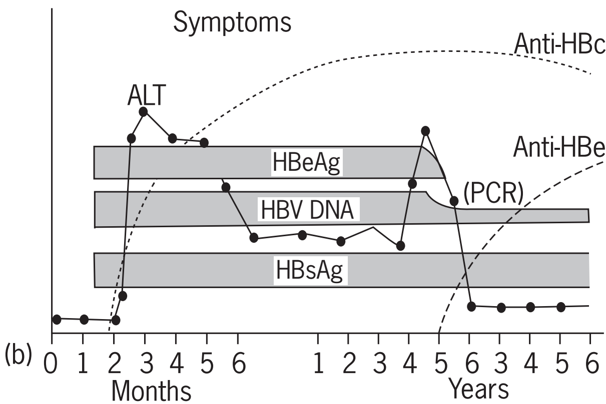


Figure 17 Serological and molecular patterns of HBV expression in (a) acute and (b) chronic HBV infections. Hatched bars show patterns of antigen expression; broken lines indicate periods in which HBV DNA (as noted) is detectable by PCR. Expression of alanine aminotransferase (ALT) used as control. (From IARC, 1994, p. 60.)

A hypothesis, following on from infection through to chronic illness and HCC, as outlined in **Figure 18** (Chisari *et al*., 2000), draws on all the factors noted above, and others, in predicting a course for disease. Such a complex scenario would emphasize the need for more than viral surveillance for the eradication of HCC.

Immune Biology of HBV and Immunoprophylaxis

The host–cell interactions that allow for the persistence of a virus, and the failure of the immune system to eliminate it in an immunocompetent individual, is a topic of considerable relevance for the DNA tumour virus field. For largely noncytopathic viruses, such as HPV, EBV, KSHV and HBV, they must either overwhelm an effective immune response or adopt mechanisms that allow for avoidance, as suggested by one or more of the hypothesized routes for progression from infection to generation of hepatocellular carcinoma (**Figure 18**). One approach for EBV therapy, as discussed elsewhere, assumes that the immune system may need to, and can, be stimulated specifically to recognize viral genes that might be expressed in its associated tumours, with beneficial effects. As noted, however, realistically such an approach is aimed at reducing morbidity, rather than effecting cure (Khanna *et al*., 1999). Such an approach may be even more valid for HBV, which can infect virtually all the hepatocytes in the liver, suggesting that the number of infected cells might actually outnumber relevant antigen-specific T-cells by several orders of magnitude in acute hepatitis infections and HCC. Thus, in individual situations where there are not sufficient cytotoxic T cells to contain infection, stimulation of the immune system might prove effective. On the other hand, there are studies to suggest that in some individuals with acute hepatitis, even in the presence of a vigorous T cell response, not all virus may be cleared. This has been explored using sensitive PCR assays where, several decades after complete clinical and serological recovery from this disease, low levels of viral DNA were detected in sera and PBMCs in some cases. Whether this result represents individuals at risk of reinfection and/or developing HCC for other, possibly genetic reasons, remains to be assessed. Notably, in HBV vaccine studies in Taiwan on children of different ethnic origins, unidentified host factors were postulated to explain the hyporesponsiveness seen among some populations (Hsu, 1996).

For HBV, it was earlier demonstrated that serum containing HBsAg retained its immunogenic properties even after heat inactivation. This observation proved the basis for plasma-derived, and later recombinant, s-antigen vaccines against HBV. In the 1980s, large-scale vaccination studies on children were initiated in both Taiwan and China, countries with the highest rates of HBV endemicity in the world, with the reasonable expectations that universal childhood immunization would allow HBV infections to be controlled in these areas within a few generations. In some high-risk areas, vaccination programmes to immunize every newborn child have been initiated and subsequent HCC incidence in these areas is being carefully monitored. As this malignancy, as with

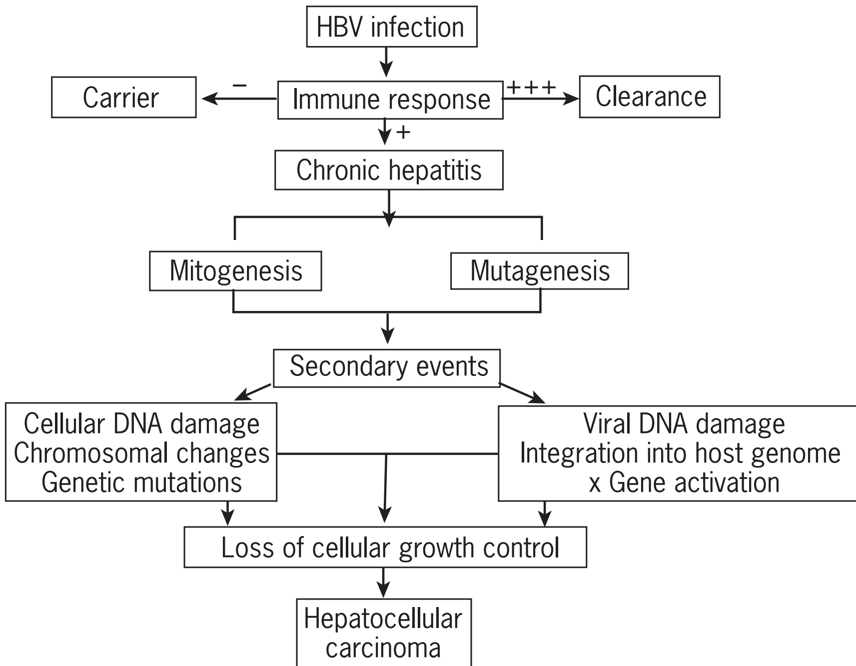


Figure 18 Hypothesis. Route from HBV infection to the genesis of hepatocellular carcinoma (HCC). (Adapted from Chisari, 2000.)

EBV-associated NPC, is mainly confined to adult populations, the data on this topic are not yet available. However, in Taiwan 10 years on, for children who completed a complete course of vaccination (four doses), anti-HBV antigen antibodies were detectable in a high proportion (82%) of them. Interestingly, in China it has now been found that tree shrews (*Tupaia belangeri chinenesis*) can be infected with human HBV and the infection passed to offspring in a high proportion of cases. Thus, a useful model may evolve for studying many of the unanswered questions about the relationship of HBV infection, the development of HCC and the variable responses to vaccination that have been observed.

OTHER DNA VIRUSES

Studies on growth changes in cells in culture or tumours produced by other DNA viruses in model (frequently immunoincompetent) animals, or their isolation from human tumours, have led to their tentative assignment as oncogenic, or potentially oncogenic, viruses. These include the ubiquitous human Polyomavirus, BK, that has been detected in brain tumours of different histological types, and also in KS, osteosarcomas and kidney carcinomas. Another human Polyomavirus, JC, causally associated with a pathological condition, PML (progressive multifocal leucoencephalopathy), has also been explored with regard to tumour formation. Both BK and JC encode large T antigens (LTs) that are related to that of SV40, and *in vitro* their genes stimulate cell growth (Barbanti-Brodano *et al.*, 1998). However, any significance in malignant growth in humans has not yet been established, although BK remains a candidate human oncogenic virus. The Adenoviruses, of which there are many distinct strains, have genes (*E1A* and *E1B*) that act as oncogenes in culture, interacting with tumour-suppressor genes. They, particularly, Ad12, can induce tumour formation, at least in animal models. Likewise, the human herpesvirus 6 (HHV-6) can transform mouse and human epidermal keratinocytes in culture, generating cell lines that are tumorigenic in athymic mice. Herpes-simplex virus (HSV) sequences have been found in human cervical cancer, but research on this topic was curtailed with the more definitive discovery of papillomaviruses in this malignancy. To date, roles for neither HSV nor HHV-6 in human malignancies have been established. The best evidence for viral causation of human malignancies appears to lie with the four DNA viruses dealt with in detail above, and with some of the RNA viruses, considered elsewhere, although it would be unwise to assume that other, undiscovered, candidate human oncogenic viruses do not exist. Newer, more sensitive methods of analysis make the discovery of novel cancer-related viruses a challenging research field.

REFERENCES

Arrand, J. R. (1994). Molecular genetics of human papillomaviruses. In: Stem, P. L. and Stanley, M. A. (eds), *Human Papillomavirus and Cervical Cancer: Biology and Immunology.* Chap. 2, 28–40. Oxford Medical Publications, Oxford.

Baer, R., *et al*. (1984). DNA sequence and expression of the B95-8 Epstein–Barr virus genome. *Nature*, **310**, 207–211.

Barbanti-Brodano, G., *et al*. (1998). BK and JC human polyoma viruses and simian virus 40: natural history of infection in humans, experimental oncogenicity and association with human tumors. *Advances in Virus Research*, **50**, 69–99.

Chang, Y., *et al*. (1994). Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science*, **266**, 1865–1869.

Chisari, F. V. (2000). Viruses, immunity and cancer: lessons from hepatitis B. *American Journal of Pathology*, **156**, 1118–1132.

Chow, L. T. and Broker, T. R. (1994). Papillomavirus DNA replication. *Intervirology*, **37**, 150–158.

Dane, D. S., *et al*. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet*, **1**, 695–698.

Frazer, I. H. (1996). Immunology of papillomavirus infection. *Current Opinion in Immunology*, **8**, 484–491.

Ganem, D. (1998). Human herpesvirus 8 and its role in the genesis of Kaposi's sarcoma. *Current Clinical Topics in Infectious Diseases*, **18**, 237–251.

Glenn, M., *et al*. (1999). Identification of a spliced gene from Kaposi's sarcoma-associated herpesvirus encoding a protein with similarities to latent membrane proteins 1 and 2A of Epstein–Barr virus. *Journal of Virology*, **73**, 6953–6963.

Griffin, B. E. (1998). Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis. *Reviews in Medical Virology*, **8**, 61–66.

Griffin, B. E. (2000). Epstein–Barr virus (EBV) and human disease: facts, opinions and problems. *Mutation Research*, **462**, 395–405.

Griffin, B. E. and Xue, S.-A. (1998). Epstein–Barr virus infections and their association with human malignancies: some key questions. *Annals of Medicine*, **30**, 249–259.

Hagensee, M. E., *et al.* (1994). Three-dimensional structure of vaccinia virus produced human papillomavirus type 1 capsids. *Journal of Virology*, **68**, 4503–4505.

Hitt, M. M., *et al*. (1989). EBV gene expression in an NPC tumor. *EMBO Journal*, **8**, 2639–2651.

Howley, P. M. (1996). Papillomavirinae and their replication. In: Fields, B. N., *et al*. (eds), *Virology*, 3rd edn. Chap. 65, 2045–2076 (Lippincott-Raven, Philadelphia).

Hsu, L. C. (1996). Ethnic differences in immune responses to hepatitis B vaccine. *American Journal of Epidemiology*, **143**, 718–724.

IARC (1994). *Hepatitis Viruses*, Monograph **59** (IARC, Lyon).

IARC (1995). *Human Papillomaviruses*, Monograph **64** (IARC, Lyon).

IARC (1997). *Epstein–Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8*, Monograph **70** (IARC, Lyon).

Khanna, R., *et al*. (1999). Vaccine strategies against Epstein–Barr virus-associated diseases: lessons from studies on cytotoxic T-cell-mediated immune regulation. *Immunological Review*, **170**, 49–64.

Labrecque, L. G., *et al*. (1999). Expression of Epstein–Barr virus lytically related genes in African Burkitt's lymphoma: correlation with patient response to therapy. *International Journal of Cancer*, **81**, 6–11.

Leight, E. R. and Sugden, B. (2000). EBNA-1: a protein pivotal to latent infection by Epstein–Barr virus. *Reviews in Medical Virology*, **10**, 83–100.

Morse, S. S. (1993). *Emerging Viruses* (Oxford University Press, Oxford).

Nakagawa, M., *et al*. (1997). Cytotoxic T lymphocyte responses to E6 and E7 proteins of human papillomavirus type 16: relationship to cervical intraepithelial neoplasia. *Journal of Infectious Diseases*, **175**, 927–931.

Pfister, H. and Fuchs, P. G. (1994). Anatomy, taxonomy and evolution of Papillomaviruses. *Intervirology*, **37**, 143–149.

Russo, J. J. (1996). Nucleotide sequence of the Kaposi's sarcoma-associated herpesvirus (HHV 8). *Proceedings of the National Academy of Sciences of the USA*, **93**, 14862–14867.

Shah, K. V. (1990). Biology of human genital tract papillomaviruses. In: Holmes, K. G., *et al*. (eds), *Sexually Transmitted Diseases*, 2nd edn. Chap. 37, 425–431 (McGraw-Hill, New York).

Shah, K. V. and Howley, P. M. (1996). Papillomaviruses. In: Fields, B. N., *et al*. (eds), *Virology*, 3rd edn. Chap. 66, 2077–2109 (Lippincott-Raven, Philadelphia).

Stöppler, H., *et al*. (1994). Transforming proteins of the papillomaviruses. *Intervirology*, **37**, 168–179.

Sun, R., *et al*. (1999). Kinetics of Kaposi's sarcoma-associated herpesvirus gene expression. *Journal of Virology*, **73**, 2232–2242.

Tarpey, I., *et al*. (1994). Human cytotoxic T lymphocytes stimulated by endogenously processed human papillomavirus type 11 E7 recognize a peptide containing a HLA-A2 (A^*0201) motif. *Immunology*, **81**, 222–227.

Van Ranst, M., *et al*. (1992). Phylogenetic classification of human papillomavirus: Correlation with clinical manifestations. *Journal of General Virology*, **68**, 4503–4505.

FURTHER READING

Arrand, J. R. and Harper, D. R. (1998). *Viruses and Human Cancer.* (BIOS Scientific Publishers, Oxford).

Boshoff, C. (1998). Kaposi's sarcoma associated herpesvirus. In: Newton, R., *et al*. (eds), *Cancer Surveys* 33. Infections and Human Cancer.157–190. (Imperial Cancer Research Fund, London).

Evans, A. S. and Kaslow, R. A. (eds) (1997). *Viral Infections of Humans. Epidemiology and Control*, 4th edn. (Plenum Medical Books, New York).

Fields, B. N., *et al*. (eds) (1996). *Fields Virology*, 3rd edn, Vol. 2. (Lippincott-Raven, Philadelpia).

Reitz, M. S., *et al*. (1999). Perspectives on Kaposi's sarcoma: facts, concepts and conjectures. *Journal of the National Cancer Institute*, **91**, 1453–1458.

Rezza, G., *et al*. (1999). Human herpesvirus 8 seropositivity and risk of Kaposi's sarcoma and other acquired immunodeficiency syndrome-related diseases. *Journal of the National Cancer Institute*, **91**, 1468–1474.

Yan, R. Q., *et al*. (1996). Human hepatitis B virus and hepatocellular carcinoma. I. Experimental infection of tree shrews with hepatitis virus. *Journal of Cancer Research and Clinical Oncology*, **122**, 283–288.

Websites

The human papillomavirus database: http://hpv-web.lanl.gov.

Virus databases on-line: http://www.ncbi.nlm.nih.gov/ICTV; http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy; http://www.virology.net.

Chapter 6

RNA Viruses

Mitsuaki Yoshida
Banyu Tsukuba Research Institute, Tsukuba, Ibaraki, Japan

CONTENTS

- General Description of RNA Tumour Viruses
- Human Tumorigenic Retroviruses: HTLVs

GENERAL DESCRIPTION OF RNA TUMOUR VIRUSES

Introduction

It is well established that tumour formation proceeds through multiple steps, each of which consists of mutation and selection of the mutated cells. Various signals for induction of cellular proliferation and fixation of the abnormal phenotypes by genetic mutation promote cellular conversion into a more malignant state. External cellular signals are transduced into the nucleus through multiple pathways and finally induce specific gene expression and cellular responses. Genetic alterations of these signalling pathways, transcriptional machinery or the target genes themselves are the origin of cancers.

Infection by some specific viruses represents one of these genetic alterations and triggers these multiple-step processes for the final induction of specific cancers in both animals and humans. In human cancers, the real causes of the cancers are mostly unknown, and therefore there is no effective way to diagnose healthy individuals who will develop cancers in their near future. This is possible, however, when the cancers result from specific viral infection since the virus carriers are easily identified by their antibodies against specific viruses. Such situations provide an opportunity not only to prevent the cancers, but also to investigate the mechanism of cancers directly in humans.

As human tumour viruses, human T cell leukaemia virus (HTLV) type 1 (HTLV-1), hepatitis C-type virus (HCV), hepatitis B-type virus (HBV), human papillomavirus (HPV) and Epstein–Barr virus (EBV) are well established. Among these, the first two, HTLV-1 and HCV, are the RNA viruses, i.e. they have RNA as their genomes. These two viruses, however, are classified into unrelated groups, HTLV-1 to the retroviridae and HCV to the flabiviridae, which are different in many respects. In this chapter, retroviruses are the main subject.

Viral Replication

The term RNA tumour viruses generally represents retroviruses that contain reverse transcriptase (RT), which transcribes genomic RNA into DNA upon infection. Cores of the viral particles contain two copies of a single-stranded, positive RNA, Gag protein and reverse transcriptase. The core is enveloped with membrane similar to the plasma membrane of host cells, on which the viral Env (envelope) glycoprotein is exposed. The interaction of the Env protein with a receptor on a target cell membrane is required for infection. Interaction of the Env with the receptor induces membrane fusion between the viral particle and cell, allowing the core to be incorporated into the cell, then the genomic RNA is reverse transcribed into complementary DNA (cDNA) by the particle RT (**Figure 1**). The cDNA is then converted to double-stranded DNA and integrated into the host chromosomal DNA forming a 'provirus.' During these processes, a long terminal repeating sequence (LTR) is formed at both ends which contains many elements essential to viral gene

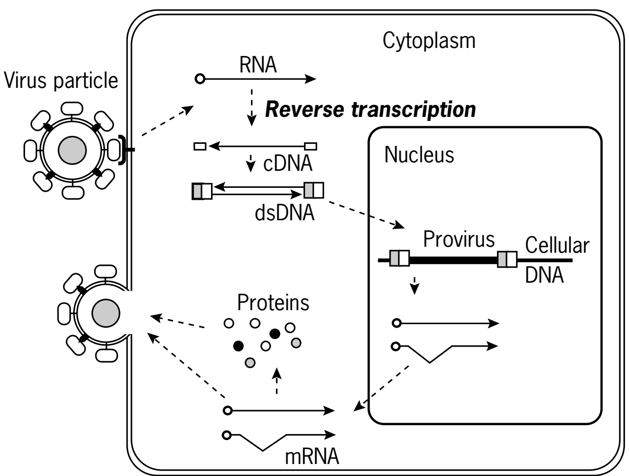


Figure 1 Life cycle of retroviruses.

expression and replication. Details of the mechanism of replication have been reviewed elsewhere (Weiss *et al*., 1985). Once the proviruses are integrated into germ cells, the retroviruses can be transmitted vertically to successive generations of hosts by transmission of proviral genomes (endogenous viruses), in addition to the standard viral replication and infection (exogenous viruses).

The proviruses are transcribed into RNA by the cellular transcription machinery and a subpopulation of the viral transcripts is spliced into subgenomic mRNA. Both spliced and unspliced viral RNA species are translated into viral proteins; genomic RNA (unspliced RNA) into Gag and Pol proteins and subgenomic RNA into Env protein. The viral proteins and genomic RNA are assembled at specific sites under the plasma membrane into the particles and released by budding. Generally, retroviral replication is not harmful to the host cells except in some cases such as human immunodeficiency virus (HIV). The receptors for a few retroviruses have been identified as membrane proteins, and their expression in many types of cells is consistent with the array of cells at risk of retroviral infection.

Genome Structures and Tumorigenesis

There are two types of retroviral genome; one type carries and the other does not carry host-derived oncogenes (**Figure 2**). Replication-competent retroviruses generally have *gag*, *pol* and *env* genes but do not carry oncogenes. Some of these viruses induce leukaemia or lymphoma after a long latency and are thus called chronic leukaemia viruses. This type of retrovirus induces leukaemia through activation of expression of the adjacent cellular genes by the integrated LTRs. When a provirus is, by chance, integrated in the vicinity of a proto-oncogene and induces abnormal expression of the proto-oncogene it may lead to tumorigenesis. This 'promoter-insertion mechanism of viral carcinogenesis' operates in a variety of tumour systems; *Myc* activation by avian leukaemia virus (ALV) in B cell lymphoma, *erbB* activation by ALV in chicken erythroblastosis and *int1* and *int2* activation by mouse mammary tumour virus (MMTV) in mouse mammary tumours. Because proviral integration is not site specific, repeated integration through viral replication is usually required before the provirus appears in a tumorigenic site. This explains the long latency after infection for tumorigenesis; however, once it integrates into the right place of the host genome, viral replication is no longer required for tumorigenesis.

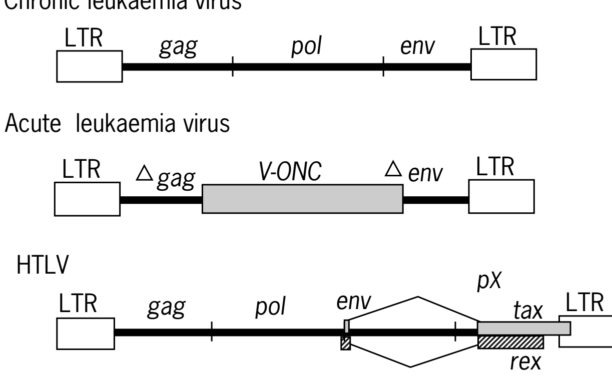


Figure 2 Genome structure of proviruses.

Endogenous viruses, which are proviruses vertically transmitted through germ cells, mostly have this type of genome. Many copies of the endogenous viruses are maintained in human cells, as many as thousands in some cases, but they are not very replicative even though they have complete genomes. Various indications of possible participation of endogenous viruses in human cancers have been reported, but, it is still a subject for further study.

The other type of retroviruses carry an oncogene derived from a cellular proto-oncogene (Weiss *et al*., 1985). Capture of a oncogene in the viral genome, which is derived from host cell DNA, in turn results in a deletion of some portion of the viral genome. Consequently, acute sarcoma/leukaemia viruses are generally replication defective. This type of virus has to be infected together with a chronic leukaemia virus as a 'helper' for their replication. One exception is Rous sarcoma virus, which has the oncogene *src* between *env* and the *3′LTR* and is competent in replication. The viral oncogene of the acute leukaemia viruses is responsible for transformation of infected cells and does not require viral replication or site-specific integration, inducing specific tumours with short latency. In contrast to the broad specificity of retroviral infection *in vitro*, these viruses are able to induce tumours in relatively few tissues *in vivo*. Specificity of retroviral tumorigenesis is restricted by either the class of viral oncogene or the tissue-specific promoter activity of the integrated *LTR*, in addition to the nature of the viral receptors on the host cells.

HUMAN TUMORIGENIC RETROVIRUSES: HTLVs

Human retroviruses include HIV, human endogenous viruses and human foamy viruses. The replication and pathology of HTLV-1 are considered throughout this chapter, but with some exceptions, observations on HTLV-1 are applicable to the other members of the HTLV group (see the following section).

HTLVs and Disease

HTLV-1 is an aetiological factor in adult T cell leukaemia (ATL) (Poeisz *et al*., 1980; Hinuma *et al*., 1981; Yoshida *et al*., 1982). ATL is a unique T cell malignancy with a CD4-positive phenotype. HTLV-1 prevalence and ATL

are clustered in southwestern Japan. The sera of ATL patients contain antibodies that react specifically with cell lines established from the ATL patient. The antigens in these cell lines were later proved to be HTLV-1 proteins by molecular characterization of the viral genome. The extensive epidemiological studies on antibodies and ATL clearly established that HTLV-1 is closely associated with ATL. The nationwide and worldwide epidemiology indicated that HTLV-1 infection is also associated with a neurological disease, HAM/TSP (HTLV-1-associated myelopathy/tropical spastic paraparesis).

The genome of HTLV-1 provirus consists of *LTR–gag–pol–env–pX–LTR* (Seiki *et al.*, 1983) and is distinct from the standard retroviral genome, *LTR–gag–pol–env–LTR* (**Figure 2**). The existence of a functional *pX* sequence is unique to this virus, thus HTLV-1 forms an independent retroviral group, HTLV. After characterization of HTLV-1, another virus, HTLV-2, was isolated from a patient with hairy T-cell leukaemia (Chen *et al.*, 1983). The genome is about 60% homologous to HTLV-1, but its pathogenicity is not yet established.

Viruses similar to HTLV-1, simian T cell leukaemia viruses (STLV), were isolated from various species of nonhuman primates, including chimpanzee. The STLVs share a 90–95% identity of genomic sequence with HTLV-1 and some isolates show more homology than those among human isolates. Although STLV is widely distributed in monkeys, no typically leukaemic animals have been observed. A few cases of a leukaemia-like disease have been noted in STLV-infected monkeys in zoos, but an aetiological connection between STLV and the disease remains to be established. Another member of the HTLV group is bovine leukaemia virus (BLV). BLV infects and replicates in B cells of cows and induces B cell lymphoma. BLV also infects lymphocytes of sheep and induces leukaemia after a short latent period.

Epidemiology

Nearly all ATL patients carry antibodies against HTLV-1 proteins. These antibodies are easily detected by indirect immunostaining of cells infected with HTLV-1, by an enzyme-linked immunosorbent assay (ELISA), by a particle-agglutination assay or by Western blotting. Some populations of healthy adults also carry HTLV-1 antibodies and these sero-positive persons are defined as the viral carriers. In fact, HTLV-1 can be detected in such individuals using the polymerase chain reaction (PCR). In considering the epidemiology of HTLV-1, geographic clustering, age-dependent prevalence, genomic variation and geographic origin of HTLV-2 are informative.

Geographic Clustering

HTLV-1 antibodies are recognized in 5–15% of adults clustered in southwestern Japan, the Caribbean islands and South America, Central Africa and Papua New Guinea and the Solomon Islands in Melanesia. The prevalence of sero-positive adults varies significantly from one district to another within these areas of endemicity. For example, in a particular isolated island in Kyushu, Japan, 30–40% of people over 40 years of age might be infected, whereas on a neighbouring island, the prevalence may be far lower. Significantly, ATL and HAM/TSP are also clustered, overlapping HTLV-1 in distribution. ATL patients and healthy carriers found sporadically in nonendemic areas mostly originate from one of the endemic areas.

HTLV-2 is frequently isolated in the United States and other countries from intravenous drug abusers and persons infected with HIV. HTLV-2 is endemic in South America, and also in Pygmy populations in Africa.These regions are likely to be the natural reservoir of HTLV-2 sporadically observed in the other places.

Age-dependent Prevalence

The prevalence of virus carriers increases with age after 20 years increasing sharply around 40–50 years of age and reaching a maximum in people aged between 50 and 60 years. The prevalence is significantly (1.6 times) higher in females than in males, but the incidence of ATL is similar in both sexes. The increase in prevalence among females can be attributed to sexual transmission of HTLV-1 from husbands to wives and also suggests that such infection is not leukaemogenic.

Although global epidemiology identified the age-dependent increase of antibody prevalence, cohort studies over 10 years in Japan revealed that sero-conversion of adults from antibody negative to positive is very rare. These observations are unable to explain the sharp increase of sero-prevalence in the 40–50 years age group. After extensive epidemiological studies of antibody prevalence, it is now accepted that the age-dependent increase is a reflection of the reduction of the infection risk at the early stage of life. Artificial milk became popular around 40 years ago in these areas in Japan, and thus reduced the incidence of breast milk-born infection of HTLV-1 (discussed below).

Genomic Stability

The viral genome is well conserved (over 96%) in Japan and the Caribbean area. Viral isolates from Africa and Papua New Guinea may vary somewhat more, but the variations are very limited. Retroviral genomes are thought to be unstable relative to those of other viruses because reverse transcription has no proofreading mechanism. Such genomic stability stands in sharp contrast to the highly labile genome of HIV. This may be associated with the very low competency of HTLV-1 replication *in vivo*.

Infection of HTLV-1

Infection In Vitro

Viral particles of HTLV-1 show extremely low infectivity *in vitro*, but co-cultivation with virus-producing cells can transmit HTLV-1 to a variety of human cells, including T and B lymphocytes, fibroblasts and epithelial cells, and also cells from monkeys, rats, rabbits and hamsters but, curiously, not mice. In these infected cells, the provirus is integrated into random sites in the chromosomal DNA, and most of the viral genes are successfully expressed. However, in non-T cell lines, the integrated proviruses become latent in expression, otherwise usually inducing the fusion of infected cells forming syncytia that ultimately die.

Only CD4-positive T cells are frequently immortalized upon infection with HTLV-1 (Miyoshi *et al.*, 1981) and the immortalized cells express high levels of IL-2Rα, proliferating in an IL-2-dependent fashion. Animal retroviruses that do not carry an oncogene generally do not immortalize cells *in vitro*, and therefore immortalization by HTLV-1 appears to be unique and suggests that the virus may have a particular function. Accordingly, a contribution of *pX* to this effect has been proposed.

In contrast to *in vitro* infection, the cells of both ATL patients and asymptomatic viral carriers infected *in vivo* are almost exclusively T cells with the CD4 + phenotype. Furthermore, infected cells *in vivo* do not express viral information at significant levels. Reverse transcriptase-mediated PCR (RT-PCR) on mRNA indicates that over 95% of infected cells fail to express viral genes *in vivo* irrespective of whether they are in a transformed or non-transformed state.

Natural Transmission

HTLV-1 can be transmitted *in vivo* through (1) blood transfusion, (2) nursing with breast milk, and (3) sexual relations.

Blood Transfusion

Retrospective studies of blood transfusions showed that 60–70% of recipients of fresh, sero-positive blood were infected with HTLV-1. Transfer of infected cells from donor to recipient is required for viral transmission, and therefore fresh, sero-positive plasma does not support the infection. Blood transfusion-mediated transmission of HTLV-1 seems not to induce ATL (see Sexual Transmission, below), but does induce HAM/TSP. Therefore, rejection of HTLV-1-positive blood can protect recipients against both HTLV-1 infection and development of HAM/TSP. The blood-mediated transmission of HTLV-1 explains a high prevalence of up to 20% of abusers of intravenous drugs by the sharing of unsterilized needles.

Mother to Child

Viral transmission from mother to child was originally suggested by epidemiological evidence: most mothers of sero-positive children were carriers of the virus and about 30% of the children of sero-positive mothers were themselves sero-positive. Neonatal infection was initially suspected, but surveys of lymphocytes in cord blood from a large number of children born to sero-positive mothers have virtually, but not completely, excluded this possibility. Instead, breast milk was found to be a likely source of transmissible virus. Supporting this, milk taken from sero-positive mothers and given to adult marmosets leads to the appearance of antibodies in these monkeys. More direct evidence stems from a practical trial demonstrating that cessation of breast-feeding by sero-positive mothers drastically reduced the sero-conversion rates of their children (see the last section).

Sexual Transmission

Wives with sero-positive husbands are very frequently sero-positive. Conversely, the husbands of sero-positive wives show the same frequency of sero-positivity as do men of the region under study. On these grounds, it seems that the virus can be transmitted from husband to wife but not vice versa. Infected T cells have been found in semen from men infected with HTLV-1 and these cells are considered to transmit the virus from male to female. The higher rate of sero-positivity in female (1.6 times) is explained by this transmission. The sex-specific incidence of ATL does not mirror this difference, suggesting that HTLV-1 infections sexually transmitted to females are not leukaemogenic to ATL.

Viral Gene Expression

Proviral Genome Unique to HTLV-1

The HTLV-1 proviral genome cloned from leukaemic-cell DNA from an ATL patient is 9032 bp long and consisted of *LTR–gag–pol–env–pX–LTR* (Seiki *et al.*, 1983). The presence of a *pX* region on the 3′ side of the *env* gene distinguishes the HTLV-1 genome from those of other retroviruses (**Figure 2**).

In general in retroviruses, the LTRs function as units regulating viral gene expression and replication. Furthermore, the *pX* region of HTLV contains additional, overlapping, regulatory genes (**Figure 3**): *tax*, *rex* and a gene whose function remains unknown, which encode p40tax, p27rex, and p21x. The Tax protein, p40tax, is a potent *trans*-activator of proviral transcription (Sodroski *et al.*, 1984; Fujisawa *et al.*, 1985) and thus is essential to viral gene expression (Yoshida, 1995). The second protein, p27rex, is a *trans*-acting modulator of RNA processing, which allows expression of the unspliced *gag* and *env* mRNAs in the cytoplasm. Expression of these unspliced

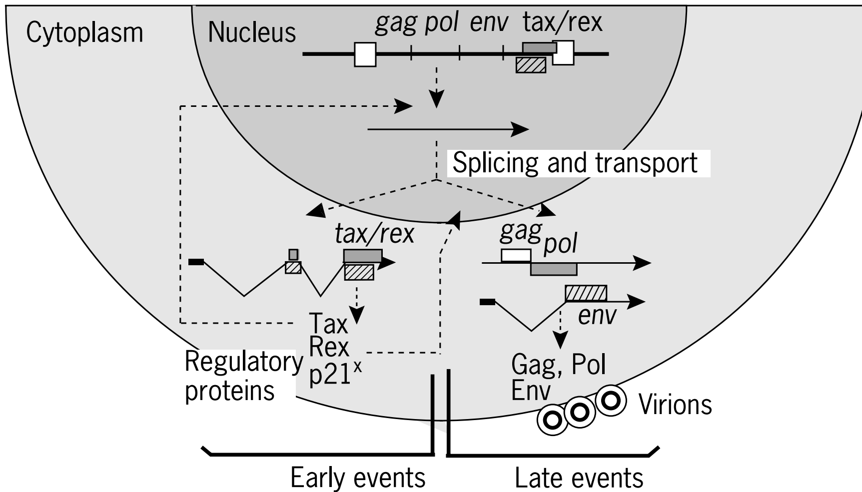


Figure 3 Transcription and viral RNA processing.

mRNAs is essential for expression of the viral structural proteins. Thus, Rex is also essential for HTLV-1 gene expression and replication. These systems are unique among retroviral regulations and have similarities to Tat and Rev systems of HIV, which are also essential for the replication of HIV.

The *pX* sequence is also able to encode various proteins when the sequence is alternatively spliced. The function of some of these products was characterized using expression vectors, but the physiology of these proteins remains to be analysed.

Transcriptional Activation

Retroviral LTR elements contain a TATA box, a transcriptional enhancer and a poly(A) signal, all of which are essential for viral gene expression and replication of RNA tumour viruses. These elements are recognized by cellular transcriptional factors for RNA polymerase II and retroviral gene expression depends upon the cellular machinery of the host. In addition to LTR-mediated regulation, HTLV-1 contains the *tax* gene, which acts in trans to stimulate viral transcription (**Figure 3**). Tax function depends on three direct repeats of the 21-bp sequence transcriptional enhancer, in the LTR. The interaction of Tax with the enhancers has been proposed to be indirect, that is, Tax interacts with a cellular protein that binds to the enhancer DNA to mediate trans-activation.

In addition to the viral genome, Tax activates cellular genes. Since the transcription of the gene for the α-chain of the IL-2 receptor (IL-Rα) was reported to be activated by Tax, many other genes were identified also to be activated but through different mechanisms (see later). These include the gene for GM-colony-stimulating factor (GM-CSF), the proto-oncogenes c-*fos* and c-*jun*, the genes for parathyroid hormone-related protein (PTHrP), MHC class I antigen and many others.

RNA Processing

The primary transcript, the viral genome, of HTLV-1 contains genes for Gag, Pol, Env, Tax, Rex and a few others in this order with some overlapping. Only the first coding frame on an mRNA is translated into protein in eukaryotic cells, and therefore the retrovirus needs to splice the primary transcript into various species of viral mRNA to encode other proteins. This splicing is regulated by Rex, a pX protein of HTLV-1 (**Figure 3**). Viral replication requires three species of mRNA: 8.5-kb genomic (unspliced) RNA as *gag* and *pol* mRNA, a 4.2-kb singly spliced sub-genomic RNA as *env* mRNA and a 2.1-kb doubly spliced mRNA for the expression of Tax and Rex proteins. The viral transcripts early after infection or induction are all spliced into completely spliced *tax*/*rex* mRNA. Tax and Rex are then expressed, and viral transcription is enormously enhanced and at the same time, unspliced RNA for Gag, Pol and Env are expressed in the cytoplasm. This regulation is essential for viral protein expression, since host-cell mRNAs are generally all spliced into mRNA to be transported to the cytoplasm.

Rex function in turn reduces the level of spliced mRNA that encodes regulatory proteins, including Tax protein, and eventually reduces viral transcription. Thus, Rex enhances the expression of viral structural proteins, but suppresses total viral gene expression. In short, Rex exerts feedback control of the viral gene expression, thus resulting in a transient expression of HTLV-1.

For this regulation, target RNA requires to have a cis-acting element (RxRE) consisting of 205 nucleotides located in the 3′ region of the viral RNA. The unique secondary structure of this element allows Rex protein to bind to it. A nuclear export signal in the Rex protein suggests transport of a Rex–RNA complex into the cytoplasm without processing. HIV has a *rev* gene, strikingly similar in function to *rex* of HTLV-1. In the HIV system, Rev protein binds to RvRE in the *env* coding sequence.

Adult T Cell Leukaemia (ATL)

It is now established that HTLV-1 is associated with ATL, HTLV-1-associated myelopathy or tropical spastic paraparesis (HAM/TSP) and uveitis. Other diseases such as chronic lung disease, monoclonal gammopathy, chronic renal failure, strongyloidiasis and nonspecific dermatomycosis are also suggested to associate with the viral infection; however, further systematic studies are required to establish the exact relationships. Here, only ATL is described.

The Leukaemic Cells

ATL cells are T cells with the CD4+ phenotype and, usually, a highly lobulated nucleus. These cells always carry HTLV-1 proviruse(s) and the site of integration is monoclonal in a given ATL patient. In 70–80% of cases of ATL patients examined, one copy of the complete provirus was integrated into each leukaemic cell. Occasionally, one or two copies of defective provirus are integrated into the DNA of a single cell. Even in the defective genomes, preservation of the pX region in defective proviruses suggests its importance in tumorigenesis.

The leukaemic cells express a high level of IL-2Rα on their surfaces. Production of PTHrP, IL-1β or GM-CSF by tumour cells has also been described. In almost all cases, leukaemic cells carry aberrant chromosomes, and there are frequently multiple abnormalities, such as trisomy of chromosome 7 and 14q11, 14q32 and 6/q15 translocations. The abnormality involving 14q32 was found in 25% of ATL patients, but the others appeared less frequently.

Clinical Features

ATL (Uchiyama *et al*., 1977) is classified into three phases, smouldering, chronic and acute phases, depending on clinical features. In smouldering ATL, patients commonly have from one to several per cent of morphologically abnormal T cells in their peripheral blood, but do not show other signs of severe illness and are therefore thought to be in an early stage of ATL development. In smouldering ATL, the abnormal cells are not aggressively malignant, but are HTLV-1 infected and expanded clonally. The onset of ATL is observed between 20 and 70 years of age, the peak rate of onset being in the 40s and 50s. The male-to-female ratio of ATL is 1.4:1.0. Symptoms of ATL vary from patient to patient, but are frequently complicated by skin lesions, enlargement of lymph nodes, liver and/or spleen and infiltration of leukaemic cells into the lungs and other organs. Patients usually have antibodies to HTLV-1 proteins, show an increased level of serum LDH and suffer from hypercalcaemia. The acute form, or phase, of ATL is aggressive and resistant to treatment; consequently, most patients in this phase die within 6 months of its onset.

Molecular Mechanism of Pathogenesis

Viral Function in Leukaemogenesis

The aetiological role of HTLV-1 in ATL has been demonstrated by sero-epidemiology and molecular biology of HTLV-1 and ATL. The bases of this are as follows: (1) ATL and HTLV-1 geographically overlap (population level); (2) most ATL patients are infected with HTLV-1 (individual level); (3) leukaemic cells from ATL patients are infected with HTLV-1 (cell level); (4) more importantly, the leukaemic cells show monoclonal integration of proviral DNA (molecular level); and (5) HTLV-1 has the capacity to immortalize human T cells *in vitro* (biochemical level). The evidence in (4) indicates that the leukaemic cells originated from a single HTLV-1-infected cell and, thus, that HTLV-1 plays a causative role in leukaemogenesis. There are estimated to be approximately one million carriers of HTLV-1 in Japan, and about 500 new cases of ATL are reported each year. About 2–5% of all carriers of HTLV-1 are thought to develop ATL during their life span (Tajima, 1990).

The site for provirus integration is monoclonal in ATL cells, but not the same among ATL patients (Seiki *et al*., 1984). Therefore, the promoter insertion model is unlikely since it requires a common integration adjacent to a proto-oncogene. Consequently, a 'trans-acting function' of HTLV-1 is postulated in leukaemogenesis. Molecular biology studies of HTLV-1 showed that the Tax protein functions as a 'trans-acting factor.' Consistent with these observations, Tax was found to immortalize T cells in an IL-2-dependent fashion, to transform rat embryonic cells in cooperation with c-*ras*, and to induce mesenchymal tumours in Tax transgenic mice. The central role of Tax in leukaemogenesis is thus proposed.

Trans-activation of Transcription

Activation of Enhancer Binding Protein

Tax trans-activates transcription via specific enhancers such as the 21-bp enhancer in the LTR, the NF-κB binding site in the gene for interleukin-2 receptor α and serum responsive element (SRE) in the c-*fos* gene . However, Tax is unable to bind directly to the enhancer DNA sequence. Instead, Tax binds to enhancer binding proteins; the first group includes CREB (cyclic AMP-responsive element (CRE) binding protein, CREM (CRE modulator protein), ATF-1 and ATF-2, which bind to the 21-bp enhancer in the HTLV-1 LTR. The second group is the family of NF-κB such as p50, p65, c-Rel and p52, which bind to the NF-κB binding site in IL-2 receptor α gene, and the third group is SRF (serum response factor) which binds to SRE (serum response element) in the c-*fos* or c-*egr* gene. These transcription factors are regulated by signal-dependent phosphorylation in normal cells; however, Tax binding permits

activation of these factors without specific phosphorylation, thus establishing constitutive activation of these genes.

It is now demonstrated that Tax binding to these transcription factors recruits CBP (CREB binding protein) on to the enhancer–CREB complexes (Kwok *et al*., 1994), because of Tax's affinity for CBP (**Figure 4**). CBP has a histone-acetylating activity and is normally unable to bind to CREB protein unless it is phosphorylated by protein kinase A. Therefore, the ternary complex of the enhancer–CREB–Tax–CBP without phosphorylation would acetylate histones bound to DNA nearby and thus activate transcription initiation. This simple binding hypothesis, however, may need further careful characterization.

Inactivation of Transcriptional Inhibitors

NF-κB proteins are activated by Tax binding in the nucleus (see the previous section). However, in resting cells, NF-κB proteins are retained in the cytoplasm by forming complexes with IκB. Stimulation of cells induces phosphorylation of IκB protein and its degradation. The released NF-κB then migrates into the nucleus because of its nuclear translocation signal and binds to its specific DNA sequences.

Tax also binds to IκBalpha which results in destabilization of IκB–NF-κB complexes. It is uncertain whether Tax induces phosphorylation of IκB and /or degradation of IκB. Activation of a protein kinase activity by Tax suggests the former mechanism, but the binding of Tax to a component of the proteasome may suggest the latter mechanism. Whatever the mechanism, Tax is able to down-regulate the transcriptional inhibitor, IκB, in the cytoplasm and thus activates NF-κB-dependent transcription. Tax appears to exert its effects via two independent mechanisms targeting an activator, NF-κB, and its inhibitor, IκB.

Trans-repression of Transcription

When the trans-activating function of Tax was extensively investigated, Tax was reported to trans-repress transcription of DNA polymerase beta (Jeang *et al*., 1990), which is required for repair of damaged DNA. Tax was then demonstrated to trans-repress the transcription of a set of growth-inhibitory genes such as *p18INK4c*, *NF-1*, *lck* and a apoptotic gene *bax*. Furthermore, Tax was also shown to trans-repress p53-dependent transcription, which affects the tumour-suppressor function of p53. These effects, therefore, suggest an abnormally enhanced proliferation of HTLV-1-infected cells. It is of interest to know how a trans-activator, Tax, is able to trans-repress other sets of genes.

The underlying mechanism was in fact shown to be rather simple, that is, inhibition of a transcriptional coactivator family, CBP/P300 (**Figure 4**). An E-box binding protein E47 in *p18INK4c* expression and p53 in p53-dependent transcription are essential to interact with CBP/p300 to achieve efficient expression. In these systems, Tax binds to CBP/p300 and interferes with the interaction between CBP/p300 and enhancer binding proteins, E47 or p53, resulting in trans-repression of specific transcription. This implies that Tax would be able to suppress many other genes since the CBP/p300 protein serves as a coactivator for a huge number of genes. It is therefore suggested that the cascade of Tax activity in transcriptional regulation

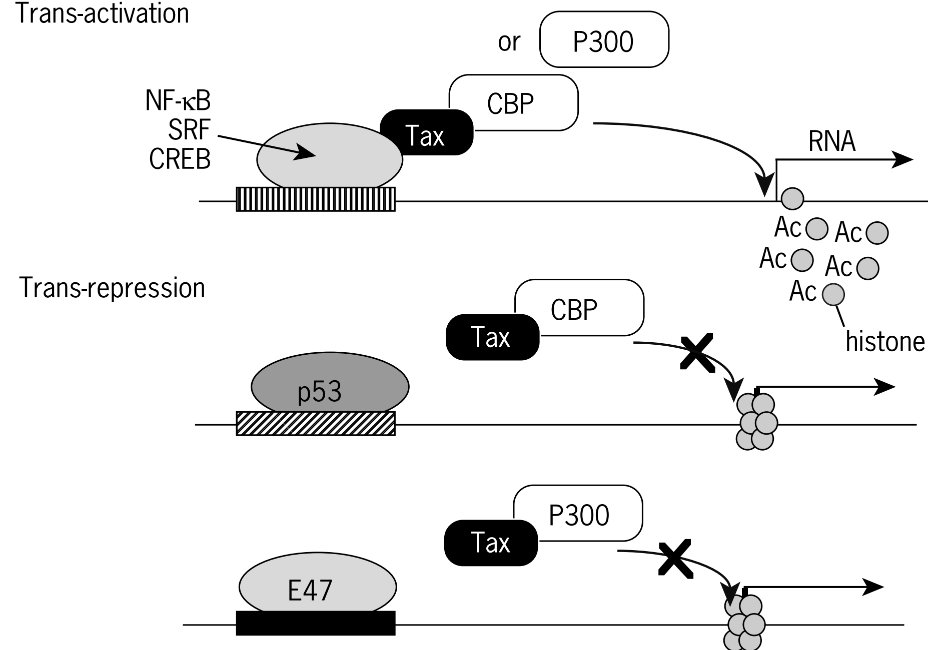


Figure 4 Tax binding to CBP or p300 to activate and repress the specific transcription.

would be unexpectedly wide in its targets and highly variable in its effect, depending on the level of expression.

Inhibition of Tumour-Suppressor Proteins

Independently of transcriptional regulation, the Tax protein was also found to interact with an inhibitor of cyclin-dependent kinase 4 (CDK4), p16INK4a (Suzuki *et al*., 1996) and a *Drosophila* large disc tumour-suppressor protein, Dlg. Furthermore, the direct binding inhibits the function of tumour-suppressor proteins. These suppressive effects on tumour-suppressor functions strongly suggest that Tax protein contributes to the development of ATL.

Cell Cycle Inhibitor

p16INK4a and p15INK4b are inhibitors of CDK4/6 and their inactivation results in activation of the kinases. Upon activation of CDK4/6, Rb is phosphorylated and is no longer able to bind E2F, thus releasing the active form of E2F, which then binds to target DNA sequences and initiates expression of various genes important for DNA synthesis. Tax binding to and inhibition of both p16INK4a and p15INK4b is able to activate CDK4 and promote cell entry into S phase (**Figure 5**). p16INK4a is frequently deleted in many human tumour cells, particularly in melanoma and haematopoietic tumours, and the deletion has been suspected to play a critical role in tumour induction and progression. Therefore, the functional knockout of the tumour-suppressor protein mimics the effect of gene deletion and may contribute to development of ATL.

With respect to the inactivation of Rb signalling pathway by p16INK4a, it is of interest to point out that DNA tumour viral proteins such as SV40 T-antigen, adenovirus E1A and papillomavirus E6/7 target Rb protein in their transformation (Nevins, 1992). The observation that developmentally unrelated tumour viruses share the common signalling pathway by targeting different molecules suggests that the Rb pathway is critical for the normal regulation of cell proliferation.

hDlg that Associates with APC, Another Tumour-suppressor Protein

Another target of Tax is hDlg, which is a signalling molecule downstream of Wnt/Frizzled and upstream of β-catenin in their signalling pathways. hDlg binds to the C-terminus of the tumour-suppressor protein APC involved in transducing cytostatic signals. Tax binds to the same domain of hDlg *in vitro* and *in vivo* and competes with APC, thus abrogating the growth-retarding signalling. In addition to the competitive displacement of APC from hDlg, Tax further induces hyperphosphorylation of hDlg as demonstrated by its slower migration in gel electrophoresis. It is noteworthy that APC and hDlg are significantly expressed in normal T cells (T. Suzuki and M. Yoshida, unpublished observation), hence these interactions might have roles in T cells, although these are not well understood.

Cell Cycle Check Point Protein, HsMAD1

Cell cycle processes are inspected at checkpoints to determine whether the scheduled processes are verified. Thus, once the checkpoint system is compromised, damaged cells can go through their cell cycle and proliferate, fixing genetic abnormality in the daughter cells. Tax of HTLV-1 can bind to the human homologue (HsMAD1) of yeast mitotic checkpoint protein MAD1 (Jin *et al*., 1998). HsMAD1 is a component of the mitotic

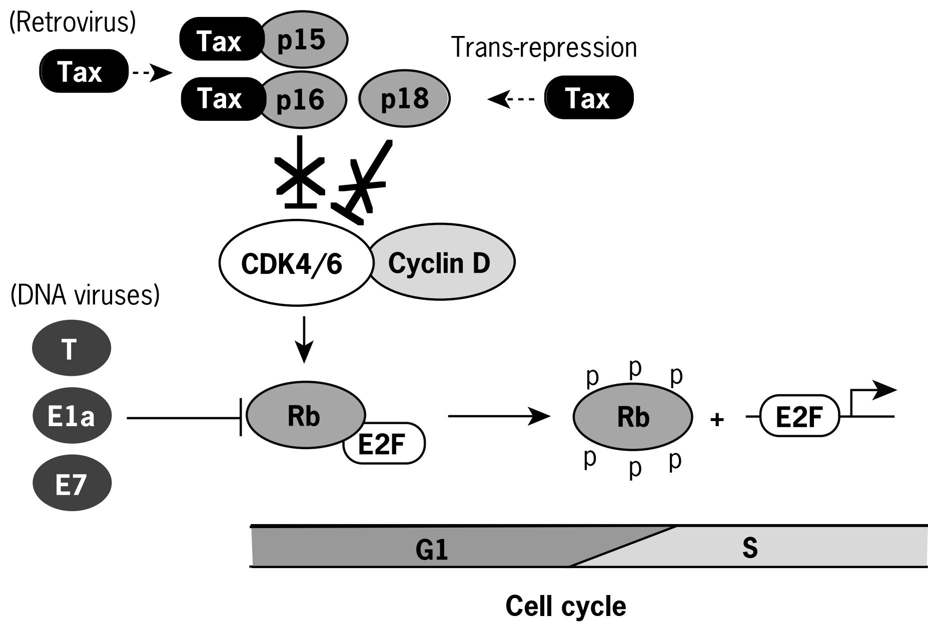


Figure 5 Tax binding to and inactivation of tumour-suppressor proteins, p16INK4a and p15INK4b.

checkpoint system which prevents anaphase and commitment to cellular division until chromosomal alignment is properly completed. Therefore, abrogation of the mitotic checkpoint function of HsMAD1 may be linked to chromosomal abnormalities which are observed at unusually high frequency in ATL cells.

Common Targets with DNA Tumour Viruses

The Tax protein appears to be pleiotropic through interacting with so many cellular regulators: activation and repression of transcription of different sets of genes, inhibition of CDK4 inhibitors and inhibition of tumour-suppressor proteins. Some of the target molecules and target pathways of Tax protein are shared by transforming proteins of DNA tumour viruses, T antigens of SV40, E1A/B of adenoviruses and E6/7 of papillomaviruses. The most striking shared target is the Rb signalling pathway: Tax targets p16INK4a and p15INK4b and the transforming proteins of DNA tumour viruses target Rb, both resulting in activation of E2F family and promotion of cells into the S phase of the cell cycle (**Figure 5**). Other examples shared are the transcriptional target coactivators CBP/p300 and hDlg; CBP/p300 is affected by T antigen and E1A, and hDlg by E6 of high-risk HPV and E4 9ORF1 of adenoviruses.

The genes for p16INK4a and Rb proteins are frequently mutated or deleted in spontaneous human cancer cells and are believed to play critical roles in tumorigenesis. Therefore, functional knockout of these gene products by viral proteins is similarly expected to play crucial roles in specific tumour induction. It is also interesting that developmentally unrelated viruses and spontaneous genetic mutations target the same signalling pathway for malignant transformation. By analogy with these spontaneous mutations or deletions, the virally induced functional inactivation of either p16INK4a or Rb protein would be primarily responsible for the induction of specific tumours in humans.

Possible Mechanism of Pathogenesis

Cooperation of the Pleiotropic Function of Tax Protein

As summarized in the previous section, Tax protein effects are pleiotropic and function at different levels, including transcriptional regulation, cell cycle promotion, tumour-suppressor function and genomic stability. It is reasonable to speculate that the pleiotropic functions of Tax would be mostly cooperative in ATL development (**Figure 6**). This interesting possibility is discussed below.

Cell proliferation. It has been demonstrated that various genes activated by Tax are in fact able to promote cell proliferation *in vitro*; transcriptional activation of lymphokines such as IL-6, GM-CSF, TGF-β and some others, lymphokine receptors such as IL-2Rα and some nuclear oncogenes such as c-*fos*, c-*egr* and c-*jun*, are all growth promoting. These transcriptional trans-activations are mediated independently through NF-κB or SRF and some others through CREB. Furthermore, trans-repression of *p18INK4c*, *lck* and *NF-1* would result in promotion of cell growth, since these genes are mostly growth-retarding. Direct inhibition of tumour-suppressor proteins such as p16INK4a, p15INK4b, cyclin D3 and hDlg all result in abnormal cell proliferation of infected cells. Indeed, most of the genes and proteins targetted directly or indirectly by Tax are likely to cooperate in promotion of cell proliferation upon infection by HTLV-1. This consideration is consistent with the fact that normal cell regulation is protected by redundant mechanisms and also that tumorigenesis proceeds through accumulation of multiple gene abnormalities.

Fixation of genetic abnormality. In addition to the genes discussed in the previous section, Tax also affects

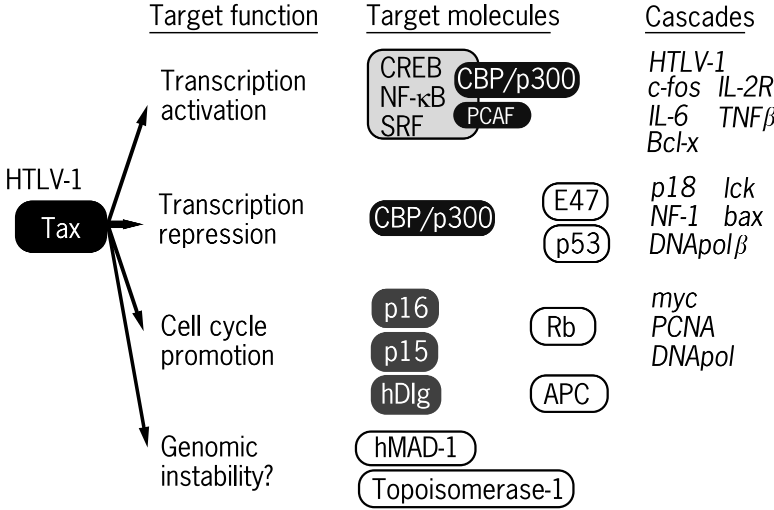


Figure 6 Summary of the pleiotropic function of Tax.

genes which are not directly linked to cell proliferation. These include trans-repression of DNA polymerase-β and Bax, trans-activation of Bcl-X_L and inhibition of topoisomerase I, PCNA and HsMAD. DNA polymerase-β DNA topoisomerase I and PCNA are involved in repair of damaged DNA and therefore reduction of these activities through any mechanism would result in higher mutation rate fixing genetic abnormality in infected cells. In fact, a mutagenic effect of Tax on host cell chromosomes has been directly demonstrated. In relation to DNA damage and its repair, p53 function is also affected by Tax; p53-dependent transcription is trans-repressed through interfering with p53 binding to CBP (see the previous section). Furthermore, p53 function is impaired by phosphorylation at Ser15. Therefore, Tax is able to abolish the cell cycle checkpoint function of p53, thus leading cells to accumulate DNA mutations. This would also be bypassed through Tax binding to HsMAD, a component of the mitotic checkpoint, suggesting promotion of abnormal cell division.

Escape from apoptosis. It is widely accepted that unbalanced activation of some signalling pathway for cell growth or undesired mutation of critical genes would induce apoptosis and eliminate these abnormal cells. Therefore, tumorigenic mutation or viral function would be counteracted were apoptosis induction to be effective. Surprisingly, the Tax protein was shown to trans-activate Bcl-X_L and trans-repress Bax, an inhibitor and mediator of apoptosis, respectively, thus preventing apoptosis of these cells. These effects of Tax seem to be critical for finalizing its effects *in vivo*, otherwise potent activities of Tax would be cancelled by elimination of the Tax-expressing cells by apoptosis.

In total, Tax induces abnormal cell growth, genomic instability and fixation of the abnormality through its pleiotropic functions. Tax is therefore concluded to be a tumour initiator and promoter in the development of ATL.

Low Expression of Tax In Vivo

The pleiotropic functions of Tax thus far identified *in vitro* are all consistent with its implication as an aetiological factor for ATL. However, the extremely low expression of Tax *in vivo* offers a different impression from those expected from its properties *in vitro*. Infected T cells in peripheral blood do not produce significant amounts of Tax mRNA and protein, irrespective of whether they are transformed or not. Expression is detected only by sensitive PCR, but not by any other techniques. Semiquantitative PCR indicated that over 95% of infected cells are absolutely negative for expression of the viral message. Such extremely inefficient expression of the viral genes might be beneficial for the virus to escape from host immune responses, but raises a question concerning the role of Tax in tumorigenesis. It should be emphasized that continuous expression of the viral proteins in infected individuals are suggested by the persistent prevalence of the antibodies.

These somewhat mysterious observations may be explained in several ways: first, Tax plays essential roles in the early stages of transformation, but it is no longer required for maintenance of the transformed phenotype. Second, very low levels of Tax may be sufficient for the maintenance of abnormal phenotypes of infected cells. However, this is unlikely because most infected cells are absolutely negative for Tax expression. Third, Tax is transiently expressed in a small population of infected cells at one time and in another cell population at other times. This would be possible in T cells, some of which are stimulated by antigens or other signals, but its activation would be soon terminated unless the stimulation was continuous. Different specificities of stimulation would induce Tax expression in different populations of T cells. However, it is not easy to distinguish these possibilities experimentally.

The other possibility for low expression of HTLV-1 genes is defective proviruses *in vivo* in cells. This is, however, not the case, since the expression of viral genes is rapidly induced when primary cells from infected individuals are cultured. The mechanism for such restriction of viral expression *in vivo* is not well understood. The extremely low expression of Tax in primary tumour cells is in contrast to transforming genes of DNA tumour viruses, which are significantly expressed and responsible for maintenance of the abnormal growth of transformed cells. For example, *E6/E7* genes of human papillomavirus are expressed in the HeLa cell line maintained for a long period in culture and their repression arrests cell growth.

Clonal Expansion of Infected T cells

Another point in question is the clonal burst of infected T cells. Growth stimulation through the pleiotropic functions of Tax would result in a random population of proliferating cells, since it stimulates growth of most T cells infected with HTLV-1. However, leukaemic cells are always monoclonal. Therefore, an additional genetic event is postulated to trigger the clonal selection of infected T cells. However, not much is known about the mechanism of clonal selection for leukaemic cells. It might not be associated with the viral function. The notion that further alteration is necessary is consistent with the long delay in ATL development after HTLV-1 infection.

Prevention of HTLV-1 Infection

Infection by HTLV-1 is easily detected by antibodies against HTLV-1 proteins. Assay kits for ELISA, Western blotting and particle agglutination have been produced for the diagnosis of HTLV-1 infection and are commercially available. Worldwide screening of HTLV-1 prevalence has been carried out using these systems. Transfusion of sero-positive blood results in transmission of HTLV-1 to two-thirds of the recipients. With the introduction of HTLV-1 screening systems in blood banks, sero-positive blood is now rejected in Japan and viral transmission through transfusion has been greatly reduced.

This has also resulted in an effective reduction of transfusion-related HAM/TSP. The application of these systems to populations in all endemic areas is now clearly shown to prevent HTLV-1 infection.

The major, natural route of viral transmission is from mother to child through infected T cells in breast milk (Hino *et al*., 1985). Curiously, mothers with high levels of antibodies to Tax protein transmit the virus to their offspring at a higher rate than do those with low titres of Tax antibodies. It is possible that efficient replication of HTLV-1 would stimulate antibody production at high levels but that the antibodies might not significantly inhibit viral replication. Non-breast-feeding has been examined among sero-positive mothers in Nagasaki City, Japan, to prevent the viral transmission into children. By consent, pregnant women are surveyed for HTLV-1 antibodies; those who are sero-positive are encouraged to avoid breast feeding. The trial indicated a drastic reduction in the incidence of sero-positive children, from about 30% to just a few per cent. The success of this trial provides direct evidence for viral transmission through milk and suggests the possibility of eliminating ATL in the next few generations.

Epidemiological studies have established that HTLV-1 infects individuals at an early age through breast milk. In fact, within 6–12 months of birth, one-quarter to one-third of children born to sero-positive mothers develop antibodies to HTLV-1. A dramatic increase in age-specific rates of sero-positivity is, however, observed only in those over 40 years old. Sexual transmission of HTLV-1 from husband to wife may account for a portion of this increase, but cannot explain the increase in males. One possible speculation to account for this phenomenon is that the age-dependent prevalence is a reflection of the infection in the young period of 0–1 years old. These epidemiological results suggest that prevention of milk and transfusion-mediated transmission might be adequate.

REFERENCES

Chen, I. S., *et al*. (1983). Human T-cell leukemia virus type II transforms normal human lymphocytes. *Proceedings of the National Academy of Sciences of the USA*, **80**, 7006–7009.

Fujisawa, J., *et al*. (1985). Functional activation of the long terminal repeat of human T-cell leukemia virus type I by a trans-acting factor. *Proceedings of the National Academy of Sciences of the USA*, **82**, 2277–2281.

Hino, S., *et al*. (1985). Mother-to-child transmission of human T-cell leukemia virus type-I. *Japanese Journal of Cancer Research*, **76**, 474–480.

Hinuma, Y., *et al*. (1981). Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proceedings of the National Academy of Sciences of the USA*, **78**, 6476–6480.

Jeang, K. T., *et al*. (1990). HTLV-I trans-activator protein, tax, is a trans-repressor of the human beta-polymerase gene. *Science*, **247**, 1082–1084.

Jin, D. Y., *et al*. (1998). Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. *Cell*, **93**, 81–91.

Kwok, R. P., *et al*. (1994). Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature*, **370**, 223–226.

Miyoshi, I., *et al*. (1981). Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature*, **294**, 770–771.

Nevins, J. R. (1992). E2F: a link between the Rb tumor suppressor protein and viraloncoproteins. *Science*, **258**, 424–429.

Poeisz, B. J., *et al*. (1980). Detection and isolation of the C retrovirus from fresh and cultured lymphocytes of a patient with T cell lymphoma. *Proceedings of the National Academy of Sciences of the USA*, **77**, 7415–7419.

Seiki, M., *et al*. (1983). Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proceedings of the National Academy of Sciences of the USA*, **80**, 3618–3622.

Seiki, M., *et al*. (1984). Nonspecific integration of the HTLV provirus genome into adult T-cell leukaemia cells. *Nature*, **309**, 640–642.

Sodroski, J. G., *et al*. (1984). Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science*, **225**, 381–385.

Suzuki, T., *et al*. (1996). HTLV-1 Tax protein interacts with cyclin-dependent kinase inhibitor p16INK4A and counteracts its inhibitory activity towards CDK4. *EMBO Journal*, **15**, 1607–1614.

Tajima, K. (1990). The 4th nation-wide study of adult T-cell leukemia/lymphoma (ATL) in Japan: estimates of risk of ATL and its geographical and clinical features. The T- and B-cell Malignancy Study Group. *International Journal of Cancer*, **45**, 237–243.

Uchiyama, T., *et al*. (1977). Adult T-cell leukemia: clinical and hematologic features of 16 cases. *Blood*, **50**, 481–492.

Weiss, R., *et al*. (1985). *RNA Tumor Viruses* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Yoshida, M., *et al*. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proceedings of the National Academy of Sciences of the USA*, **79**, 2031–2035.

Yoshida, M. (1995). HTLV-1 oncoprotein Tax deregulates transcription of cellular genes through multiple mechanisms. *Journal of Cancer Research and Clinical Oncology*, **121**, 521–528.

FURTHER READING

Yoshida, M. (2001). Multiple viral strategy of HTLV-1 for dysregulation of cell growth control. *Annual Review of Immunology*, **19**, 475–496.

Chapter 7

Genomic Instability and DNA Repair

Kurt W. Kohn
National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Vilhelm A. Bohr
National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

CONTENTS

- Introduction
- Genomic Instability
- The Need for DNA Repair
- Repair by DNA Alkyltransferase
- Repair of Single-Base Damage
- Repair of Base-Free Sites
- Nucleotide Excision Repair (NER)
- Mitochondrial DNA Repair in Mammalian Cells
- DNA Mismatch Repair (MMR)
- Recombinational Repair: Repair of DNA Double-Strand Breaks (DSB) and Cross-links
- DNA Repair Involving RecQ-Family Helicases
- DNA Repair Polymerases
- The Hus1:Rad1:Rad9 Sliding Clamp and Rad17 Clamp-Loader Model
- Multimolecular Assemblies and Nuclear Foci
- The DNA Replication Checkpoint

INTRODUCTION

Preservation of genome integrity is critical for the functional preservation of dividing cells. Genomic instability and DNA damage set the stage for carcinogenesis, both in the initiation stage and in the evolution toward malignancy. The frequency estimates for various types of DNA damage normally incurred in mammalian cells are summarized in **Table 1**. Multiple mechanisms have evolved to repair DNA damage and preserve genome integrity. Several mechanisms are conserved widely or even universally among organisms and thus play a fundamental role in maintaining the integrity of living species.

Table 1 Estimated frequencies of DNA lesions normally occurring in mammalian cells

Damage	Events per cell per day
Single-strand breaks	55 000
Depurinations	13 000
Depyrimidinations	650
Guanine-O6 methylation	3100
Cytosine deamination	200
Glucose-6-phosphate adduct	3
Thymine glycol	270
Thymidine glycol	70
Hydroxymethyluracil	620
Guanine-8 oxygenation	180
Interstrand cross-link	8
Double-strand break	9
DNA-protein cross-link	Unknown

Multistage carcinogenesis, the stepwise accumulation of genetic changes favouring malignant behaviour, is brought about jointly by chemical mutagenesis and genomic instability mechanisms. Spontaneous mutations, estimated at perhaps three per cell during a human lifetime, however, are insufficient by themselves to account for cancer incidence (Coleman and Tsongalis, 1999). The probability of accumulating the several (perhaps five or more) genetic changes thought necessary for malignancy is greatly enhanced by the development of genomic instability, which can function as a mutator phenotype. For most malignant tumours, therefore, genomic instability plays an important part in causality (see reviews by Breivik and Gaudernack, 1999).

Genomic instability and/or defective DNA repair are characteristic of several human genetic disorders, including ataxia telangiectasia, Fanconi anaemia, Bloom syndrome, Werner syndrome, xeroderma pigmentosum,

Cockayne syndrome and Nijmegen breakage syndrome. Most of these genetic defects predispose affected individuals to the development of malignant tumours (see the chapter *Inherited Predispositions to Cancer*).

GENOMIC INSTABILITY

Genomic instability implies an abnormally high rate of genomic alterations. Not only do tumours contain genome abnormalities, they also have increased genomic heterogeneity among their cells. Two types of genomic instability syndromes have been established in tumours: microsatellite instability and chromosome instability. Malignant tumours almost always have one or the other, but rarely both. Thus malignancy can develop by one route or the other, but does not need both (see review by Coleman and Tsongalis (1999)).

Both types of genomic instability have been observed early in tumour development while the lesions are still small and benign. Genomic abnormalities and variation increase as tumours progress toward malignancy. Genomic variation progresses by a combination of increased rate of variation (genomic instability) and selection of cell clones adapted to the malignant lifestyle (see the chapter *Cell Proliferation in Carcinogenesis*).

Genomic abnormalities can arise by several mechanisms having different regional characteristics in DNA. Local alterations include sequence amplifications, deletions, insertions and point mutations. Chromosomal translocations and rearrangements arise by breakage/rejoining or other recombination events. Mitotic abnormalities give rise to aneuploidy by unequal chromosome segregation and to multiploidy by failure of nuclear division. These processes lead to genomic heterogeneity, which becomes increasingly rampant during the progress of malignancy; the most autonomously replicating genotypes become selected as tumours progress. Thus genomic instability is implicated in both the origin and progression of most malignant tumours.

Microsatellite Instability

Microsatellite instability implies variation in the length of homopolymer regions (particularly poly(A) sequences) or of dinucleotide or trinucleotide repeat regions. Microsatellite instability is caused by a defect in mismatch repair (MMR). For a compilation of microsatellite instability reports for various tumours and a discussion of mechanism of production, see Coleman and Tsongalis (1999).

Microsatellite instability is discussed further in the section on MMR.

Chromosome Instability

Chromosome instability implies not only abnormal variation in gross chromosome number (aneuploidy), but also an increased rate of chromosomal alterations. Aneuploid tumour cells have been estimated to have a 10–100-fold increased rate of chromosome gain or loss and a similar increase in the rate of loss of heterozygosity at specific genomic sites (Lengauer *et al*., 1998). Sometimes the loss of a chromosome is balanced by the duplication of the remaining allelic chromosome. Loss of heterozygosity of a given genetic region can be detected by DNA electrophoresis and hybridization even when the corresponding chromosome is present in the euploid 2 copies.

For a compilation of frequencies of gain or loss of DNA from each chromosome arm in various tumours (as determined by comparative genomic hybridization), see Rooney *et al*. (1999). Interestingly, some chromosome arms show gains much more often than losses, some show the reverse and many show nearly equal frequencies of gain or loss. For a given tumour type, on the other hand, each chromosome arm may show frequent gain or loss, but rarely both. This may reflect distinctive preferred patterns of selection among different tumour types in the context of chromosome instability.

Although the association of microsatellite instability with defective MMR is well established, the molecular origin of chromosome instability is only beginning to be elucidated. Unlike microsatellite instability, which can be complemented by cell fusion with a chromosome-instability cell type, the converse is not the case: whereas microsatellite instability is a recessive character, chromosome instability is dominant and may be due to a gain-of-function mutation of a single protein (Lengauer *et al*., 1998). Chromosome instability is not a consequence of increased chromosome number *per se*, because tetraploid cells resulting from the fusion of two cells having microsatellite instability does not yield chromosome instability. Sometimes gross chromosome instability is accompanied by p53 mutation, but the early stages of chromosome instability often appear much earlier than p53 mutation. Moreover, mutation or inactivation of p53 does not by itself affect chromosome stability (Lengauer *et al*., 1998).

Centrosome Abnormalities

Abnormal centrosome function in cancer cells can cause chromosome instability. Sometimes the centrosomes become overduplicated, causing multifocal spindles and grossly abnormal mitoses. Centrosome dysfunction however is not the only route to aneuploidy. Improper behaviour of the mitotic spindle can be due to a variety of as yet poorly defined defects in mitotic checkpoints, which can cause unequal chromosome segregation between daughter cells.

The molecular basis of centrosome dysfunction is not yet clear. Although abnormal centrosome function is sometimes associated with p53 loss or mutation, this is not a strict association, because normal centrosome function together

with aneuploidy has been observed in the presence of mutated p53. Thus, p53 mutation does not by itself cause centrosome dysfunction, but may be implicated in other routes to aneuploidy.

THE NEED FOR DNA REPAIR

Living organisms are constantly exposed to stress from environmental agents and from endogenous metabolic processes. An important factor is exposure to oxidative reagents or oxidative stress, largely arising as a side effect of mitochondrial energy metabolism. The resulting reactive oxygen species (ROS) attack proteins, lipids and DNA. Since proteins and lipids are readily degraded and resynthesized, the most significant consequence of oxidative stress is thought to be DNA modifications, which can become permanent via the formation of mutations and other types of genomic damage.

Many different types of DNA base changes have been observed following oxidative stress, and these lesions are widely considered as instigators of cancer, development, ageing and neurological disorders (for review, see Wiseman and Halliwell, 1996). The endogenous attack on DNA by ROS generates a low steady-state level of DNA adducts that have been detected in the DNA of human cells (Dizdaroglu, 1991). Over 100 oxidative base modifications in DNA have been detected in human cells (Wiseman and Halliwell, 1996). The best known and most widely studied is 8-hydroxyguanosine (8-oxoG). Oxidative DNA damage accumulates in cancerous tissues and is thought to contribute to carcinogenesis. For example, higher levels of oxidative base damage were observed in lung cancer tissue compared with surrounding normal tissue and a ninefold increase in 8-oxoG, 8-hydroxyadenine and 2,6-diamino-4-hydroxy-5-formamidopyrimidine in DNA in breast cancer tissue compared with normal tissue has been reported (Wiseman and Halliwell, 1996). DNA damage can also occur after direct attack by external or exogenous sources. Radiation from various sources can directly damage bases in DNA. For example, ultraviolet (UV) irradiation from exposure to sunlight creates certain DNA lesions. The main ones are the cyclobutane pyrimidine dimers formed usually between two adjacent thymine bases in DNA and the pyrimidine–6,4–pyrimidine dimer photoproducts. Irradiation from γ-ray sources or X-rays creates many different kinds of lesions in DNA, including base modifications, sites with a loss of base, and breaks in a DNA strand. DNA breaks can be single- or double-stranded. Many food constituents can directly damage DNA. These include carcinogens or chemicals that react directly with DNA or do so after metabolic modification. Some of these agents alkylate DNA bases, some forming bulky adducts. For example, aromatic amines are found in a variety of foods and are known to cause DNA damage and to be highly mutagenic. A number of poisons attack DNA directly. An example is mustard gas or nitrogen mustard which chemically modifies DNA bases and produces cross-links between bases on the same or on opposite DNA strands (Kohn, 1996). Interstrand cross-links cause havoc in the cell by completely blocking the progress of polymerases. DNA repair pathways have evolved to deal with all these lesions in DNA. Some are listed in **Figure 1**, which shows the general pathways of DNA repair, including those that are thought to exist in mitochondria.

Some DNA lesions can be repaired directly back to the original DNA structure, as in the case of the alkyltransferase

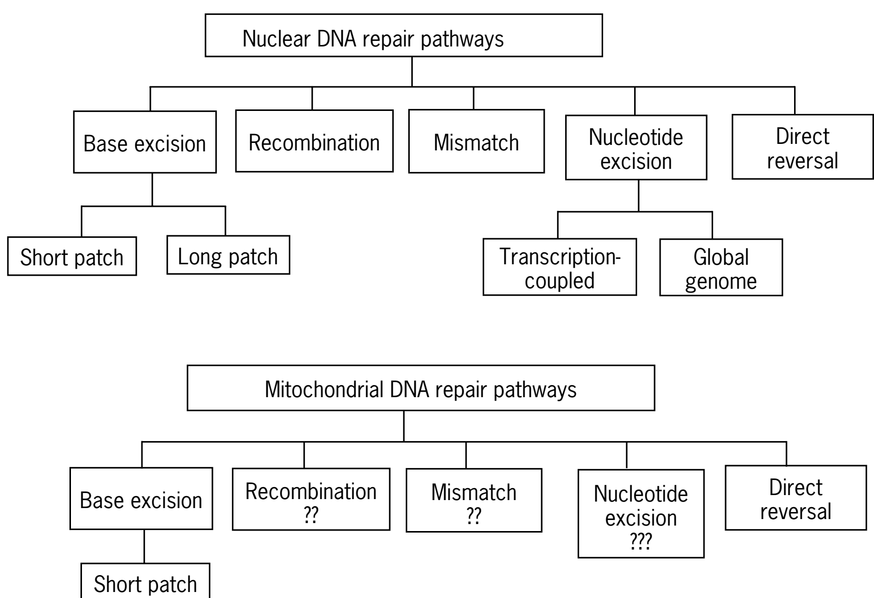


Figure 1 Repair mechanisms for nuclear and mitochondrial DNA in mammalian cells.

and photolyase reactions. Most lesions, however, require more complex repair mechanisms, such as base excision repair, nucleotide excision repair, mismatch repair and recombinational repair. In general, the more bulky DNA base modifications are removed by nucleotide excision repair, while less bulky ones are repaired by base excision repair pathways. These are not firm distinctions, however, since there is much more overlap between these pathways than previously thought. (See the chapters on *The Formation of DNA Adducts* and *Physical Causes of Cancer*.)

REPAIR BY DNA ALKYLTRANSFERASE

Of all DNA repair mechanisms, the simplest and most perfect is mediated by O^6-alkylguanine-DNA alkyltransferase. It is also the most specialized, applying only to a particular type of chemical damage to DNA bases, namely chemical adducts at the guanine-O6 or cytosine-O4 positions. These types of adducts are produced by chemical carcinogens such as nitrosamines and by chemotherapeutic agents such as nitrosoureas (see the chapter *The Formation of DNA Adducts*). Adducts at the guanine-O6 or cytosine-O4 positions alter the base-pairing preferences and therefore are highly mutagenic. Cells normally contain an alkyltransferase, which efficiently removes alkylations from these positions and restores the original chemistry of the base. Thus the repair is perfect in that the DNA is brought back precisely to its original state.

The alkyltransferase that accomplishes this feat is an unusual enzyme, because it is, in part, its own substrate: it transfers an adduct from a DNA base to a sulfhydryl group at the enzyme's active site. The alkylated sulfhydryl group on the enzyme is so stable that it cannot be removed to regenerate the active enzyme molecule: each enzyme molecule can act only once, whereupon it becomes permanently inactivated. The reaction therefore is stoichiometric rather than catalytic, and fails to meet a classical criterion for an enzyme. Nevertheless, in terms of structure and mechanism, this is clearly an enzyme in that it lowers the energy barrier of a reaction. Its unusual feature is that one of the stable products of the reaction happens to be part of the enzyme protein itself. This same feature, however, makes the repair very fast and efficient. It is one of the few DNA repair processes that can be accomplished in essentially a single step. However, the number of DNA adducts that can be removed is limited by the number of active enzyme molecules present in the cell.

REPAIR OF SINGLE-BASE DAMAGE

Base excision repair (BER) is thought to be the major way in which the cell deals with most types of damage to single bases in DNA, although nucleotide excision repair (NER) (see later) may also play a part. An exceptional case, however, is presented by adducts at the O6 position of guanine; as described above, these adducts can be removed by an alkyltransferase while leaving the normal guanine base in place. The alkyltransferase-mediated repair of guanine-O6 adducts is unusual and remarkable in that it is error free. On the other hand, adducts at the guanine-N7 position, which is the most common alkylation site on DNA, lead to loss of the guanine base. This can occur either through the action of a repair glycosylase or by spontaneous hydrolysis of the glycosidic bond (alkylation at N7 facilitates spontaneous release of guanine from DNA). Either way, the loss of the base leaves the DNA with an unsubstituted deoxyribose unit, known as a base-free site or AP site (AP stands for apurinic/apyrimidinic). A particularly important BER mechanism removes uracil residues that normally arise spontaneously due to the occasional hydrolysis of the 4-amino group of cytosine. This is carried out by a uracil glycosylase.

Hydrolysis of 5-methylcytosines yields a normal base, thymine, which would not be recognized as abnormal by a repair glycosylase. 5-Methylcytosine in mammalian DNA occurs only at 5′-CpG-3′ dinucleotides, which consequently are strongly disfavoured in most regions of the genome. Some DNA regions, usually outside of coding sequences, however contain islands rich in CpG sequences which, when methylated, are subject to GC → AT transitions.

REPAIR OF BASE-FREE SITES

Two independent mechanisms exist for repair of base-free (or 'abasic') sites: single-nucleotide gap-filling and long-patch repair. Either process may be preceded by the action of a glycosylase which cleaves the glycosidic bond between the base and sugar moieties of DNA. An AP-endonuclease then cuts the DNA strand containing the base-free site immediately on the 5′ side of the lesion, and yields a 5′-sugar-phosphate terminus and a 3′-OH terminus. A repair polymerase (typically Pol β) then extends the 3′ end and displaces the base-free sugar residue. In the long-patch mechanism, the displacement of the damaged strand extends to include between 2 and about 10 nucleotide residues. The displaced DNA segment (sometimes referred to as a 'flap') is removed by a flap endonuclease (FEN1). The DNA strand can then be made whole by the action a DNA ligase. Long-patch repair may be carried out via Pol β, which does not require proliferating cell nuclear antigen (PCNA), or via Pol δ which is PCNA dependent (discussed by Prasad *et al.*, 2000). These authors recently found in a reconstituted system of long-patch BER that FEN1 and Polβ can cooperate in the linked processes of strand displacement (a Pol β function) and displaced-strand cleavage (a FEN1 function). The two enzymes mutually stimulated each other in this system.

A PCNA-dependent reconstituted long-patch BER system was found to require replication protein A (RPA) for optimum activity; PCNA and RPA seemed to function coordinately (Dianov *et al*., 1999).

It is not yet clear how the two base excision repair pathways (involving short- or long-patch repair of the consequent base-free sites) are regulated. It may depend on the type of glycosylase involved or on the type of DNA polymerase. The BER complex situated at the AP site in DNA involves DNA polymerase, FEN1, AP-endonuclease and PCNA, and the organisation of this complex is very important. In addition to the short- and long-patch BER pathways, there appears to be another BER pathway of transcription coupled repair which deals with repair of active genes. As mentioned above, there is much emerging evidence for overlap and interaction between the repair pathways. For example, the xeroderma group G (XPG) protein participates in both BER and NER.

NUCLEOTIDE EXCISION REPAIR (NER)

NER is the most versatile of the DNA repair mechanisms. It repairs a variety of bulky adducts that distort the DNA helix, but only if both chemical damage and helix distortion are present. One of the most important functions of NER is to repair photoproducts due to sunlight UV exposure. One of these products, pyrimidine–6,4–pyrimidine dimers, (representing about 20% of photoproducts), causes large DNA helix distortions and is efficiently repaired by NER. Another photoproduct, cyclobutane–pyrimidine dimers, is more abundant (about 80% of the total), but causes less helix distortion and therefore is less efficiently repaired. The repair of the latter type of lesion may be aided specifically by the XPE protein, but the exact mechanism is not clear. NER also repairs bulky carcinogen adducts, such as DNA adducts of polycyclic aromatic hydrocarbons, and DNA cross-links produced by anti-cancer drugs such as cisplatin. In addition, NER can function (albeit inefficiently) as a backup for base damage that evades BER or the alkyltransferase repair mechanism (see reviews by Balajee and Bohr (2000), Batty and Wood (2000) and de Boer and Hoeijmakers (2000)).

The proteins associated with the seven complementation groups of xeroderma pigmentosum (XP) (XPA to XPG) play parts in the NER mechanism. XP is a highly cancer-prone disease in which the patients suffer from high incidence of skin and internal cancers, due to defective DNA repair. The NER process is depicted in an interaction diagram in **Figure 2**. The diagram shows how the proteins involved in NER assemble at the site of a DNA lesion and cooperate to excise a DNA single-strand segment containing the lesion.

The global form of NER operates throughout the genome. A modified form of NER, known as transcription-coupled repair (TCR), however, operates preferentially at sites of transcription. TCR differs from global NER in the mechanism by which the DNA helix at the lesion site is opened to permit access to the repair machinery. In global NER, the lesion is recognized by the XPC:HR23B heterodimer which then recruits the transcription factor IIH (TFIIH) complex. TFIIH contains two DNA helicases (XPB and XPD) that unwind the DNA locally in opposite directions from the lesion site. At about the same time, XPA comes into play and serves as a nucleus for the assembly of other repair components at the lesion site. XPA may also participate in the recognition of lesions. The DNA unwinding may be assisted by the single-strand binding protein RPA. In addition, RPA can bind several key proteins, including XPA, and thus may assist in assembling the repair complex.

In the case of TCR, a DNA lesion is detected in the course of transcription by RNA polymerase II and its associated proteins. The polymerase stalls at the site of the lesion, which somehow leads to the assembly of the repair complex. Although the details of how this happens are not clear, the polymerase presumably dissociates and is replaced by TFIIH and XPA; it is thought that the Cockayne syndrome group B (CSB) protein plays a part in this process.

TFIIH has at least two modes of action. First, it functions in transcription initiation through the action of its associated helicases, as well as through the action of cyclinH:cdk7 (another component of TFIIH) which phosphorylates the C-terminal tail of RNA polymerase II and thereby allows transcription to start. (Another function of cyclinH:cdk7 is to phosphorylate cyclin-dependent kinases in the course of cell cycle regulation (Kohn, 1999).) Second, the helicases XPB and XPD (which are also part of the TFIIH complex) play an important part in NER and TCR (the helicases bind more tightly in the TFIIH complex than does cyclinH:cdk7, which may not be present in the repair form of TFIIH).

The subsequent steps of NER or TCR, leading to the excision of an oligonucleotide containing the lesion, are summarized in **Figure 2** and its caption. After excision, the resulting gap in the damaged DNA strand must be filled. This is accomplished by DNA polymerse δ or ε, followed by DNA ligase I. The assembly of the polymerase at the $3'$ terminus of the stand gap first requires PCNA, which clamps on to the DNA and RFC, which loads PCNA on to the $3'$ terminus.

When an NER complex has assembled at a DNA lesion, the complex may send a signal indicating the presence of DNA damage. A linchpin in this communication is p53. The ability of p53 to bind RPA could be part of this communication link.

Transcription-coupled NER

As already mentioned, NER has two pathways: (1) transcription-coupled repair (TCR) and (2) global genome repair (GGR). The TCR pathway repairs lesions in the

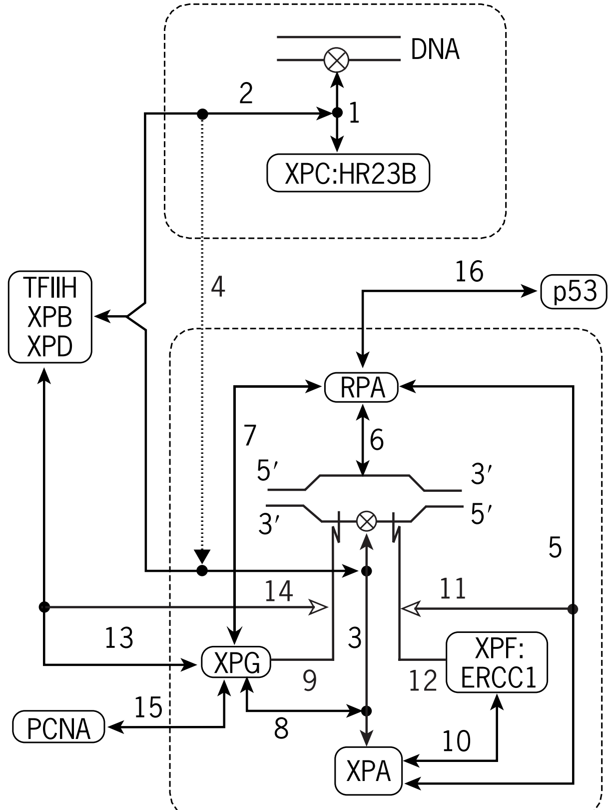


Figure 2 Molecular interactions in nucleotide excision repair (NER). The symbols used in this and subsequent diagrams are defined in (Kohn, 1999) (or see http://discover.nci.nih.gov/kohnk/interaction_maps.html). The upper dashed box depicts a normal double-stranded DNA helix bearning a lesion which is recognized and bound by a heterodimer consisting of XPC and HR23B. A DNA segment surrounding the lesion is then unwound by the XPB and XPD helicases (components of the TFIIH transcription factor complex). The lower dashed box depicts the region of unwound helix and the excision of a DNA segment containing the lesion. The transition from the closed to the unwound state of the DNA (with its associated proteins) is indicated by the hatched arrow with the solid triangle arrowhead. During this transition, the XPC:HR23B complex is replaced by XPA. XPA serves to assemble several proteins that participate in the excision of the lesion. In the case of transcription-coupled repair (TCR), the DNA helix around the lesion is already unwound due to the transcription process, and therefore repair can begin with XPA (lower box) and does not require XPC or HR23B. The numbered steps are as follows. (1) A lesion in one strand of an intact DNA helix becomes bound by a dimer consisting of the XPC and HR23B proteins. (2) This dimer (XPC:HR23B) binds the TFIIH complex which contains the DNA helicases XPB and XPD (Yokoi *et al.*, 2000). All of these proteins are needed for the initial opening of the DNA helix at the site of the lesion. XPB and XPD function with opposite polarity: ($3' \rightarrow 5'$ and $5' \rightarrow 3'$, respectively) to unwind the DNA for a short distance on both sides of the lesion (no diagram symbol is available for unwinding). (3) XPA is then recruited to the lesion (however, in the case of TCR, the helix is already open owing to the presence of RNA polymerase II, and XPA can bind to the lesion without the aid of XPC:HR23B). (4) The transition arrow (solid triangle arrowhead, shown here hatched) indicates that the DNA helix opens and XPC:HR23B is replaced by XPA, and the TFIIH complex now is bound to XPA instead of to XPC:HR23B. (5) XPA binds the DNA single-strand-binding protein RPA. (6) RPA binds the undamaged strand where the helix has been opened. Thus RPA helps to stabilize the XPA complex at the site of the lesion (the length of the unwound region in NER is similar to the 30 nucleotides required for optimum binding of RPA to DNA single strand). (7) RPA recruits endonuclease XPG. (8) XPG binds XPA while XPA is bound to the lesion (XPG may be required for XPA to replace XPC:HR23B at the site of the lesion; this is not shown in the diagram). (9) XPG incises the lesion-containing strand on the 3′ side of the lesion (approximately 6–14 nucleotide residues away from the lesion). (10) XPA recruits the XPF:ERCC1 heterodimer to the lesion site. (11) RPA interacts with XPF and directs the endonuclease activity of XPF to the 5′ side of the lesion. (12) XPF incises the lesion-containing strand on the 5′ side of the lesion (approximately 16–25 nucleotides away from the lesion) (a single-strand segment containing the lesion is thereby released whose modal length is 27 (24–32) nucleotides, independent of the type of lesion. The sites of incision by XPF and XPG may be at or close to the transitions between unwound and helical DNA, since these two enzymes are structure specific endonucleases). (13) TFIIH interacts strongly with XPG. (14) TFIIH may help to position XPG on the 3′ side of the lesion. (15) XPG recruits PCNA which is required for the subsequent DNA repair synthesis that fills the gap left by the excised strand segment. (16) RPA can bind p53 and thereby perhaps serve to signal the presence of DNA damage. For references, see reviews by Balajee and Bohr (2000); Batty and Wood (2000); and de Boer and Hoeijmakers (2000) and the annotation list for DNA repair in Kohn (1999).

transcribed strand of transcriptionally active genes and is dependent on RNA polymerase II (RNA pol II) (Balajee and Bohr, 2000). The GGR pathway removes lesions from genes regardless of whether they are transcriptionally active or inactive.

Cells of a rare genetic disease, Cockayne syndrome (CS), are defective in TCR, but proficient in GGR of UV-induced DNA damage. Affected individuals suffer from postnatal growth failure resulting in cachectic dwarfism, photosensitivity, skeletal abnormalities, mental retardation, progressive neurological degeneration, retinopathy, cataracts and sensorineural hearing loss. Two complementation groups, CS-A and CS-B, have been identified and the corresponding genes have been cloned. The cellular phenotype of CS includes increased sensitivity to a number of DNA-damaging agents including UV radiation, ionizing radiation and hydrogen peroxide (Friedberg, 1996).

A characteristic feature of CS cells is that they do not recover the ability to synthesize RNA efficiently after UV damage; this phenotype is consistent with a defect in TCR. The *CSA* and *CSB* genes have been cloned and their products characterized. The *CSA* gene product is a 44-kDa protein that belongs to the 'WD repeat' family. Members of this protein family are structural and regulatory proteins,

but usually lack enzymatic activity. The *CSB* gene product is a 168-kDa protein that belongs to the SWI/SNF family, which are DNA and RNA helicases with seven conserved sequence motifs. CSB has an acidic amino acid stretch, a glycine-rich region and two putative NLS sequences. CSB is a DNA-stimulated ATPase, but is not able to unwind DNA in a conventional strand displacement assay (Selby and Sancar, 1997).

The precise molecular role of CSB is not clear at present. CSB may facilitate repair of active genes by recruiting DNA repair proteins to actively transcribed regions. *In vitro,* CSB forms a complex with RNA polymerase II, DNA and the RNA transcript in a manner that requires ATP hydrolysis (Tantin *et al*., 1997). This quaternary complex recruits another molecular complex including the TFIIH core subunits p62 and XPB. TFIIH is a complex factor thought to promote local DNA unwinding during transcription initiation by RNA pol II and promoter escape, as well as in NER (Balajee and Bohr, 2000).

It is also possible that CSB indirectly stimulates TCR by facilitating transcription. Members of the SWI/SNF family are involved in regulating transcription, chromatin remodelling and DNA repair, including such actions as disruption of protein–protein and protein–DNA interactions. The *CSB* gene product could have a similar function. In fact, it is still a matter of debate whether CS is due to a primary defect in transcription or DNA repair (Friedberg, 1996). Some evidence suggests that CSB may indirectly stimulate TCR by facilitating the process of transcription (Balajee *et al*., 1997; Selby and Sancar, 1997). Thus, CSB may be a transcription elongation factor and a repair-coupling factor acting at the site of RNA pol II-blocking lesions, and the CS phenotype may arise from deficiencies in both transcription and DNA repair. The biological function of CSB in these different pathways may be mediated by distinct functional domains of the protein.

It is well established that the *CSB* phenotype involves a defect in TCR of UV-induced DNA damage, although

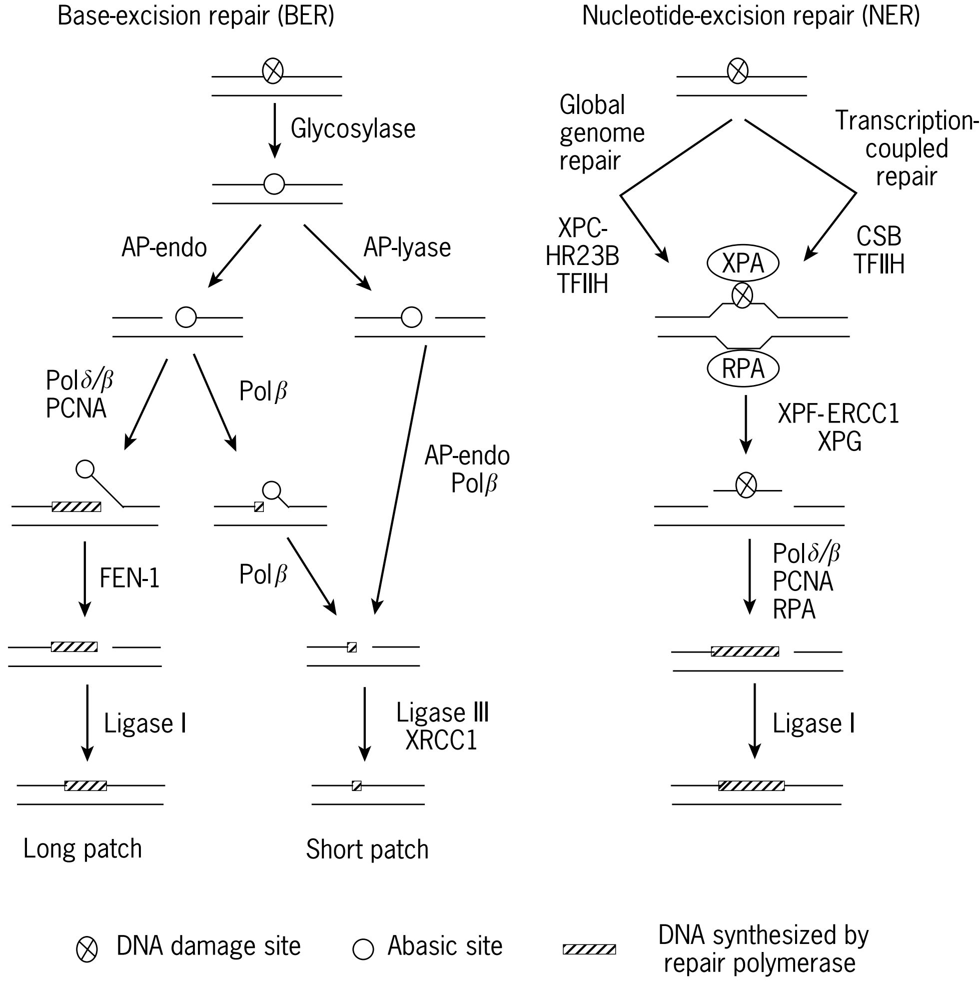


Figure 3 The pathways of base excision repair (BER) and nucleotide excision repair (NER). Some of the proteins involved are shown. The general steps are recognition, incision, replication and ligation. NER has a subpathway called transcription-coupled repair (TCR) and within BER there are two pathways, long-patch and short-patch BER. The individual pathways are described in the text.

CSB may also function in TCR of oxidative damage (Le Page *et al.*, 2000).

The various types of BER and NER are summarized in **Figure 3**.

MITOCHONDRIAL DNA REPAIR IN MAMMALIAN CELLS

Oxidative phosphorylation is an essential metabolic pathway that takes place in the mitochondrion and produces reactive oxygen species (ROS), such as superoxide, hydrogen peroxide, and hydroxyl radicals. Mitochondrial DNA (mtDNA), because of its close proximity to the electron transport chain, is at risk of damage from the ROS produced by oxidative phosphorylation. The most common oxidative DNA lesions in mtDNA are uracil, 8-oxoG and thymine glycol. If unrepaired, uracil can cause GC–AT transition mutations. Thus, it is predicted that insufficient DNA repair capacity in the mitochondrion could lead to mitochondrial dysfunction and degenerative disease including altered energy balance and other pathophysiological states associated with ageing and cancer.

For many years it was thought that mitochondria had no DNA repair capacity, and that this could explain the accumulation of DNA lesions found here with ageing. Recently, however, it has become evident that there is, indeed, efficient DNA repair in these organelles, but that they do not appear to have the same variety of repair pathways that are found in the nuclear DNA. The known or suspected DNA repair mechanisms present in mitochondria are indicated in **Figure 1**. The many question marks at repair pathways indicate lack of concise knowledge; there is renewed interest in the exploration of mitochondrial DNA repair.

There have been reports of recombinational repair and mismatch repair pathways in mitochondria and the general repair pathways in mtDNA were recently reviewed (Croteau *et al.*, 1999). In mtDNA, UV-induced lesions, but not oxidative lesions, are repaired. 8-Oxoguanine is efficiently removed from mtDNA, as it is from actively transcribed genes in the nucleus, and the repair efficiency is similar in all regions of the mitochondrial genome (Croteau *et al.*, 1999).

The mitochondria contain enzymes that participate specifically in BER in the mitochondrion. An early indication of the existence of mitochondrial BER (mtBER) was the isolation of a mammalian mitochondrial endonuclease that recognizes and cleaves AP sites. Later, an *in vitro* reconstituted repair assay was performed using mitochondrial enzymes from *Xenopus laevis* and an abasic site-containing DNA substrate (Pinz and Bogenhagen, 1998). A DNA ligase was also purified from mitochondria that may be related to nuclear DNA ligase III. Recently, it was confirmed that the gene encoding human DNA ligase III produces two forms of the ligase, one nuclear and one mitochondrial (Lakshmipathy and Campbell, 1999). The mitochondrial DNA Pol β possesses a 5′-deoxyribose phosphate lyase activity via β-elimination, suggesting that Pol β may play a role in mitochondrial BER (Longley *et al.*, 1998). PCNA has also been shown to stimulate Pol β-mediated DNA synthesis, suggesting that PCNA may be an auxiliary factor in mitochondrial mtDNA replication and repair.

DNA MISMATCH REPAIR (MMR)

MMR deals with at least two types of replication errors: (1) single base-pair mismatches and (2) insertion or deletion loops that arise by slippage between the template and replicating strands. Slippage tends to occur in sequence-repeat regions, such as segments of poly(A) or of dinucleotide repeats such as $(AC)_n$, where base pairing near the replication point can easily shift position. Repeated sequences sometimes occur incidentally in coding regions where slippage could cause frame-shift mutations.

A marker for defective repair of insertion/deletion loops is microsatellite instability. Since microsatellites generally occur outside of genes, abnormal length variation of these sequence repeats does not usually cause mutation. Microsatellite instability, however, serves as a sensitive indicator of the harder to detect changes brought about by the same mechanism within genes.

The significance of the MMR mutator phenotype in tumours with microsatellite instability is shown by the finding of frame-shift mutations in numerous tumour-suppressor genes (*APC*, *TGF*β*-RII*, *IGF-IIR*, *BAX*, *BRCA1*, *BRCA2*) and some DNA repair genes (*hMSH3*, *hMSH6*, *BLM*) (Buermeyer *et al.*, 1999). These frame shifts usually occur within mononucleotide tracts.

In human cells, MMR is carried out by a choreography of multiprotein complexes made up by hMSH2, hMSH3, hMSH6, hMLH1, hPMS2 and probably hMLH3 (**Figure 4**). Also involved are excision and replication proteins: Pol δ, PCNA, RFC, EXO1 and FEN1, which degrade and replace the error-containing segments of newly synthesized strand. The system must also include a DNA ligase which remains to be identified. How the system distinguishes the newly synthesized strand is unknown, although it may well be through recognition of a strand terminus.

Single base-pair mismatches and insertion/deletion loops are recognized by hMSH2:hMSH6 heterodimers (also known as hMutSα complex) which serve as the starting point for the assembly of a ‘repairosome’ typically containing hMLH1, hPMS2 and the excision and replication machinery proteins. An alternative route is through hMSH2:hMSH3 heterodimers (hMutSβ) which recruits a repairosome via hMLH1 and hMLH3; this complex can repair insertion/deletion loops, but not single base-pair mismatches. Loss of *hMSH2* or *hMLH1* thus abrogates the repair of both types of defects. As expected, cells defective

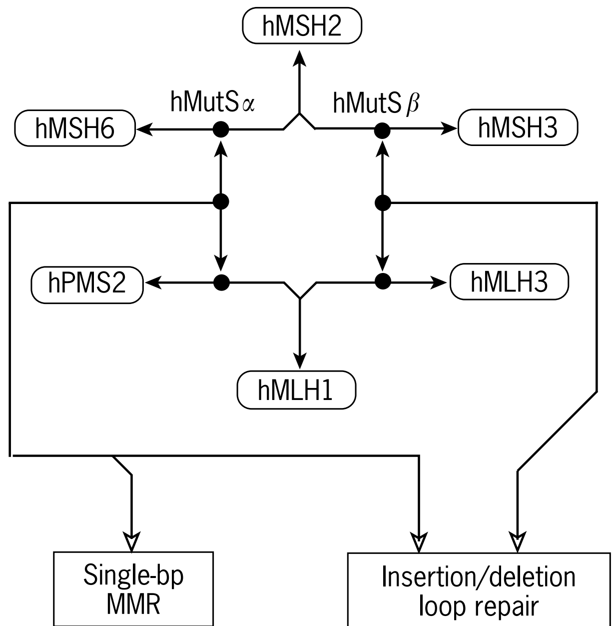


Figure 4 Molecular interactions and functional capabilities of proteins implicated in DNA mismatch repair (MMR). The hMutSα complex, consisting of hMSH2 and hMSH6, combines with hMLH1 and hPMS2 to form the nucleus for a repairosome complex that repairs both insertion/deletion loops and single base-pair mismatches. The hMutSβ complex, consisting of hMSH2 and hMSH3, combines with hMLH1 and hMLH3 and leads to a repairosome complex that can repair insertion/deletion loops, but not single base-pair mismatches.

in *hMSH6* display a mutator phenotype but exhibit little or no microsatellite instability (Jiricny and Nystrom-Lahti, 2000)(because insertion/deletion loops can still be repaired via the hMutSβ path).

It has been proposed that the hMSH2:hMSH6 complex functions akin to a molecular switch, owing to its ADP/ATP-binding ability and ATPase activity (Fishel, 1998). In its ADP-bound state, hMSH2:hMSH6 binds strongly to DNA mismatches. This binding then facilitates the exchange of ADP for ATP (which otherwise occurs only very slowly). In the ATP-bound state, hMSH2:hMSH6 can dissociate from DNA lesions. Its ATPase activity then recycles hMSH2:hMSH6 to the active ADP state. At some point in the cycle, the components of the repairosome are assembled.

MMR, Apoptosis and Chemotherapy Resistance

In addition to its role in DNA repair, the MMR system seems to signal the presence of DNA damage to the apoptosis-initiating system, which may be why MMR-defective tumour cells tend to have increased resistance to DNA-damaging drugs such as cisplatin (see reviews by Li (1999) and Jiricny and Nystrom-Lahti (2000)). Treatment of a mixture of MMR-proficient and deficient cells with cisplatin resulted in enrichment of the MMR-deficient population (Fink *et al*., 1998). Moreover, cisplatin resistance in ovarian cancer recently was reported to be linked to suppression of *hMLH1* due to hypermethylation in the gene's promoter region (Strathdee *et al*., 1999).

MMR-deficient cells also resist killing by alkylating agents that methylate DNA guanine-O6 positions. Such alkylations are mutagenic, because these alkylated guanines base-pair preferentially with T. MMR-deficient cells are alkylation tolerant: they retain the alkylations, but are not killed by them. The cost of survival, however, is mutagenesis. Treatment of *MSH2*-knockout mice with agents that methylate DNA guanine-O6 positions failed to induce apoptosis in the small intestine (a prominent response in wild-type animals) (Toft *et al*., 1999). This *MSH2*-dependent apoptosis was partially mediated by a p53-dependent pathway.

The apoptosis resistance also carries over to other DNA damaging agents including 6-thioguanine (which becomes incorporated into DNA as a bogus base), cisplatin and topoisomerase blockers. Resistance to these agents is conferred by loss of *MSH2, MSH6, MLH1* or *PMS2* functions (but not by loss of *MSH3* function) in several mammalian systems. In addition to loss of apoptotic response, the resistant cells do not exhibit the usual G2/M cell cycle arrest. The components of the MMR system thus appear to have an essential role in the transmission of DNA damage signals (see reviews by Buermeyer *et al*. (1999) and Li (1999)).

The role of MMR in apoptosis signalling may have relevance for chemotherapy with DNA-damaging agents, because drug resistance may develop by loss of MMR function in a single selection step (Aebi *et al*., 1996). Loss of MMR may also confer resistance to low doses of ionizing radiation (Fritzell *et al*., 1997; DeWeese *et al*., 1998) (see review by Li (1999)).

The route by which signals from the MMR system induce apoptosis remains to be elucidated; it may in part involve phosphorylation of p53 and/or the related p73 family proteins (Duckett *et al*., 1999; Li, 1999). The function of p73 in the induction of apoptosis in cisplatin-treated cells may be regulated by tyrosine kinase c-Abl (Gong *et al*., 1999). Since MMR is targeted exclusively to newly synthesized DNA strands (or to strand regions containing nearby strand breaks (Duckett *et al*., 1999)), base damage in the template strand could not be removed: the MMR system could sense the mismatch caused by the base damage, but would attempt to repair the wrong strand. This futile repair cycle is one model proposed as the initiator of the apoptosis signal. Alternatively, the MMR recognition complex might assemble at damage-induced mismatches near replication forks, block replication and thereby induce apoptosis (Li, 1999).

Thus the MMR system corrects DNA mismatches caused by base damage in newly synthesized DNA strands (or in

strands near break sites). However, when presented with damage that it cannot repair, the system sends out an apoptosis-inducing signal. Loss of components of the MMR system allows cells to survive and proliferate while retaining an accumulation of DNA damage. Treatment of MMR-defective tumours with drugs that alkylate DNA at guanine-O6 positions may therefore be ineffective or even detrimental (Li, 1999).

MMR and Colon Cancer

For reasons unknown, MMR defects are associated mainly with cancer of the colon (predominantly right colon), endometrium and ovary. MMR is most closely associated with HNPCC, the most common cancer predisposition syndrome; 70% of HNPCC kindred have germ-line mutations in one of the MMR-associated genes. About 60% of the mutations are in *hMLH1* and about 35% in *hMSH2* (Jiricny and Nystrom-Lahti, 2000; http://www.nfdht.nl). Tumours of HNPCC kindred with mutations in *hMSH2* or *hMLH1* have strong mutator phenotypes and high microsatellite instability (Buermeyer *et al*., 1999). One copy of the gene is mutated in the germline of HNPCC patients, and both copies are mutated in their tumours (which do not exhibit loss of heterozygosity). Loss of MMR requires inactivation of both copies of one of the critical genes, and occurs with increased probability if one of the copies is already mutated in the germ line. In contrast to other cancers, which usually have rampant aneuploidy and loss of chromosome arms, HNPCC tumours have few allelic losses and often have a diploid karyotype (Rosen, 1997). Although chromosome instability and aneuploidy are early events in the development of most other cancers, the initial tumorigenic event in HNPCC is loss of mismatch repair, which induces more localized genome alterations.

Microsatellite instability is also present in 15% of colon cancer patients who have no family history of colon cancer. As in HNPCC, these sporadic tumours occur predominantly on the right side of the colon. In most of these cases, MMR is inactivated, not by mutations, but by transcriptional silencing of the *hMLH1* gene. The *hMLH1* gene is subject to silencing by hypermethylation of its promoter, which is a primary factor in sporadic gastrointestinal tumours having microsatellite instability. In cell lines derived from such tumours, this methylation can be reversed by treating the cells with 5-azacytidine, which eventually restores hMLH1 expression. In most of the tumours, both copies of *hMLH1* are distinguishably suppressed by hypermethylation, although occasionally one of the alleles is inactivated by mutation. For further details and references, see Buermeyer *et al*. (1999) and Markowitz (2000).

MMR and TGFβ Receptor

TGFβ-RII, a receptor in the TGFβ tumour-suppressor pathway, has a special relationship with MMR-deficient colon and stomach cancers. This relationship is due to two circumstances. First, the coding sequence of the human *TGFβ-RII* gene contains a homopolymer tract of 10 adenines that is subject to frame-shift mutation in MMR-deficient cells. This causes premature transcript termination with loss of most of the cytoplasmic domain of the receptor. Among colon tumours, these *TGFβ-RII* frame shifts are found exclusively and almost universally in those tumours that exhibit microsatellite instability. Second, TGFβ functions in intestinal crypts to cause cell cycle arrest and apoptosis when these cells reach the luminal region of the crypts, which is where TGFβ is concentrated. If signalling through the TGFβ pathway is abrogated, crypt cells can continue to proliferate and can initiate the sequence of changes that eventually leads to malignancy. By contrast, MMR-deficient endometrial cancers have a much lower frequency of *TGFβ-RII* frame-shift mutations than do the MMR-deficient gastrointestinal cancers (reviewed by Markowitz (2000)).

The significance of the TGFβ pathway in colon cancer is confirmed by the occurrence, in 15% of microsatellite-stable colon cancers, of inactivating mutations of the *TGFβ-RII* gene outside the poly(A) tracts. Some microsatellite-stable colon cancers with normal *TGFβ-RII* genes bear mutations in downstream components of the TGFβ pathway, such as Smad2 and Smad4. For a recent review, see Markowitz (2000).

A molecular interaction diagram of the essentials of the TGFβ pathway leading to cell proliferation is shown in **Figure 5** and explained in the caption.

Frame-shift mutation in the poly(A) tract of the *TGFβ-RII* gene is often the initiating lesion in microsatellite instability-high non-familial colon cancers. This was shown by a tight correlation between the adenoma-to-carcinoma transition and poly(A) tract length alteration in retrospective tissue pathology samples (Grady *et al*., 1998). A more recent study suggested that the initial gene alteration could be in either the *TGFβ-RII* or the *BAX* gene, or in both (Calin *et al*., 2000). Since BAX is a proapoptotic factor, mutation of this gene may contribute to the apoptosis resistance of some microsatellite instability-high tumours. Nevertheless, TGFβ itself has a proapoptotic effect (pathway unknown) on normal intestinal epithelial cells in culture (references cited by Grady *et al*. (1998)). (See chapter *Signalling by TGF-β*.)

RECOMBINATIONAL REPAIR: REPAIR OF DNA DOUBLE-STRAND BREAKS (DSB) AND CROSS-LINKS

DSB constitute a common type of DNA damage, produced by ionizing radiation, replication blocks and certain DNA-reactive drugs such as bleomycin. The formation and repair of DSB are also part of the immune system's

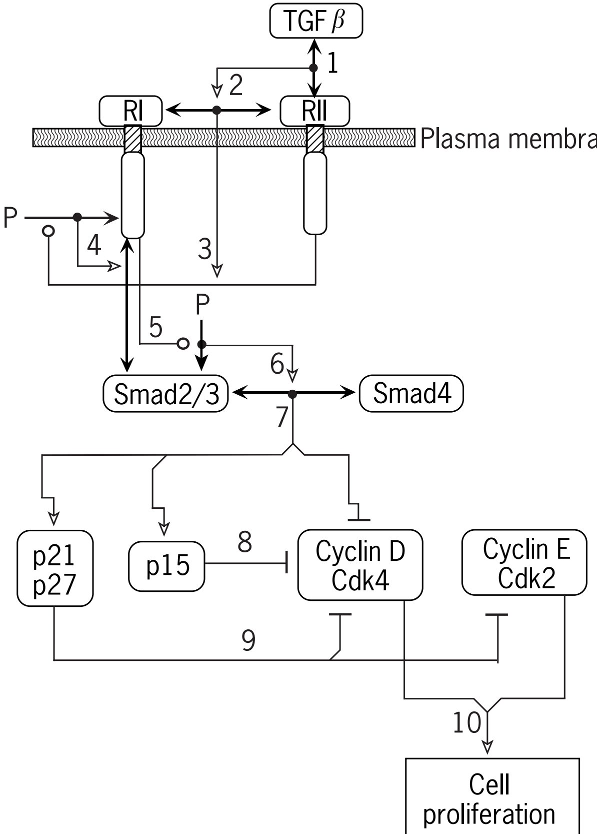


Figure 5 Molecular interaction diagram of the essentials of TGFβ signalling to cell proliferation. (1) TGFβ binds to the type II receptor, TGFβ-RII. (2) This stimulates heterodimer formation between TGFβ-RII and TGFβ-RI. (3) The RII subunit can then phosphorylate the cytoplasmic domain of RI. (4) A phosphorylated site on RI binds Smad2, thereby recruiting this protein to the plasma membrane. (5) This permits the kinase domain of RI to phosphorylate Smad2. (6) Phosphorylated Smad2 binds Smad4. (7) The Smad2:Smad4 heterodimer translocates to the nucleus, where it stimulates the expression of G1/S phase inhibitors p15, p21 and p27, and inhibits the G1/S phase stimulators cyclin D and cdk4 (Massague, 1998; Massague and Wotton, 2000). (8) p15 inhibits Cdk4. (9) p21 and p27 inhibit both Cdk4 and Cdk2 kinase activities. (10) Both cyclin D- and cyclin E-dependent kinases are required for entry of cells into S phase. (For further details, see the chapter *Signalling by TGF-β*.)

V(D)J recombination process. Defects in DSB repair can cause translocations and other DNA rearrangements (Flores-Rozas and Kolodner, 2000). There are two types of recombinational repair of DSB, differing in whether or not the DNA ends to be joined require extensive sequence homology.

Non-homologous End Joining (NHEJ): Repair of DNA Double-strand Breaks

The NHEJ mechanism repairs DNA double-strand breaks without the need for extensive sequence homology between the DNA ends to be joined, although a few complementary base pairs are needed to provide cohesive ends. NHEJ is an error-prone repair process, because it usually creates small deletions. NHEJ is responsible for the rejoining of DSB during V(D)J recombination in the processing of immunoglobulin genes. Defects in NHEJ in mice cause severe combined immune deficiency and radiation sensitivity.

Small deletions are a by-product of NHEJ, because of the need for cohesive DNA ends ('microhomology'), generated by resecting a few nucleotides from one of the DNA strands. The required microhomology may occasionally be as short as a single base pair, although two or three is more efficient. This is in marked contrast to the extensive homology needed by the 'single-strand annealing' repair mechanism, which produces large deletions (see below). NHEJ usually requires resection of only a short region to reach a sequence of microhomology by chance. The exonuclease that resects one of the DNA strands and stops when it detects microhomology is thought to be Mre11, a component of the Mre11:Rad50:Nbs1 module (see the heading The DNA Replication Checkpoint) (Paull and Gellert, 2000).

NHEJ proceeds with the binding of Ku (Ku70–Ku80 dimer) to the ends of the broken DNA (the physical relationship between Ku and Mre11 is unknown). The DNA-bound Ku protein recruits a tight dimer consisting of XRCC4 and Lig4. The DNA ligase activity of Lig4 is thereby activated and efficiently seals the broken strands (**Figure 6**). Ku can also recruit and activate the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs) to the break site. DNA-PKcs, however, is not as essential as Ku, and its exact role in NHEJ is not clear. DNA-PK may phosphorylate Ku70, or RPA (a single-strand binding heterotrimer that may play a role in the process) or DNA-PK itself (see Nick McElhinny *et al.* (2000) and a review by Karran (2000)).

Homologous Recombination, Another Mechanism of DSB Repair

In mammalian cells, homologous recombination was until recently thought to be much less common than NHEJ. Indeed, transfected DNA usually integrates nonhomologously. However, it now appears that homologous recombination does play a major role in DSB repair (see review by Jasin (2000)).

Repair by homologous recombination is not subject to the loss of nucleotides at the joining site that makes NHEJ error-prone. However, when double-strand breaks occur

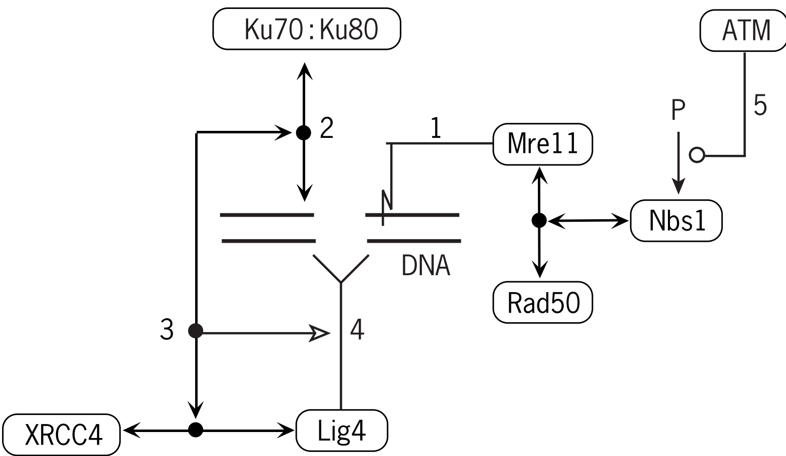


Figure 6 The NHEJ mechanism of DNA double-strand break repair, as currently understood. (1) One DNA strand is resected a short distance by the exonuclease Mre11 (probably functioning as a Mre11:Rad50:Nbs1 trimer) until a point of microhomology is reached, which provides cohesive ends. (2) The double-strand ends bind Ku70:Ku80 dimer. (3) The DNA-bound Ku70:Ku80 recruits Lig4:XRCC4, a tight dimer. (4) Lig4 is thereby activated and brought to the site where it can ligate the strands. Other molecules, such as DNA-PKcs and RPA, may have roles in this process, but their functions remain to be defined. (5) Nbs1 is phosphorylated by ATM in response to DNA damage, as a result of which DNA replication is inhibited (mechanism unknown).

within repetitive sequences (such as Alu-family elements, of which there are about 10^6 copies scattered in the human genome), recombination can occur between Alu elements in different parts of a chromosome or between different chromosomes. This can cause deletion or expansion of regions within a chromosome or translocation between chromosomes. Nevertheless, the preference for recombination between aligned sister chromatids presumably minimizes major chromosome aberrations.

When recombination occurs between homologous chromosomes one of which contains a defective critical gene, the normal copy may be lost. The resulting loss of heterozygosity often can be readily detected. A classical example is the retinoblastoma gene, *Rb*, a recessive tumour suppressor. Heterozygous carriers of a mutant *Rb* gene are susceptible to loss of the functional copy of the gene in an occasional cell, an event that starts the cell on the road to malignancy.

Repair of double-strand breaks or cross-links by homologous recombination requires complementary sequences between the damaged DNA and an undamaged homologue, such as a sister chromatid. First, the 5′-terminating strands of the double-strand break are resected by an exonuclease, so as to leave a 3′-terminated protruding single strand. Repair of DNA cross-links is thought also to begin with processing of the lesion to yield a 3′-terminated protruding strand, but the details of how this is accomplished are unclear.

The protruding 3′-terminus then binds Rad52, a large heptameric doughnut-shaped protein that protects the strand from degradation. Seven Rad52 molecules, assembled as a symmetrical ring, bind specifically to DNA single-strand ends and prevent further exonuclease attack. Although the Rad52 ring has a large central hole, there is no evidence of DNA within the channel. The length of single-strand tail associated with the Rad52 ring is estimated as 36 nucleotides. This terminal region may be exposed and configured to facilitate base pairing with a complementary strand (Parsons *et al.*, 2000).

The single-stranded region behind the Rad52 ring is then covered by a contiguous array of Rad51 molecules to form a nucleoprotein filament, which is capable of strand exchange. Rad51 is a structural and functional homologue, of bacterial RecA that is conserved from yeast to humans. Unlike RecA, however, the eukaryotic Rad51 requires ATP hydrolysis to bind properly on DNA. One Rad51 monomer binds per three nucleotides of DNA single strand. Also, unlike RecA, Rad51 requires a 3′ or 5′ extension of DNA single strand to initiate strand exchange. (see review by Karran (2000) and references cited by Namsaraev and Berg (2000)).

The Rad51 nucleoprotein filament is a loose helix which facilitates invasion of the damaged donor strand into a homologous double-strand region that may be located on another chromatid. The strands of the undamaged recipient DNA become locally separated, while base-paired heteroduplex forms with the donor single strand from the damaged DNA. The pairing between donor and recipient strand (‘heteroduplex’ region) is extended by a process called ‘branch migration’ in which the recipient double strand opens, and its original base pairs are replaced by heteroduplex. DNA replication machinery assembles at the 3′-terminus of the donor strand and extends the strand while further displacing the original complementary strand.

Finally, the displaced strand is cleaved by an endonuclease and the donor strand is ligated to the 5′-terminus of the recipient strand.

The second strand of the damaged DNA may be processed by way of a second recombination event. Alternatively, a new complementary strand is synthesized from the template provided by the displaced recipient strand. A distinction between the two mechanisms is that the former involves only leading-strand synthesis, whereas the latter involves both leading- and lagging-strand syntheses.

The rad51–DNA complex may include additional components that provide, as yet undefined functions, conceivably involved in the processing of chromatin structures. Five sequence relatives ('paralogues') of Rad51 have been demonstrated to engage in a pattern of mutual interactions: Rad51B, Rad51C, Rad51D, XRCC2 and XRCC3. As will be described later, these molecules may function together as multiprotein complexes (Schild *et al*., 2000). Knockout of any of the components conferred a high degree of genomic instability and enhanced sensitivity to the DNA cross-linking agents mitomycin and cisplatin (Schild *et al*., 2000). Tumours that might have acquired genomic instability by loss of function of one of these paralogues would be predicted to be sensitive to DNA cross-linking and double-strand break-inducing agents.

The Rad51 System for Homologous Recombination Repair

As already mentioned, Rad51 and its relatives (Rad51B, Rad51C, Rad51D, XRCC2 and XRCC3) engage in a pattern of mutual interactions and may function together as multiprotein complexes (Schild *et al*., 2000). The demonstrated pattern of interactions is summarized in **Figure 7**.

In cells subjected to ionizing radiation or mitomycin, Rad51 aggregates in foci in the nucleus, made visible by immunostaining. The Rad51 foci occur at DNA damage sites in S phase cells, particularly in regions of postreplicative chromatin. This preference conforms with the preferential double-strand break repair when cells are in late S phase or G_2. The foci may represent repair assemblies that function when sister chromatids are available for homologous recombination (Karran, 2000).

The formation of these foci requires Rad51B (Takata *et al*., 2000), consistent with the ability of Rad51B to enhance interactions between members of the Rad51 complex (**Figure 7**). Another family member, Rad54, is also required (Karran, 2000). Rad54 binds Rad51 (Tan *et al*. (1999), cited by Karran (2000)) and promotes DNA double-strand break repair carried out by homologous recombination with sister chromatids (Dronkert *et al*., 2000).

More extended complexes are possible, since Rad51, directly or indirectly, can bind Rad52, Rad54, p53, BRCA1, BRCA2 and c-Abl. Rad52 can bind the single-strand binding protein RPA. Larger assemblies are possible if other protein–protein interactions can exist simultaneously. For example, p53 has been reported to have binding sites (in order from N- to C-terminus) for MDM2, DP1, PARP, c-Abl, RPA, XPB/D, p19ARF, p300/CBP, BRCA1 and 14-3-3. These, and potential chains of further binding interactions, have been summarized in a molecular interaction map (Kohn, 1999) (see also http://discover.nci.nih.gov/

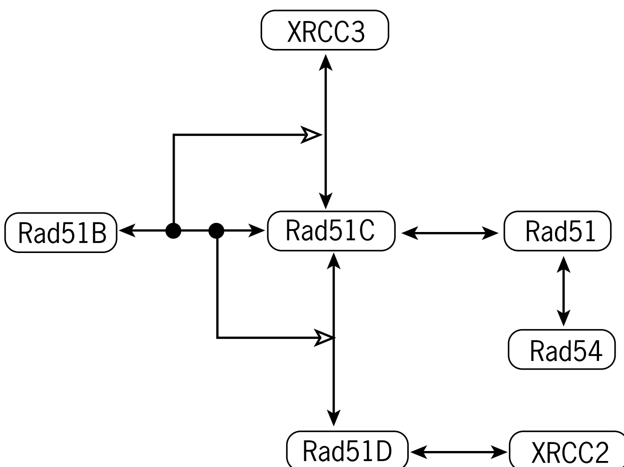


Figure 7 Interaction pattern among Rad51 family members, as reported by Schild *et al*. (2000) and Tan *et al*. (1999) (cited by Karran (2000)). The double-arrowed lines indicate demonstrated binary interactions. The interactions seem capable of building up multimolecular assemblies, as suggested by binding experiments using three components and by yeast three-hybrid experiments. The lines with open triangular arrowheads represent the finding that Rad51B enhances the Rad51C interactions with Rad51D or XRCC3. An additional finding (not included in this diagram) was that Rad51B could bind two molecules of Rad51C in the yeast three-hybrid system, suggesting that the complex could in effect become doubled (Schild *et al*., 2000).

kohnk/interaction_maps.html). Large multimolecular assemblies consisting of different subsets of binding interactions may be formed by remodelling for different functions. The interaction set includes cycles (e.g. Rad51-p53-BRCA1-Rad51 or p53-cAbl-Rad51-BRCA1-p53) which might form a network of molecular chains and perhaps account for the observably large nuclear foci within which DNA repair sites seem to be localized.

Single-strand Annealing

An alternative mechanism of homologous recombination repair, called single-strand annealing, is important when no sister chromatid is available. This mechanism involves resection or realignment of the broken DNA duplex to regions of homology that may be some distance from the break site. Completion of the repair then only requires DNA strand scission of ligation events. Consequently, the segment of DNA intervening between the realigned regions of homology is deleted, which makes this repair mechanism highly error prone.

DNA REPAIR INVOLVING RecQ-FAMILY HELICASES

The RecQ DNA-helicase of *Escherichia coli* is implicated in the suppression of illegitimate recombination and the repair of DNA double-strand breaks. Five human helicases homologous to RecQ are known: BLM, WRN, RecQ4 (also called RecQL4), RecQ5 and RecQ1 (**Figure 8**). Defects in the first three have been recognized as causes of rare genetic diseases: Bloom syndrome, Werner syndrome and Rothmund–Thomson syndrome, respectively (see the chapter *Inherited Predispositions to Cancer*). All three are associated with genetic instability which appears to be due to loss of a helicase activity. The function of these helicases in mammals is only beginning to be elucidated.

The DNA Helicase Defective in Bloom Syndrome (BLM)

The BLM helicase has been the most extensively studied. It is a large protein (1417 amino acids) which contains the motifs characteristic of DNA and RNA helicases. Bloom syndrome cells exhibit a high frequency of chromosome breaks and exchanges. A characteristic of the BLM defect, not shared by the other helicase defects, however, is an increase in reciprocal exchanges between sister chromatids. The cells have a prolonged S phase, thought to be due to difficulty in dealing with stalled replication forks or abnormal replication fork configurations which may occasionally form in the normal course of events. BLM can unwind relatively short DNA duplexes in an ATP-dependent reaction. It can unwind longer helices when aided by the single-strand binding protein, RPA, to which it binds tightly (via the 70-kDa subunit of the RPA trimer) (Brosh *et al.*, 2000). BLM, however, does not efficiently unwind DNA from a blunt end, suggesting that a single-stranded tail is needed (to which RPA might anchor) in order to initiate unwinding of relatively long duplexes. RPA binding and cooperative function is a characteristic

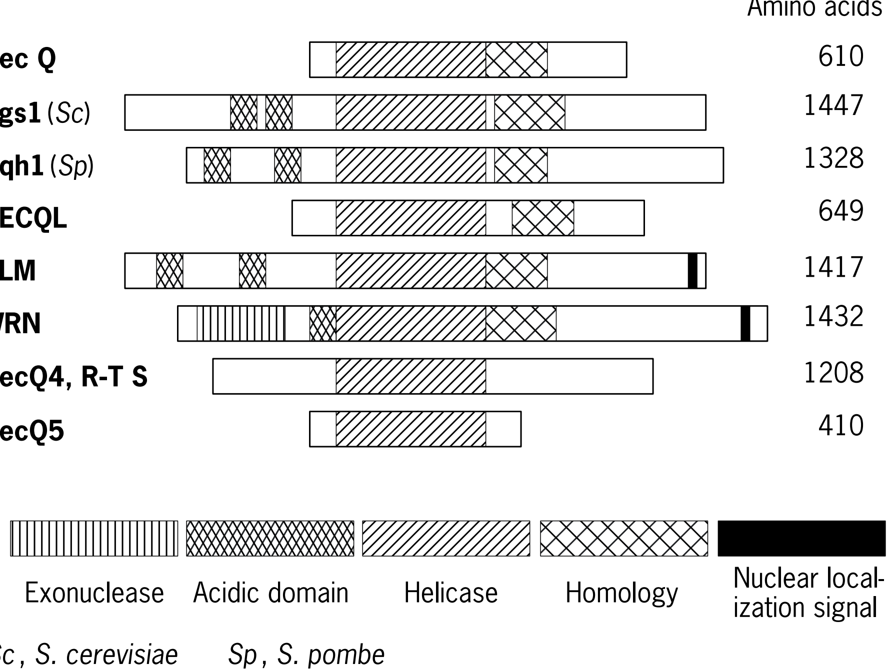


Figure 8 The RecQ family of helicases.

also of the WRN helicase. BLM preferentially unwinds four-stranded DNA helices consisting of Gs stabilized by Hoogsteen pairing; such structures sometimes occur in regulatory regions of genes, and related structures occur in telomeres (Brosh *et al.* (2000) and references cited therein).

BLM recognizes and binds to Holliday junctions (crossover structures between two double helices) and promotes branch migration (movement of the crossover junction along the helices). A possible configuration of a stalled replication fork may form by rewinding of the template strands with displacement of the two newly replicated strands which then pair with each other. The consequent structure is in fact a Holliday junction and could be acted upon by nucleases that normally 'resolve' Holliday junctions. The result would be a double-strand break (Karow *et al.*, 2000) (**Figure 9**). The duplex formed by the pairing of the newly replicated strands could be recognized as a double-strand end by the homologous recombination repair system (e.g. involving Rad52 and Rad51) and lead to homologous recombination between sister chromatids. A role of the Rad52–Rad51–Rad54 repair system is supported by the recent finding that the enhanced sister chromatid exchange in *BLM−/−* cells requires Rad54 (Wang *et al.*, 2000).

BLM is concentrated in part in nuclear foci and, during S phase, also in nucleoli. Nucleoli contain highly repetitive ribosomal DNA sequences which perhaps tend to form homologous crossovers during replication. BLM perhaps helps to prevent or reverse such crossovers, thereby reducing the chance of loss or expansion of ribosomal DNA regions. BLM has been found in a large multimolecular protein contain BRCA1 and several DNA repair-related proteins (Wang *et al.*, 2000).

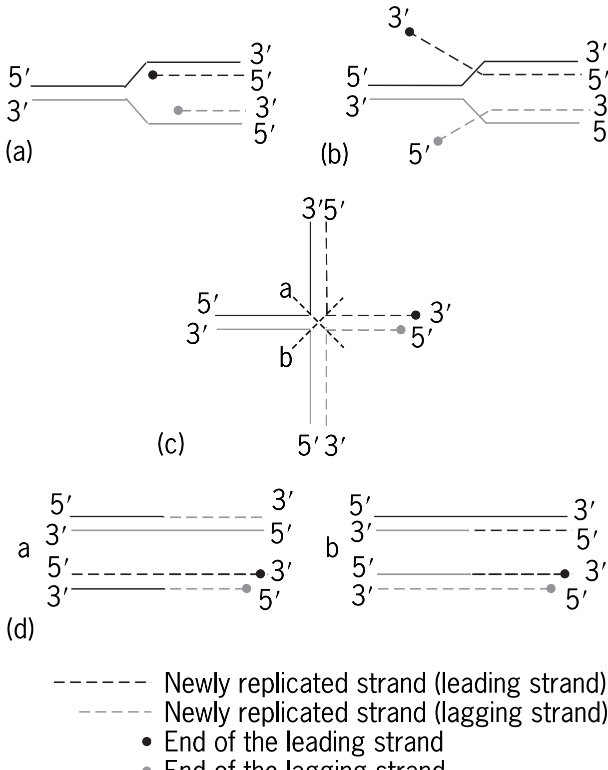


Figure 9 How a helicase defect can lead to a double-strand break involving a parental DNA strand in the vicinity of a stalled replication fork. (a) A normal replication fork. The newly replicated strands are shown by dashed lines with filled circles depicting strand termini. (b) After rewinding of the template helix. This may happen when there is a defect in a helicase that normally operates to unwind the template. The replication fork has backed up, displacing parts of the newly replicated strands. (c) After the displaced segments of newly replicated strands have base-paired with each other. The strands are shown re-configured in order to display the symmetry of a Holliday junction. The Holliday junction can be enzymatically 'resolved' by breakage and rejoining of strands. This can happen in two ways, indicated by the dashed diagonal lines labelled a and b which intersect the strands that are cut and recombined. (d) The final products after the Holliday junction has been resolved in the two possible ways, a and b. In both cases, one of the parental DNA strands becomes joined to newly replicated DNA ending in double-strand termini, equivalent to a double-strand break.

The Protein Defective in Werner Syndrome (WRN)

Werner syndrome (WS) is a homozygous recessive disease characterized by early onset of normal ageing including wrinkling of skin, greying of hair, cataracts, diabetes and osteoporosis. Neoplasms, particularly sarcomas, are observed at higher prevalence in WS patients than in normal individuals of the same age. The symptoms of WS begin to appear near puberty, and most patients die before reaching age 50. Because the clinical features of WS are similar to symptoms of ageing in normal individuals of more advanced age, WS is considered to be a segmental progeria. One of the motivations to study WS is its resemblance to ageing; thus, knowledge of the mechanism and molecular basis of WS might give insight into normal ageing and ageing-associated diseases such as cancer.

WS is caused by mutation in a single gene, *WRN*. The protein product of *WRN* is a member of the RecQ family of helicases, which also includes the yeast protein Sgs1 and the Bloom syndrome protein (BLM). Biochemical evidence shows that the WRN protein is an NTP-dependent DNA unwinding enzyme (and a DNA-dependent NTPase). It was recently shown that WRN is also a 3′–5′ exonuclease (reviewed in Bohr *et al.*, 2000). Thus, WRN is a helicase, exonuclease and ATPase. WRN is the only member of the RecQ protein family that is a DNA exonuclease.

WRN and Genomic Instability

Despite extensive study of cells from WS patients, the precise molecular deficiencies involved in WS remain to be defined. Genomic instability of WS cells has been well documented and is consistent with a defect in replication, recombination or DNA repair. Some WS cells undergo premature replicative senescence and delayed progression of S phase, and some WS cells are hyper-recombinogenic. The *WRN* homologues *S. cerevisiae Sgs1*, *S. pombe rqh1* and *E. coli recQ* suppress illegitimate recombination. One possibility is that the WRN helicase is an anti-recombinase, but this hypothesis has not been tested directly.

It has been suggested that WS cells are genetically unstable because alternate DNA structures are not properly resolved. One such alternate structure is the DNA triplex, which can form in a DNA sequence-dependent manner. Sequences that can form triple helices are abundant in the human genome. WRN helicase unwinds a 3′ tailed triple helix DNA substrate in an NTP-dependent manner. Thus, it is possible that triplex structures persist in WS cells and contribute to the variegated translocation mosaicism observed in WS cells.

WRN and DNA Repair

Some evidence suggests a role for the WRN protein in DNA repair. For example, WS cells are sensitive to the carcinogen 4-nitroquinoline 1-oxide (4-NQO). However, WS cells are not hypersensitive to UV light or several other DNA-damaging agents. WS cells may also be partially defective in transcription. Several observations suggest that WRN may be a general activator of transcription by RNA polymerase II (Balajee *et al*., 1999).

WRN can proficiently unwind short DNA duplexes ($\sim$30 bp) in a reaction dependent on nucleoside triphosphate hydrolysis. In the presence of RPA, however, WRN can unwind long DNA duplexes (up to 800 bp). Moreover, WRN can bind directly to RPA (Brosh *et al*., 1999).

WRN does not preferentially bind DNA damage. It could, however, play an important role as a molecular sensor of DNA damage. It may interact with the DNA minor groove in its action as a helicase, as suggested by studies using the minor groove-binding drugs netropsin and distamycin (Brosh *et al*., 2000a,b).

The progress of the 3′–5′ exonuclease reaction catalysed by WRN is blocked by bulky DNA adducts and oxidative lesions, such as 8-oxoguanine (Cooper *et al*., 2000; Machwe *et al*., 2000). Although WRN does not bind preferentially to DNA lesions, it may sense their presence in DNA via protein–protein interactions, perhaps as an early step in damage recognition, and may recruit DNA repair enzymes to the site of a lesion.

WRN can bind to a single-stranded site in a recombination intermediate Holliday structure, where also the RuvA protein binds. This unwinding is ATP dependent, suggesting that it is due to the helicase activity of WRN. WS cells are not deficient in *in vitro* DNA repair assays for nucleotide excision repair (NER) of various bulky adducts in DNA or in base excision repair (BER) of abasic sites, and they are not generally sensitive to tested DNA-damaging agents, with the exception of 4NQO. Thus, as mentioned above, the influence of WRN on DNA repair may be subtle.

Interaction of WRN with Ku, a Protein Needed for DNA Double-strand Break Repair

WRN interacts physically and functionally with the Ku heterodimer. Ku strongly stimulates the 3′–5′ exonuclease activity of WRN but does not affect its helicase or ATPase activities (Cooper *et al*., 2000). The Ku heterodimer and DNA-PK are key proteins in DSB repair (Featherstone and Jackson, 1999). It has been proposed that a helicase and an exonuclease are required in DSB repair to remove the single-stranded overhangs. Thus, it is tempting to speculate that WRN provides both of these enzymatic functions during DSB repair. However, this model predicts that WS cells would have a defect in DSB repair, and there is no evidence of such a defect at the cellular level at present. Alternatively, the interaction between WRN and Ku may have a role in replication, but the role of Ku in that process is still not clear. Ku-deficient transgenic mice have a distinct senescent phenotype, suggesting that there may be another pathway in which WRN and Ku may cooperate.

Ku, the Ku70 : Ku80 heterodimer, can bind DNA and the catalytic subunit of DNA-PK. In the presence of DNA, Ku activates the kinase activity of DNA-PKcs. Ku, however, may have other functions, independent of DNA-PKcs. Ku can bind at DNA double-strand breaks, or at junctions between DNA single- and double-strand regions. Once bound, Ku can move along the DNA in an ATP-independent manner. There have been some suggestions that Ku might have helicase activity, but recent studies suggest that it does not (Cooper *et al*., 2000). Since Ku strongly interacts with the WRN helicase it may attract that helicase activity to the site. Multiple Ku units can load on to the same DNA segment (Frit *et al*., 2000). In addition to DNA-PKcs, Ku may bind other repair-related proteins and perhaps serve as nucleus of a multiprotein repair focus at a point of DNA damage. Ku binds to telomeric ends and may participate in DNA repair or telomerase repair functions here.

What is the Role of WRN in DNA Repair?

In summary, several lines of evidence support the notion that WRN is involved in DNA damage recognition and processing. The WRN exonuclease may function in an early step of DNA repair to recognize DNA lesions. The enzyme recognizes and arrests at some oxidative DNA base lesions. This arrest may then attract DNA repair proteins to

the site, including RPA, p53 and PCNA, which have also been implicated in early steps of DNA damage recognition. WRN is likely to be in a protein complex involving various DNA replication proteins (Lebel *et al*., 1999) and this complex is also likely to contain proteins involved in BER.

DNA REPAIR POLYMERASES

There have been major developments in our understanding of the role of various mammalian DNA polymerases. There are now about 12 characterized DNA polymerases and more are constantly being discovered. In **Table 2** we list some of the properties of these. One of the interesting features under study is that they differ considerably in their fidelity or proofreading of the DNA template. Mutations in these polymerases have been directly associated with human cancer-prone diseases such as xeroderma pigmentosum complementation group V (Woodgate, 1999). This field is evolving very rapidly, and new polymerases are constantly being discovered.

THE Hus1:Rad1:Rad9 SLIDING CLAMP AND Rad17 CLAMP-LOADER MODEL

Hus1, Rad1, Rad9 and Rad 17 are components of a so-called 'Rad checkpoint' DNA damage response system that is conserved from yeast to humans. Although the role of these molecules in mammalian cells is not yet clear, recent evidence suggests a provocative model for their molecular mode of function (Rauen *et al*. (2000) and references cited therein). Hus1, Rad1 and Rad9 form a heterotrimer complex that resembles PCNA in structure and function. PCNA forms a trimeric clamp that encircles the DNA like a doughnut and binds DNA polymerase, with which it can slide along the DNA, keeping the polymerase from falling off. Like PCNA, the Hus1:Rad1:Rad9 heterotrimer may form a clamp around the DNA. It then perhaps slides along the DNA in search of points of damage, or perhaps it binds and keeps in place a repair polymerase. As in the case of PCNA, a 'clamp loader' is needed to open the trimeric doughnut and reassemble it around the DNA. The clamp loader for PCNA is RFC, a five-subunit protein. Recent evidence suggest that the clamp loader for Hus1:Rad1:Rad9 may be Rad17, which shows sequence homology with all five subunits of PCNA (Rauen *et al*., 2000).

Hus1:Rad1:Rad9 trimer normally is distributed throughout the nucleus. In response to DNA damage, however, it concentrates in foci and becomes difficult to extract. This may reflect the clamping of the trimer around the DNA. The association of Rad17 with the trimer, however, is transient, as would be expected for a clamp loader.

The biological functions of this system are still poorly understood. The homologous system in yeast seems to control the S and G2 cell cycle checkpoints. In mammals, Hus1 may do more than just monitor DNA damage. Hus1 is expressed in all examined tissues and throughout embryonic development. *Hus1*-knockout mouse embryos are able to complete gastrulation, but shortly afterwards develop severe abnormalities and die midway in gestation. Some of the embryos survived to the point of having a beating heart. Cells in the *Hus1*-null embryos proliferated at a normal rate, but died by apoptosis at an abnormally high rate (Weiss *et al*., 2000). Cells from *Hus1*-null mouse embryos exhibit increased spontaneous chromosomal abnormalities, suggesting that Hus1 function is needed to maintain chromosome stability (Weiss *et al*., 2000).

Table 2 Eukaryotic polymerases and their fidelity in replication of undamaged DNA. (Adapted from Wang, 1999.)

Polymerase	Error frequency (mutations per base pair)
Alpha (α)	10^{-4}
Delta (δ)	10^{-5}
Epsilon (ε)	10^{-6}
Gamma (γ)	10^{-6}
Beta (β)	5×10^{-4}
Zeta (ζ)	10^{-4}
Eta (η)	10^{-2}
Iota (ι)	10^{-1}
Theta (θ)	?
Kappa (κ)	10^{-4}
Lambda (λ)	?
Mu (μ)	10^{-1}

MULTIMOLECULAR ASSEMBLIES AND NUCLEAR FOCI

DNA repair functions may often be organized in large multimolecular structures. Often these structures are assembled in large nuclear foci that can be seen by means of fluorescent antibodies. Components of the repair systems may be localized in foci which function as repair factories where DNA lesions could be brought for processing. Exchangeable components might then be efficiently shared among several repair tasks simultaneously in progress. Nuclear foci sometimes undergo rearrangement or remodelling during the cell cycle or in response to DNA damage or other types of stress. Nuclear foci or molecular repair assemblies have been found involving the Rad51 system already discussed in the section on homologous recombination. Other repair systems that may

function as multimolecular assemblies include systems based on Ku (see earlier) and on BRCA1.

A BRCA1-associated Genome Surveillance Complex (BASC)

BRCA1 has several protein-binding domains, and may be associated with large multiprotein complexes in the nucleus, including the DNA repair-related proteins MSH2, MSH6, MLH1, ATM, BLM and the Rad50:Mre11:Nbs complex (Wang *et al*., 2000a,b). The exact structure and function of these complexes is still not clear, however, because it has been difficult to demonstrate specific direct binding in cells (Jun Qin, personal communication). BRCA1-based multimolecular foci perhaps assemble in alternative arrangements with different components, making it difficult to establish individual interactions *in vivo*. In addition, BRCA-1 directly participates in TCR of oxidative DNA damage.

THE DNA REPLICATION CHECKPOINT

DNA damage during S phase normally causes temporary arrest of DNA replication. Proteins required for this checkpoint include ATM, Mre11, Rad50 and NBS1. Genetic defects occur in ATM (ataxia telangiectasia), NBS1 (Nijmegen breakage syndrome) and Mre11 (ataxia telangiectasia-like disorder) (see the chapter *Inherited Predispositions to Cancer*). In all three syndromes, DNA damage (e.g. by ionizing radiation) fails to arrest replication and leads to extensive chromosome damage (Petrini, 2000). Although the mechanistic details are not yet in, we know a few steps in the process. DNA damage causes phosphorylation of NBS1. This is required for replication arrest. ATM senses the DNA damage and carries out the phosphorylation of NBS1. The activity of JNK (c-Jun N-terminal kinase), which is normally stimulated in response to DNA damage, fails to respond in cells derived from patients having any of the three syndromes (see review by Petrini (2000)). **Figure 6** includes the molecular interactions for which there is evidence.

REFERENCES

Aebi, S., *et al*. (1996). Loss of DNA mismatch repair in acquired resistance to cisplatin. *Cancer Research*, **56**, 3087–3090.

Balajee, A. S. and Bohr, V. A. (2000). Genomic heterogeneity of nucleotide excision repair. *Gene*, **250**, 15–30.

Balajee, A. S., *et al*. (1997). Reduced RNA polymerase II transcription in intact and permeabilized Cockayne syndrome group B cells. *Proceedings of the National Academy of Sciences of the USA*, **94**, 4306–4311.

Balajee, A. S., *et al*. (1999). The Werner syndrome protein is involved in RNA polymerase II transcription. *Molecular Biology of the Cell*, **10**, 2655–2668.

Batty, D. P. and Wood, R. D. (2000). Damage recognition in nucleotide excision repair of DNA. *Gene*, **241**, 193–204.

Bohr, V. A., *et al*. (2000). Werner syndrome protein: biochemical properties and functional interactions. *Journal of Experimental Gerontology*, **35**, 695–702.

Breivik, J. and Gaudernack, G. (1999). Genomic instability, DNA methylation, and natural selection in colorectal carcinogenesis. *Cancer Biology*, **9**, 245–254.

Brosh, R. M. Jr, *et al*. (1999). Functional and physical interaction between WRN helicase and human replication protein A. *Journal of Biological Chemsitry*, **274**, 18341–18350.

Brosh, R. M., Jr, *et al*. (2000a). Potent inhibition of Werner and Bloom helicases by DNA minor groove binding drugs. *Nucleic Acids Research*, **28**, 2420–2430.

Brosh, R. M., Jr, *et al*. (2000b). Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity. *Journal of Biological Chemistry*, **275**, 23500–23508.

Buermeyer, A. B., *et al*. (1999). Mammalian DNA mismatch repair. *Annual Review of Genetics*, **33**, 533–564.

Calin, G. A., *et al*. (2000). Genetic progression in microsatellite instability high (MSI-H) colon cancers correletes with clinico-pathological parameters: a study of the TGFβRII, BAX, hMSH3, hMSH6, IGFIIR and BLM Genes. *International Journal of Cancer (Pred. Oncology)*, **89**.

Coleman, W. B. and Tsongalis, G. J. (1999). The role of genomic instability in human carcinogenesis. *Anticancer Research*, **19**, 4645–4664.

Cooper, M. P., *et al*. (2000). Ku complex interacts with and stimulates the Werner protein. *Genes and Development*, **14**, 907–912.

Croteau, D. L., *et al*. (1999). Mitochondrial DNA repair pathways. *Mutation Research*, **434**, 137–148.

de Boer, J. and Hoeijmakers, H. J. (2000). Nucleotide excision repair and human syndromes. *Carcinogenesis*, **21**, 453–460.

DeWeese, T. L., *et al*. (1998). Mouse embryonic stem cells carrying one or two defective Msh2 alleles respond abnormally to oxicative stress inflicted by low-level radiation. *Proceedings of the National Academy of Sciences of the USA*, **95**, 11915–11920.

Dianov, G. L., *et al*. (1999). Replication protein A stimulates proliferating cell nuclear antigen-dependent repair of abasic sites in DNA by human cell extracts. *Biochemistry*, **38**, 11021–11025.

Dizdaroglu, M. (1991). Chemical determination of free radical-induced damage to DNA. *Free Radicals in Biology and Medicine*, **10**, 225–242.

Dronkert, M. L. G., *et al*. (2000). Mouse RAD54 affects DNA double-strand break repair and sister chromatid exchange. *Molecular Cellular Biology*, **20**, 3147–3156.

Drummond, J. T., *et al*. (1996). Cisplatin and adriamycin resistance are associated with MutLα and mismatch repair deficiency in an ovarian tumor cell line. *Journal of Biological Chemistry*, **271**, 19645–19648.

Duckett, D. R., *et al*. (1999). hMutSα- and hMutLα-dependent phosphorylation of p53 in response to DNA methylator damage. *Proceedings of the National Academy of Sciences of the USA*, **96**, 12384–12388.

Featherstone, C. and Jackson, S. P. (1999). Ku, a DNA repair protein with multiple cellular functions? *Mutation Research*, **434**, 3–15.

Fink, D., *et al*. (1998). Enrichment for DNA mismatch repair-deficient cells during treatment with cisplatin. *International Journal of Cancer*, **77**, 741–746.

Fishel, R. (1998). Mismatch repair, molecular switches, and signal transduction. *Genes and Development*, **12**, 2096–2101.

Flores-Rozas, H. and Kolodner, R. D. (2000). Links between replication, recombination and genome instability in eukaryotes. *Trends in Biochemical Science*, **25**, 196–200.

Friedberg, E. C. (1996). Cockayne syndrome – a primary defect in DNA repair, transcription, both or neither? *BioEssays*, **18**, 731–738.

Frit, P., *et al*. (2000). Ku entry into DNA inhibits inward DNA transactions *in vitro*. *Journal of Biological Chemistry*, **275**, 35684–35691.

Fritzell, J. A., *et al*. (1997). Role of DNA mismatch repair in the cytotoxicity of ionizing radiation. *Cancer Research*, **57**, 5143–5147.

Gong, J. G., *et al*. (1999). The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage. *Nature*, **399**, 806–809.

Grady, W. M., *et al*. (1998). Mutation of the type II transforming growth factor-beta receptor is coincident with the transformation of human colon adenomas to malignant carcinomas. *Cancer Research*, **58**, 3101–3104.

Jasin, M. (2000). Chromosome breaks and genomic instability. *Cancer Investigations*, **18**, 78–86.

Jiricny, J. and Nystrom-Lahti, M. (2000). Mismatch repair defects in cancer. *Current Opinions in Genetic Development*, **10**, 157–161.

Karow, J. K., *et al*. (2000). The Bloom's syndrome gene product promotes branch migration of Holliday junctions. *Proceedings of the National Academy of Sciences of the USA*, **97**, 6504–6508.

Karran, P. (2000). DNA double strand break repair in mammalian cells. *Current Opinions in Genetic Development*, **10**, 144–150.

Kohn, K. W. (1996). Beyond DNA crosslinking: history and prospects of DNA-targeted cancer treatment – 15th Bruce F. Cain Memorial Award Lecture. *Cancer Research*, **56**, 5533–5546.

Kohn, K. W. (1999). Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Molecular Biology of the Cell*, **10**, 2703–2734.

Lebel, M., *et al*. (1999). The Werner syndrome gene product co-purifies with the DNA replication complex and interacts with PCNA and topoisomerase I. *Journal of Biological Chemistry*, **274**, 37795–37799.

Lengauer, C., *et al*. (1998). Genetic instabilities in human cancers. *Nature*, **396**, 643–648.

Le Page, F., *et al*. (2000). Transcription coupled repair of 8-oxoguanine in murine cells: the ogg1 protein is required for repair in nontranscribed sequences but not in transcribed sequences. *Proceedings of the National Academy of Sciences of the USA*, **97**, 8397–8402.

Li, G.-M. (1999). The role of mismatch repair in DNA damage-induced apoptosis. *Oncology Research*, **11**, 393–400.

Longley, M. J., *et al*. (1998). Identification of 5′-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair *in vitro*. *Proceedings of the National Academy of Sciences of the USA*, **95**, 12244–12248.

Machwe, A., *et al*. (2000). Selective blockage of the $3' \rightarrow 5'$ exonuclease activity of WRN protein by certain oxidative modifications and bulky lesions in DNA. *Nucleic Acids Research*, **28**, 2762–2770.

Markowitz, S. (2000). TGFβ receptors and DNA repair genes, coupled targets in a pathway of human colon carcinogenesis. *Biochimica Biophysica Acta*, **1470**, M13–M20.

Massague, J. (1998). TGFb signal transduction. *Annual Review of Biochemistry*, **67**, 753–791.

Massague, J. and Wotton, D. (2000). Transcriptional control by the TGFβ/Smad signaling system. *EMBO Journal*, **19**, 1745–1754.

Namsaraev, E. A. and Berg, P. (2000). Rad51 uses one mechanism to drive DNA strand exchange in both directions. *Journal of Biological Chemistry*, **275**, 3970–3976.

Nick McElhinny, S. A., *et al*. (2000). Ku recruits the XRCC4-Ligase IV complex to DNA ends. *Molecular Cellular Biology*, **20**, 2996–3003.

Parsons, C. A., *et al*. (2000). Precise binding of single-stranded DNA termini by human RAD52 protein. *EMBO Journal*, **19**, 4175–4181.

Paull, T. T. and Gellert, M. (2000). A mechanistic basis for Mre11-directed DNA joining at microhomologies. *Proceedings of the National Academy of Sciences of the USA*, **97**, 6409–6414.

Petrini, J. H. J. (2000). The Mre11 complex and ATM: collaborating to navigate S phase. *Current Opinions in Cell Biology*, **12**, 293–296.

Pinz, K. G. and Bogenhagen, D. F. (1998). Efficient repair of abasic sites in DNA by mitochondrial enzymes. *Molecular Cellular Biology*, **18**, 1257–1265.

Prasad, R., *et al*. (2000). FEN1 stimulation of DNA polymerase β mediates an excision step in mammalian long patch base excision repair. *Journal of Biological Chemistry*, **275**, 4460–4466.

Rauen, M., *et al*. (2000). The human checkpoint protein hRad17 interacts with the PCNA-like proteins hRad1, hHus1, and hRad9. *Journal of Biological Chemistry*, **257**, 29767–29771.

Rooney, P. H., *et al*. (1999). Comparative genomic hybridization and chromosomal instability in solid tumours. *British Journal of Cancer*, **80**, 862–873.

Rosen, N. (1997). Cancers of the gastrointestinal tract. In: Devita, V. T., Jr, *et al*. (eds), In: *Cancer: Principles and Practice of Oncology*. 971–1054 (Lippincott-Raven, Philadelphia).

Schild, D., *et al*. (2000). Evidence for simultaneous protein interactions between human Rad51 parallogs. *Journal of Biological Chemistry*, **275**, 16443–16449.

Selby, C. P. and Sancar, A. (1997). Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *Journal of Biological Chemistry*, **272**, 1885–1890.

Strathdee, G., *et al*. (1999). A role for methylation of the hMLH1 promoter in loss of hMLH1 expression and drug resistance in ovarian cancer. *Oncogene*, **18**, 2335–2341.

Takata, M., *et al*. (2000). The Rad51 paralog Rad51B promotes homologous recombinational repair. *Molecular Cellular Biology*, **20**, 6476–6482.

Tantin, D., *et al*. (1997). Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Molecular Cellular Biological*, **17**, 6803–6814.

Toft, N. J., *et al*. (1999). Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *Proceedings of the National Academy of Sciences of the USA*, **96**, 3911–3915.

Vogel, H., *et al*. (1999). Deletion of Ku86 causes early onset of senescence in mice. *Proceedings of the National Academy of Sciences of the USA*, **96**, 10770–10775.

Wang, T. S. F. (1999). Cellular DNA polymerases. In: DePamphilis, M. L. (ed.), In: *DNA Replication in Eukarytoic Cells*. 461–493 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Wang, W., *et al*. (2000a). Possible association of BLM in decreasing DNA double strand breaks during DNA replication. *EMBO Journal*, **19**, 3428–3435.

Wang, Y., *et al*. (2000b). BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes and Development*, **14**, 927–939.

Weiss, R. S., *et al*. (2000). Inactivation of mouse *Hus1* results in genomic instability and impaired responses to genotoxic stress. *Genes and Development*, **14**, 1886–1898.

Wiseman, H. and Halliwell, B. (1996). Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer. *Biochemical Journal*, **313**, 17–29.

Woodgate, R. (1999). A plethora of lesion-replicating DNA polymerases. *Genes and Development*, **13**, 2191–2195.

FURTHER READING

Balajee, A. S. and Bohr, V. A. (2000). Genomic heterogeneity of nucleotide excision repair. *Gene*, **250**, 15–30.

Batty, D. P. and Wood, R. D. (2000). Damage recognition in nucleotide excision repair of DNA. *Gene*, **241**, 193–204.

Buermeyer, A. B., *et al*. (1999). Mammalian DNA mismatch repair. *Annual Review of Genetics*, **33**, 533–564.

Flores-Rozas, H. and Kolodner, R. D. (2000). Links between replication, recombination and genome instability in eukaryotes. *Trends in Biochemistry Science*, **25**, 196–200.

Haber, J. E. (2000). Partners and pathways: repairing a double-strand break. *Trends in Genetics*, **16**, 259–264.

Jasin, M. (2000). Chromosome breaks and genomic instability. *Cancer Investigation*, **18**, 78–86.

Jiricny, J. and Nystrom-Lahti, M. (2000). Mismatch repair defects in cancer. *Current Opinions in Genetics Development*, **10**, 157–161.

Lengauer, C., *et al*. (1998). Genetic instabilities in human cancers. *Nature*, **396**, 643–648.

Lindahl, T. and Wood, R. D. (1999). Quality control by DNA repair. *Science*, **286**, 1897–1905.

McCullough, A. K., *et al*. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annual Review of Biochemistry*, **68**, 255–285.

Oshima, J. (2000). The Werner syndrome protein: an update. *BioEssays*, **22**, 894–901.

Pegg, A. E. (2000). Repair of O^6-alkylguanine by alkyltransferases. *Mutation Research*, **462**, 83–100.

Vessey, C. J., *et al*. (2000). Genetic disorders associated with cancer predisposition and genomic instability. *Progress in Nucleic Acids Research*, **63**, 189–221.

Chapter 8

Telomerase

Carmela P. Morales, Woodring E. Wright and Jerry W. Shay
The University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA

CONTENTS

- Introduction
- Telomeres and the 'End-replication' Problem
- Telomeres are the Divisional Clock
- Telomerase
- Assays for Telomerase
- Association between Telomerase Activity and Cancer
- Exploitation of Telomerase in Cancer Diagnostics and Prognostics
- Telomerase Inhibition in Cancer
- Telomerase for the Development of *In Vitro* Models of Cancer Progression

INTRODUCTION

Intense interest in the enzyme telomerase has occurred in the field of cancer biology over the past 5 years. Recent evidence suggests a role for telomerase during the multi-step process of carcinogenesis. Reactivation or upregulation of telomerase is found in the majority of human cancers and appears to be responsible for the limitless replicative potential of malignant cells, a hallmark of cancer. In normal cells, the 'replicative lifespan' is tightly regulated by an internal divisional clock which limits the number of divisions that a cell can undergo during its lifetime. This divisional clock, known as the telomere, is located at each end of all linear chromosomes. By setting maximum limits on the number of times a cell can divide, telomeres may serve to prevent genetically aberrant cells from accumulating the additional mutations they need to become malignant. For the successful propagation and continued growth of malignant cells, therefore, telomere control on cell growth must be subverted. Current evidence indicates that in almost all human cancers this is achieved by reactivation of the enzyme telomerase.

TELOMERES AND THE 'END-REPLICATION' PROBLEM

Telomeres are long stretches of noncoding DNA located at the ends of all eukaryotic chromosomes (**Figure 1**). In vertebrates, telomeres are comprised of simple, repetitive noncoding DNA sequences. Human telomeres contain the six base pair sequence TTAGGG, repeated many thousands of times.

The length of telomeres varies from chromosome to chromosome. Evidence suggests that as chromosome 'caps,' telomeres have at least three critical functions: to protect chromosome ends from enzymatic degradation and abnormal fusion reactions; to serve as a buffer zone to protect against the 'end-replication' problem; and to serve as a gauge for mitotic age (the divisional clock).

The role of telomeres in maintaining chromosomal integrity was proposed by Barbara McClintock in 1941. Studying telomeres in maize chromosomes, McClintock observed that if not capped by telomeres, the ends of chromosomes had a tendency to fuse. Her observations

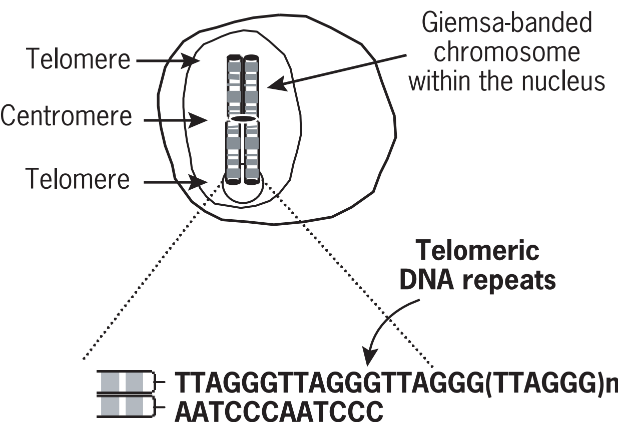


Figure 1 Location of telomeres at the ends of all eukaryotic chromosomes. Telomeres are comprised of the simple DNA sequence TTAGGG, repeated several thousand times (*n*).

were confirmed 50 years later in yeast and mice when it was demonstrated that without telomeric ends, chromosomes undergo aberrant end-to-end fusions, forming multicentric chromosomes with a propensity to break during mitosis, activating DNA damage checkpoints, and in some cases leading to widespread cell death (Zakian, 1989).

The second major function of telomeres relates to the process of semiconservative DNA replication. During each round of cell division, 50–200 base pairs are lost from the ends of linear human chromosomes (Hastie *et al*., 1990; Lindsey *et al*., 1991; Allsopp *et al*., 1992, 1995). This ‘end-replication’ problem (**Figure 2**) occurs because conventional DNA replication machinery is unable to replicate completely the 3′ ends of chromosomal DNA during the S phase of each cell cycle. The polymerases that copy parental DNA strands prior to cell division synthesize DNA only in the 5′ to 3′ direction and require a short RNA primer to begin. These primers are then degraded and filled in by DNA synthesis extending from the upstream primer. However, at the end of a linear chromosome there is no ‘upstream’ DNA synthesis to fill in the gap between the final RNA priming event and the end of the chromosome.

This replication strategy predicts that with each round of cell division, there will be progressive shortening of the 3′ end of chromosomal DNA. Telomeric DNA therefore provides a cushion of expendable noncoding sequences to protect against the potentially catastrophic attrition of important chromosomal material.

TELOMERES ARE THE DIVISIONAL CLOCK

The existence of an internal divisional clock was first suggested in 1965 by Leonard Hayflick, who demonstrated that cells maintained in culture have a finite capacity to proliferate. In 1972, Olovnikov suggested that erosion of the chromosome ends could lead to the loss of essential genes and an exit from the cell cycle. Harley *et al*. introduced a modification of the Olovnikov theory, proposing a telomere-based mechanism to account for the process of ‘cellular ageing’ (Harley *et al*., 1990). Specifically, it was proposed that after a certain number of divisions, telomeres are no longer sufficient to protect chromosome ends from degradation and aberrant fusion reactions. Through signalling mechanisms that are not entirely understood, a few short telomeres may trigger exit from the cell cycle at G1 and entry into senescence, a postmitotic state characterized not only by a lack of further cell division, but also

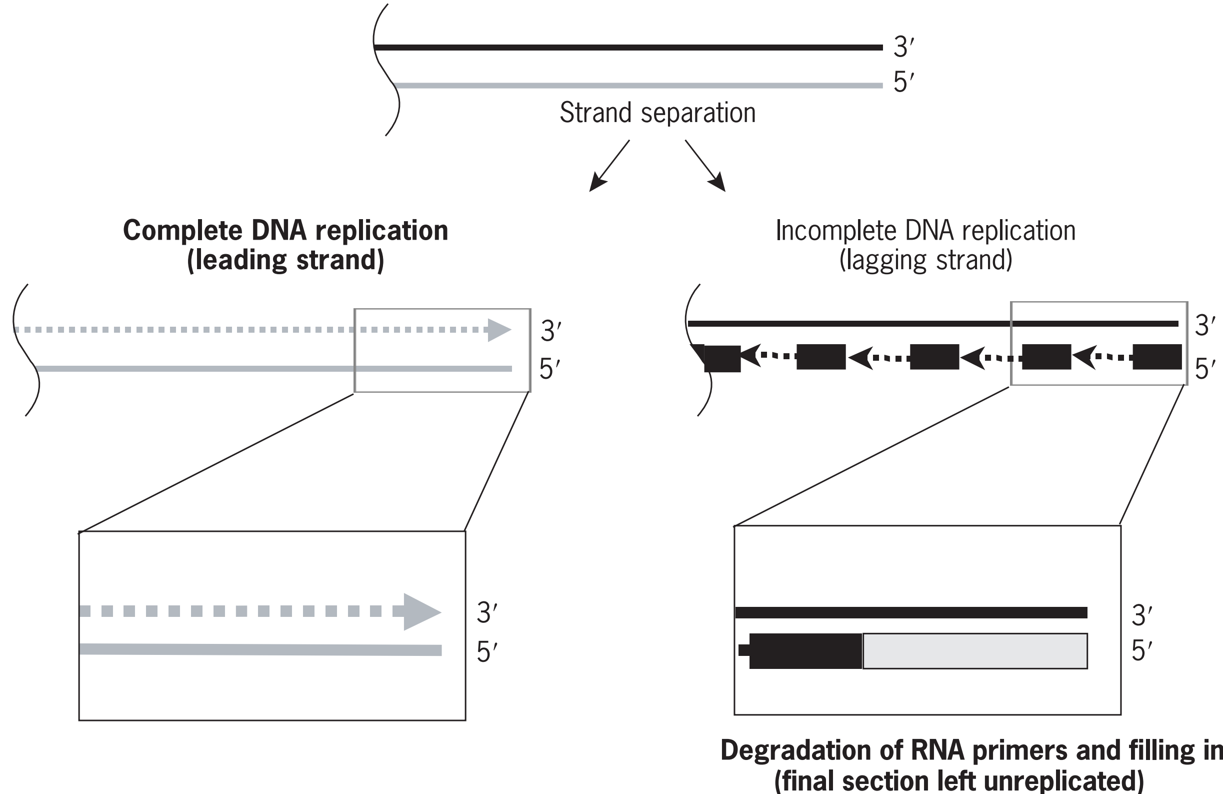


Figure 2 The end replication problem. DNA polymerases require a short RNA primer (black rectangles) to initiate DNA replication in the 3′ to 5′ direction. Since the extreme 3′ end of the chromosome cannot accommodate a primer, this part of the chromosome cannot be replicated, leading to a loss of telomeric DNA with each round of cell division.

by an altered pattern of gene expression and continued metabolic activity for long periods of time. In the case of normal human fibroblasts, a correlation was found between the number of divisions the cells could undergo in culture and initial telomere length, regardless of the age of the fibroblast donor. Additionally, average telomere length in blood and colonic mucosa was shown to decrease with biological age. In adults, sperm telomeres were found to be several kilobase pairs longer than in somatic tissues. Finally, significantly shorter telomere lengths have been demonstrated in some primary cells from patients with the premature ageing syndrome Hutchinson–Gilford progeria compared with normal age-matched controls. These cells also exhibited a reduced proliferative capacity compared with age-matched controls when maintained in culture. Although a large amount of correlative data supported the notion that telomere length determines the proliferative capacity of human cells, a direct test of this hypothesis (**Figure 3**) was lacking. The ability to elongate telomeres experimentally led to the observation that hybrid cells with artificially long telomeres had a longer lifespan than that of cell hybrids in which telomeres had not been elongated (Wright *et al*., 1996). These observations provided the first direct evidence that telomere length is the counting mechanism that limits the proliferative capacity of human cells.

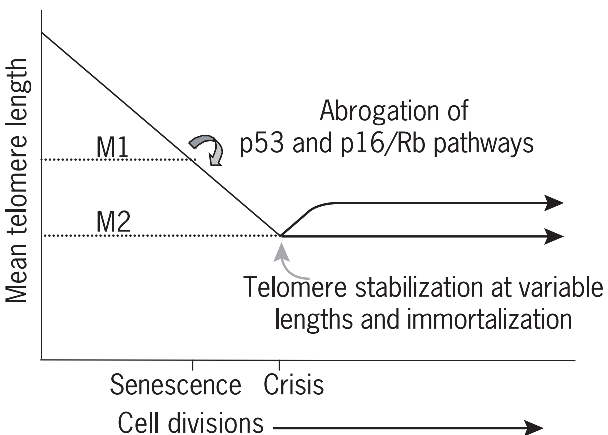


Figure 3 The telomere hypothesis. Telomeric repeats are lost with each round of cell division, leading to a decrease in mean telomere length with accumulated divisions. Upon reaching a certain length, short telomeric DNA sequences trigger entry into the senescence pathway, during which cells remain metabolically active but are no longer able to divide. This stage is also known as mortality stage 1 (M1). Cells can be forced to proliferate beyond this point by abrogating p53 and pRb or their respective pathways. During this period of extended life, cells continue to divide in the face of progressive telomere shortening. After reaching a critically short telomere length, most cells enter a crisis, undergoing widespread cell death (apoptosis). This second stage is also known as mortality stage 2 (M2). Only those rare cells which engage mechanisms that stabilize telomere length are able to continue proliferating for indefinite periods.

The ability to alter cellular proliferative capacity by manipulating telomere length provided a mechanistic basis for earlier observations of cellular lifespan *in vitro*. Normal human fibroblasts maintained in culture undergo a finite number of divisions as determined by their initial telomere length, after which they enter a state of growth arrest (senescence), also known as mortality stage 1 (M1).

Cells nearing the end of their lifespan can be forced to proliferate beyond this point by the introduction of certain viruses or oncogenes that abrogate the function of the tumour-suppressor genes *p53* and *pRb*. These observations suggest that p53 and pRb perhaps mediate cell cycle exit at G1 in response to telomere shortening. Bypass of M1 allows additional rounds of cell division until further, critical telomere shortening occurs, resulting in a state of 'crisis,' characterized by widespread cell death. This second stage is known as mortality stage 2 (M2). As a low-frequency ($\sim 10^{-7}$) event in human cells, a subpopulation of cells escapes from crisis, giving rise to cells which now have an unlimited proliferative capacity (immortalized). The characteristic feature of such immortal cells is the ability to maintain their telomeres.

The dual role of telomeric DNA as protector of chromosomal integrity and mitotic clock implicates cellular senescence as a natural and effective initial protection mechanism against the development of cancer. It is generally believed that tumours are initiated by multiple genetic events in cells which result in the inappropriate activation of growth stimulatory signals, an insensitivity to antigrowth signals and a resistance to apoptosis. However, transformation to fully malignant derivatives does not occur in most cases because the majority of these aberrant cells will have exhausted their endowment of allowed divisions.

TELOMERASE

Early studies demonstrated a significantly shorter telomere length in most cancers compared with noncancerous tissue from the same patient (Hastie *et al*., 1990). In culture, cancer cells generally have short but stable telomeres, suggesting that human cancers have developed strategies for the maintenance of telomeric DNA at a length above the critical threshold. In 85–95% of human cancers, this telomere stabilization is achieved by reactivation or upregulation of the ribonucleoprotein enzyme telomerase.

Telomerase is an RNA–protein complex which utilizes its RNA as a template for the addition of TTAGGG repeats to the 3′ ends of chromosomes, thereby compensating for losses due to the end-replication problem (**Figure 4**).

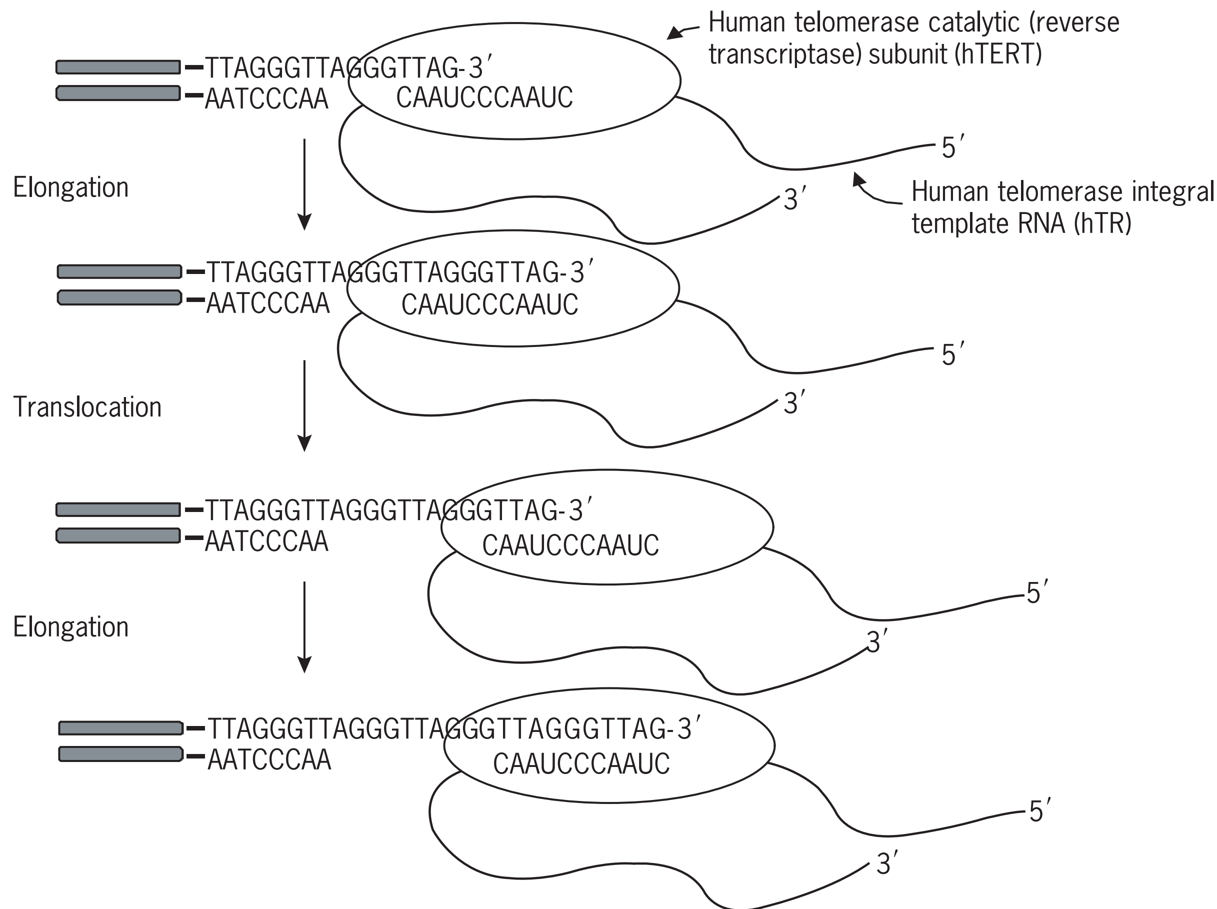


Figure 4 Telomerase is an enzyme minimally comprised of a catalytic protein component and an RNA subunit. The RNA serves as a template for the addition of TTAGGG repeats to the ends of chromosomes, leading to maintenance of telomeric DNA. This enzyme is processive, meaning that a variable number of repeats may be added to a chromosome prior to disengaging the chromosome end.

First discovered in *Tetrahymena* by Elizabeth Blackburn and Carol Greider, telomerase activity has now been detected in extracts from almost all organisms, with the exception of bacteria and viruses, which have circular genomes, and *Drosophila*, which have retrotransposons instead of telomeres. In humans, most adult somatic cells lack telomerase activity. However, telomerase is present at high levels in germ cells, early embryos, activated T and B cells and germinal centres of lymphoid organs. Telomerase activity is also detectable at lower amounts in the basal cells of renewal tissues (skin and intestine). In somatic tissues and cells (including T cells), however, the presence of detectable levels of telomerase activity is not sufficient to prevent long-term telomere attrition. The variable levels of telomerase in normal human tissues are illustrated in **Figure 5**.

In humans, telomerase is composed of two essential components: an integral RNA (hTR), which provides the template for the synthesis of telomere repeats, and a protein subunit (hTERT), which provides catalytic activity. The cloning of hTERT and hTR made it possible to test directly the hypothesis that telomere shortening regulates the entry into cellular senescence. Using an *in vitro* system, the combination of hTERT and hTR was first shown to reconstitute telomerase activity (Weinrich *et al*., 1997; Beattie *et al*., 1998). Second, the introduction of hTERT into telomerase-negative primary cells resulted in telomerase activity. Exogenous hTR was not needed because it is present even in cells that do not normally have telomerase activity. Finally, it was shown that telomere maintenance by exogenous telomerase was sufficient for the immortalization of human mammary epithelial cells, foreskin fibroblasts, retinal pigmented epithelial cells and umbilical vascular endothelial cells. Taken together, these experiments provided direct evidence that short telomere length directs entry into cellular senescence and that telomere maintenance by telomerase is sufficient to bypass this growth arrest under most circumstances. Importantly, the introduction of telomerase prior to either M1 or M2 is sufficient for immortalization, indicating that telomeres

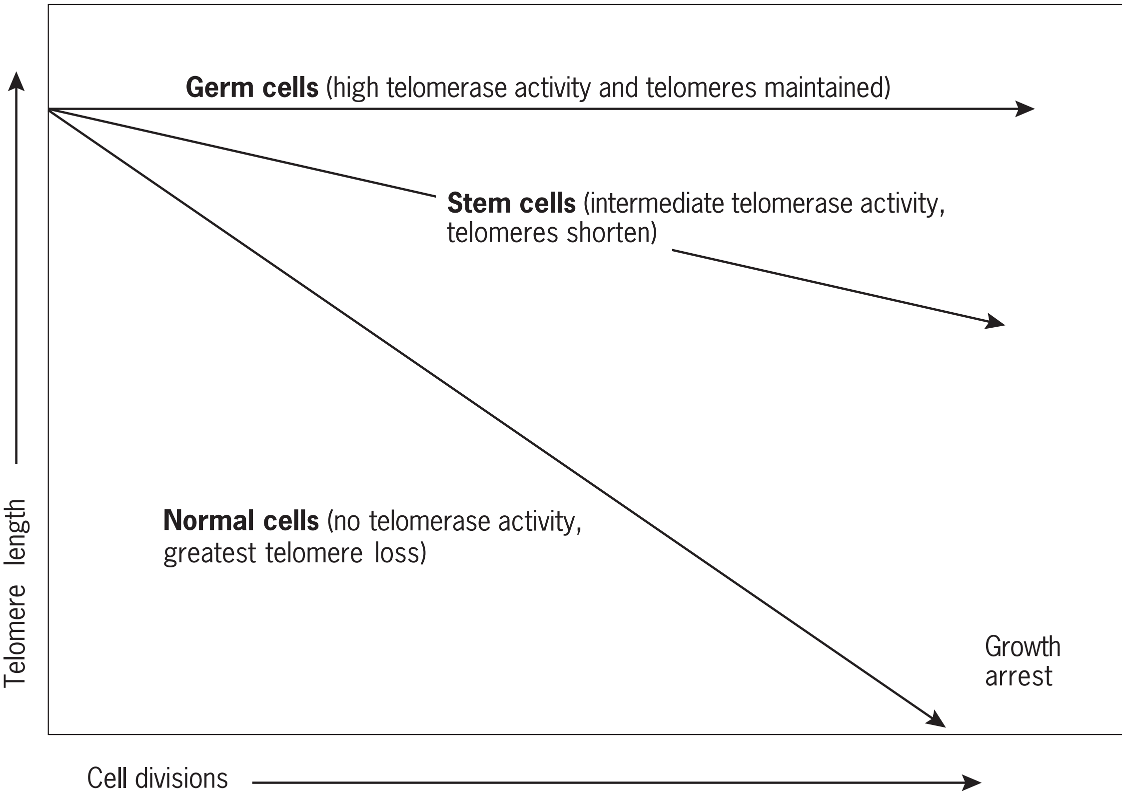


Figure 5 Variable levels of telomerase in normal human tissues. Germ cells of the reproductive system maintain high levels of telomerase activity throughout life and therefore do not sustain telomere shortening. Stem cells of renewal tissues express modest levels of telomerase, leading to a blunted rate of telomere shortening. Somatic cells and tissues lack detectable telomerase activity and sustain the greatest rates of telomere loss.

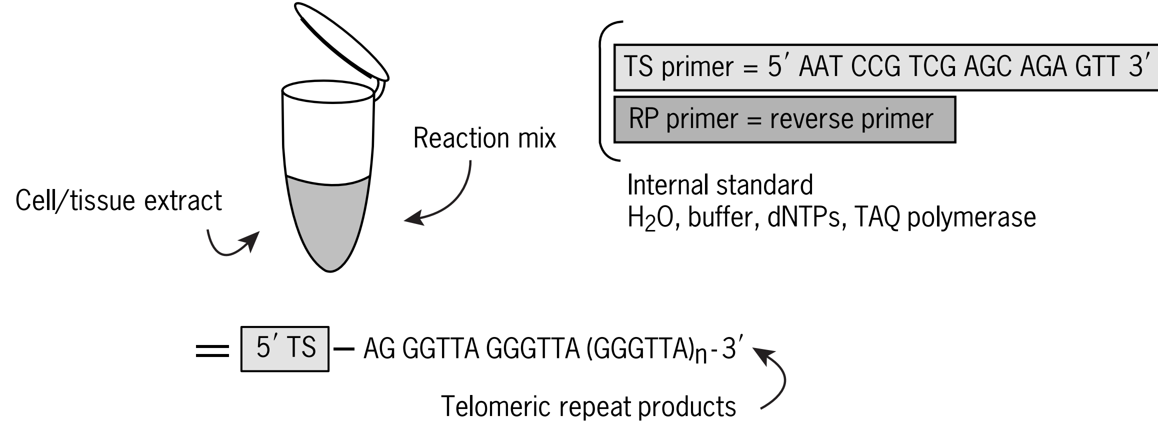
are associated with both the M1 and M2 stage of growth arrest.

Various factors such as oxidative stress, introduction of an activated Ha-*ras* oncogene and γ-irradiation have been shown to induce a senescent-like state in cells much younger than the Hayflick limit. Several reports have also suggested that in addition to telomerase, inactivation of the Rb/p16^{INK4a} pathway is required for the immortalization of some human epithelial cells. The protein p16^{INK4a} is an inhibitor of cyclin-dependent kinases, and its levels have been shown to increase after only a few passages in culture. During the establishment of human mammary epithelial cells under standard culture conditions, there appears to be a 'self-selection' process, such that only the cells which have lost p16^{INK4a} expression (usually by methylation of the promoter) are able to survive the initial culture period. In these surviving cells, exogenous telomerase expression leads to immortalization. In most cases, no appreciable decrease in telomere length can be demonstrated to account for this 'self- selection' process, termed M0. These observations suggest that the involvement of p16^{INK4a} at M0 is telomere-independent, a finding that stands in direct contrast to the role of p53 in mediating of growth arrest at the telomere-dependent M1 stage. Recent observations in keratinocytes suggest that the loss of p16^{INK4a} expression is not required if the cells are maintained under optimized culture conditions, such as co-culturing with irradiated fibroblast feeder layers (Ramirez *et al.*, 2001). These findings suggest that artificial cell culture conditions may account for the premature growth arrest (M0) in epithelial cells, and that this response is mediated by induction of p16^{INK4a}.

ASSAYS FOR TELOMERASE

The standard method for measuring telomerase activity is a highly sensitive PCR-based assay termed the TRAP (telomere repeat amplification protocol) assay (**Figure 6**).

In this assay, extracts are first prepared from primary tissue or cultured cells by lysing the cells with a detergent, releasing telomerase into the extract solution. An aliquot of this solution is then added to a reaction mixture containing a short primer and deoxynucleotide triphosphates (dNTPs). If telomerase is present, it will elongate the primer with TTAGGG repeats. The products of this elongation step are



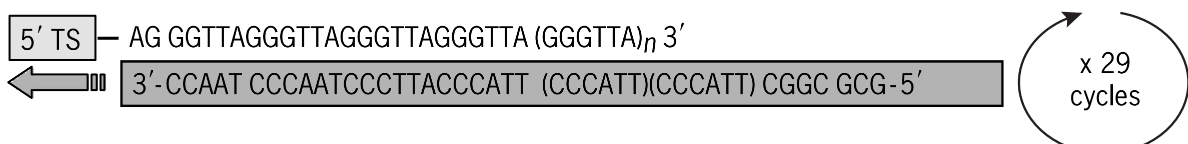


Figure 6 The telomerase activity assay. This sensitive PCR-based assay can be applied to extracts from a variety of cells and tissues. In the first step, extracts are incubated with labelled nucleotides and also a synthetic telomere end (TS primer) at room temperature. If the extract contains telomerase, the enzyme will synthesize the addition of TTAGGG repeats to the TS primer. The second step utilizes PCR to amplify the extended products in the presence of additional primers and polymerases. These amplified products are then separated by electrophoresis to generate the characteristic six-base-pair ladder indicating telomerase activity. ITAS is an internal standard included to control for the possibility of enzyme inhibitors.

heterogeneous in length, representing multiples of the six-base-pair TTAGGG sequence. Using a second primer which is complementary to the telomerase repeat and a DNA polymerase known as TAQ polymerase, each product is amplified using the polymerase chain reaction (PCR). The amplified products are then run on a polyacrylamide gel, creating a six-base-pair ladder. The laddering effect occurs because telomerase is a processive enzyme, adding telomeric repeats in multiples of TTAGGG (**Figure 7**). This amplification protocol increases the sensitivity of the assay such that telomerase activity can be detected in samples containing as few as 0.1% positive cells. An internal standard is also incorporated into the assay, since some tissue extracts contain molecules that inhibit PCR and give false-negative results. In addition, this internal standard permits semiquantitative analysis of relative telomerase activity levels.

Alternative approaches to the measurement of telomerase have recently been developed. Unlike TRAP, which is a functional assay carried out on extracts of cells or tissues, *in situ* techniques are designed to visualize the components of telomerase at a cellular level. *In situ* hybridization for the RNA component of human telomerase (hTR) can be applied to formalin-fixed, paraffin-embedded tissues as well cultured cells and cell smears. Several studies have demonstrated good concordance between telomerase activity as measured by TRAP and telomerase RNA by *in situ* hybridization. Although normal cells do contain hTR, the levels in normal tissues are sufficiently low that they do not complicate the observation of elevated hTR levels in tumours. Antibodies for the immunhistochemical detection of the telomerase protein component are now becoming commercially available, even though their utility remains to be established. Studies using other antitelomerase antibodies, however, been shown to correlate with telomerase activity by TRAP. These techniques may have some advantages over the PCR-based assay. The excellent morphological preservation of cellular detail provided by *in situ* hybridization or immunohistochemistry may be helpful in localizing telomerase to specific cell types. Furthermore, *in situ* telomerase assays could be readily adapted by clinical laboratories, many of which already utilize such techniques for the detection of other proteins.

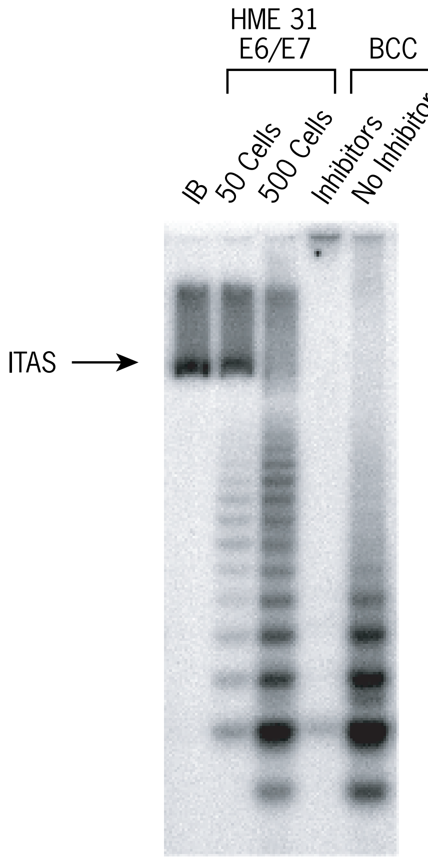


Figure 7 Telomerase activity in two different basal cell carcinomas (BCC). Human mammary epithelial cells expressing the E6 and E7 oncoproteins (HME 31 E6/E7) serve as a positive control for quantitation. The lane containing lysis buffer (LB) serves as a negative control. This representative telomerase assay gel reveals the characteristic 6 base-pair ladder indicative of enzymatic activity, and the internal standard (ITAS) that serves to normalize sample-to-sample variation. Absence of the ITAS signal, as demonstrated in the first BCC sample, indicates the presence of PCR inhibitors. Without this internal control, this sample may be misinterpreted as lacking telomerase activity. Quantitation of telomerase activity is done by determining the ratio of the internal standard to the telomerase ladder.

ASSOCIATION BETWEEN TELOMERASE ACTIVITY AND CANCER

The TRAP assay has made possible the large-scale testing for telomerase activity in a wide variety of human cancers and normal tissues. Using TRAP, telomerase activity has been detected in 85–95% of all human cancers and cancer cell lines, whereas adjacent normal tissue and mortal cells in culture are generally telomerase negative (**Table 1**).

Thousands of individual malignancies representing all of the major organ systems have been tested to date, including those originating from the head and neck, lung, gastrointestinal, pancreatic and biliary tract, liver, breast, male and female reproductive tract, kidney/urinary tract, central nervous system, skin and blood (Shay and Bacchetti, 1997). Preinvasive and preneoplastic lesions, such as colorectal adenomas, high-grade prostatic intraepithelial neoplasia, *in situ* breast carcinoma and those from the head/neck and lung tissue, are positive in 30% to almost 100% of cases. These observations provide strong evidence that most human malignancies are associated with the reactivation or upregulation of telomerase. Given that most normal human cells have the capacity to undergo 60–70 population doublings, it may at first glance seem difficult to invoke telomere shortening as a barrier to cancer formation, because after 60 doublings, a single cell would generate a tumour mass of approximately 10^{15} kg! However, not only do evolving malignancies exhibit high rates of turnover due to chronic, widespread apoptosis and differentiation within the tumour, but also many clonal expansions occur.

A schematic demonstrating the relationship between cell turnover and tumorigenesis is shown in **Figure 8**.

In this scenario, a single healthy cell would be expected to generate a population of 10^6 cells after 20 doublings. As a rare event, one cell in this population acquires a genetic mutation which confers to it a selective growth advantage over the remaining cells. Owing to its newly acquired growth advantage, this cell then generates a clone of similarly altered cells. This process is repeated several times, with each successive generation resulting from an event in a single cell from the previous generation. In some cases, an additional generation is required to convert a minimally functional recessive mutation into a strong phenotype through loss of the remaining wild-type allele (loss of heterozygosity (LOH)). Rarely, a single cell emerges which has acquired sufficient advantageous mutations to result in the development of a malignancy. Using such a scenario, it can be seen that the number of cells in the final tumour grossly underrepresents the number of cells required to produce it. Because so many doublings are required for the development of a cancer, most potentially tumorigenic cells probably senesce owing to critical telomere erosion. Thus, cellular senescence could be viewed as a powerful initial blockade against carcinogenesis.

The timing of telomerase reactivation in human tumours appears to vary considerably from organ to organ. In most cases, the mean telomere length in a variety of tumour types is substantially shorter than those in normal tissues from the same patient. Taken in the context of the telomere hypothesis, these observations suggest that telomerase reactivation or upregulation occurs only after dramatic telomere loss, and serves to stablilize shortened chromosome ends and permit continued cell proliferation. In most colorectal, oesophageal and pancreatic adenocarcinomas, telomerase reactivation appears to be a late event, occurring during the transition from low- to high-grade

Table 1 Telomerase activity in human cancers[a]

Pathology	% Positive	Pathology	% Positive
Head/neck and lung		*Breast*	
Normal oral mucosa	32	Fibrocystic disease/fibroadenoma	0
Head/neck squamous cell carcinoma	86	Carcinoma *in situ*	75
Non-small cell lung carcinoma	78	Carcinoma (ductal and lobular)	88
Small cell lung carcinoma	100	Adjacent tissue	5
Adjacent lung tissue	4	*Reproductive tract*	
Gastrointestinal tract		Normal adult ovary	33
Gastric metaplasia/adenoma[b]	27	Normal myometrium/endometrium	0
Gastric carcinoma	85	Leiomyoma	0
Adjacent gastric tissue	25	Leiomyosarcoma	100
Colorectal adenoma[b]	45	Cervical/vaginal/endometrial cancer	100
Colorectal carcinoma	89	Ovarian carcinoma	91
Adjacent and normal colon tissues[b]	25	Normal adult testis	100
Pancreas and liver		Normal prostate	0
Benign pancreatic lesions (all)	0	BPH without carcinoma[c]	5
Pancreatic carcinoma	95	BPH with carcinoma[c]	11
Adjacent pancreatic tissue	14	High-grade PIN[d]	60
Benign pancreatic brushings	0	Prostate carcinoma	90
Malignant pancreatic brushings	100	*Skin*	
Normal liver tissue	0	Normal epidermis[b]	44
Nonmalignant liver disease (all)[b]	29	Squamous cell carcinoma	83
Hepatocellular carcinoma	86	Basal cell carcinoma	95
Adjacent liver tissue	2	Melanoma	86
Kidney/urinary tract		*Haematological tissues*	
Normal urothelium	0	Myeloma	100
Dysplastic urothelium	43	Lymphoma, low grade	86
Bladder carcinoma (all stages)	92	Lymphoma, high grade	100
Bladder carcinoma (washings)	73	Tonsils, normal	100
Bladder carcinoma (voided urine)	29	Myelodysplastic syndrome	67
Renal cell carcinoma	83	CML, chronic[e]	71
Adjacent renal tissue	0	CML, early accelerated	33
Wilm tumour	100	CML, blast stage	100
Adjacent renal tissue (Wilm)	33	CLL, early[f]	14
Neural tissues		CLL, late	57
Normal retina	0	Acute promyelocytic leukaemia	100
Retinoblastoma	50	Acute lymphocytic leukaemia	80
Glioblastoma multiforme	75	Acute myelogenous leukaemia	73
Oligodendroglioma	100		
Anaplastic astrocytoma	10		
Meningioma, ordinary	17		
Meningioma, atypical	92		
Meningioma, malignant	100		
Ganglioneuroma	0		
Neuroblastoma	94		
Adjacent neural tissue	0		

[a]Adapted from Shay and Bacchetti, 1997. [b]Telomerase activity weak compared with carcinomas. [c]BPH = benign prostatic hypertrophy. [d]PIN = prostatic intraepithelial neoplasia. [e]CML = chronic myelogenous leukaemia. [f]CLL = chronic lymphocytic leukaemia.

dysplasia. Additionally, early-stage neuroblastomas lack or have low levels of telomerase activity, whereas late-stage disease has high levels of telomerase activity. In other malignancies, such as head and neck cancers, lung carcinomas and breast carcinomas, telomerase activity is present in early preneoplastic lesions, albeit at lower levels than frankly malignant tissues. Since most of these data were obtained by studying tissue extracts, it is difficult to determine whether telomerase activity in preneoplastic lesions is due to the infiltration of microscopic quantities of tumour cells or to low-level telomerase activity in preneoplastic cells. The development of *in situ* techniques, such as immunohistochemistry with telomerase antibodies or *in situ* hybridization for the telomerase template RNA

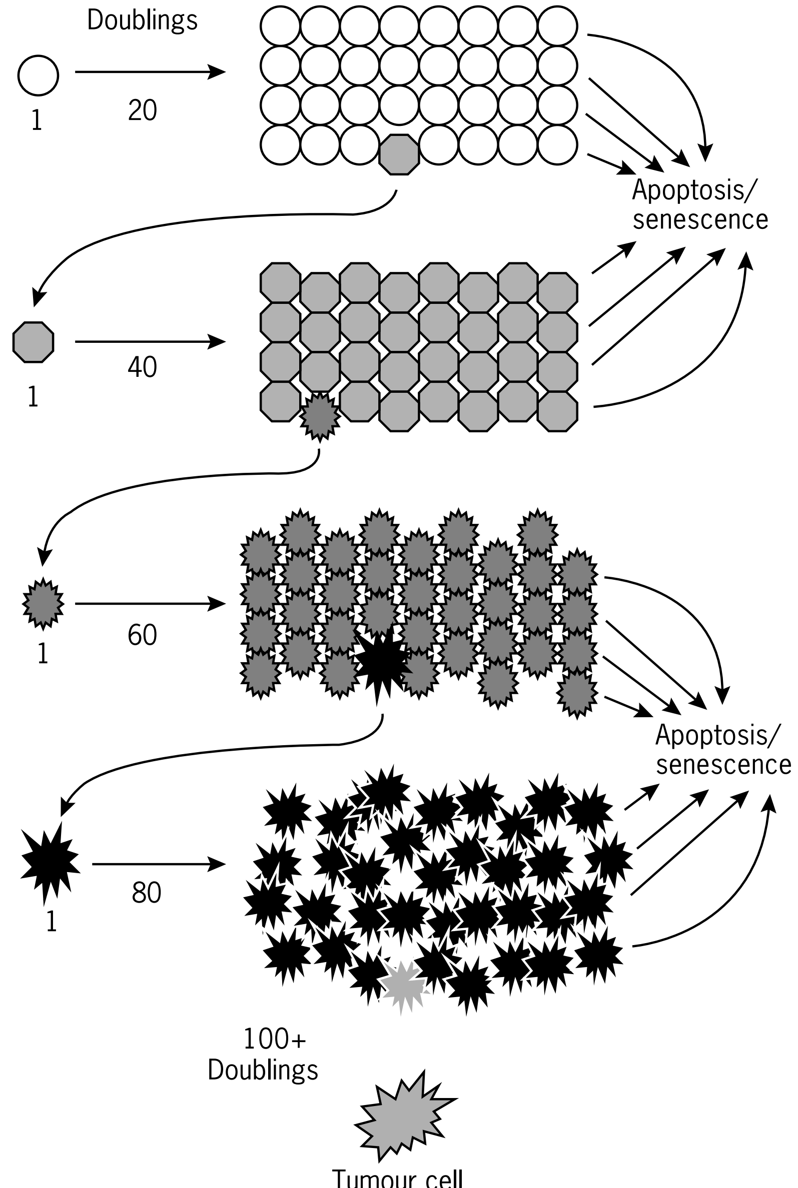


Figure 8 Relationship between cell divisions and tumorigenesis. Within a normal population of cells, a single cell acquires a mutation which endows the cell with a growth advantage. After 20 doublings, a clone of cells emerges, one of which undergoes an additional advantageous mutation. While the other cells senesce or die, this cell survives to generate a clone of similarly mutated cells. The cycle continues until one mutant emerges which has acquired all of the necessary mutations for tumorigenesis. According to this scheme, a fully tumorigenic cell can result in less than 100 doublings.

(hTR), will be important in clarifying these important issues. In the case of cervical cancer, *in situ* hybridization for hTR showed focal increases in hTR expression at the level of *in situ* carcinomas.

Although most human cancers express high levels of telomerase, a substantial portion (10–15%) are telomerase negative. There are several explanations for this observation. First, although the TRAP assay is capable of detecting telomerase activity with only 1–10 tumour cells, because tumours are heterogeneous, some sampled specimens may contain no or insufficient numbers of telomerase-positive tumour cells to be detected by the assay. Second, some cancers may not have reached a point where telomerase activity is required. Such tumours may in fact still be mortal and therefore truly telomerase negative. Third, as mentioned previously, it has been reported that some tissue extracts contain inhibitors of either the elongation or amplification steps of the TRAP assay, leading to false-negative results. Other reasons for false-negative results include technical errors, such as poor sample preservation, sampling error or misloading of the specimen into the reaction mixture. Finally, there is experimental evidence for one or more

alternative mechanisms (possibly based on recombination) for lengthening of telomeres (ALT). Some immortalized cell lines show evidence of ALT activity characterized by the absence of telomerase activity but the presence of very long and heterogeneous telomeres. Although the mechanism is not well known, the existence of an ALT pathway has important theoretical implications for telomerase inhibition as a treatment for cancer. To date, however, there is no experimental evidence suggesting that tumour cells with telomerase can be converted to the ALT pathway.

EXPLOITATION OF TELOMERASE IN CANCER DIAGNOSTICS AND PROGNOSTICS

The strong association between telomerase and most human malignancies has prompted a flurry of studies exploring the potential clinical utility of telomerase as a diagnostic cancer marker. In addition to measuring telomerase activity in tissue extracts, the TRAP assay has been successfully applied to a wide variety of samples including bladder washings, sedimented cells in voided urine and colonic effluent, oral rinses, brushes and washes, endoscopic brushings, biliary aspirates, ascitic fluid, blood, fine needle aspirates and frozen sections. Formalin-fixed, paraffin-embedded pathological material can also be tested for the presence of micrometastasis using recently developed *in situ* hybridization to the telomerase template RNA (hTR) and immunohistochemical detection of the catalytic component (hTERT). In oesophageal carcinomas, a marked increase in hTR occurs during the transition from low- to high-grade dysplasia, suggesting that telomerase is important for the development of advanced lesions in oesophageal carcinogenesis. Although initial data appear promising, these techniques need to be validated by comparison with the standard TRAP assay. However, current evidence demonstrates a good correlation between telomerase activity and *in situ* levels of hTR.

The utility of telomerase in predicting the outcome of cancer (prognosis) is based on the notion that without telomere maintenance, malignant cells will be unable to sustain long-term proliferation and eventually undergo cell death and tumour regression, contributing to a favourable outcome. The best evidence for such a scenario is in stage 4S neuroblastomas that lack detectable telomerase activity. These cases are associated with large rate of spontaneous tumour regression. This indicates that, at least in some cancers, telomerase is not absolutely required for malignancy but that tumours without telomerase may ultimately regress if they do not engage a mechanism for telomere stabilization. In ordinary meningiomas, a strong correlation was found between telomerase activity and disease relapse. Telomerase activity has also been observed to confer a worse prognosis in other malignancies such as neuroblastoma, acute myeloid leukaemia, breast cancer and some gastrointestinal cancers.

Another potential role for telomerase is in the detection of residual disease after surgical resection or adjuvant chemo- and/or radiation therapy. Telomerase activity has been detected in cells adjacent to tumours, suggesting the presence of small foci of residual malignant cells. Such information could be used to restage a lesion and identify a subset of patients who would benefit from additional therapy. Although there are reports that inflammatory cells express low levels of telomerase activity, analysis of malignant lymph nodes reveals levels of telomerase at least sixfold higher than their benign counterparts.

TELOMERASE INHIBITION IN CANCER

Several lines of evidence support the notion that inhibition of telomerase may be an effective anti-cancer strategy. As mentioned previously, telomerase is present in most human malignancies. Although the introduction of certain viral oncoproteins or the abrogation of tumour-suppressor genes may confer an extended lifespan, in the absence of a mechanism for telomere maintenance, these cells eventually reach a period of crisis and undergo widespread cell death.

Telomerase is a challenging molecule for drug development because of the long period required to reach sufficient telomere shortening. To be considered telomerase-specific, inhibitors should fulfil several criteria: (1) inhibitors should reduce telomerase activity without initially affecting proliferation; (2) treatment with inhibitors should result in telomere shortening with each round of cell division; (3) treated cells should eventually undergo growth arrest or apoptosis; (4) there should be a correlation between initial telomere length and time to growth arrest or cell death; (5) control of chemically related molecules or inhibitors lacking the ability to inhibit telomerase activity should not have an effect on cell proliferation and telomere length. Numerous conventional chemotherapeutic agents have been reported to inhibit telomerase. These reports are based on observations that treatment with such agents resulted in widespread cell death and loss of telomerase activity. However, the interpretation of the results is suspect since they do not fulfil the criteria expected for telomerase inhibitors.

There are several important theoretical considerations associated with telomerase inhibitor therapy. First, during the initial period of telomere shortening, continued cell proliferation could result in clinically significant tumour growth. Second, discontinuation of therapy for even short periods of time during the treatment period could result in the rapid induction of telomerase and telomere relengthening. Third, selective pressure on tumour cells being treated with telomerase inhibitors could lead to drug resistance due to the emergence of cells with alternative

mechanisms of telomere maintenance. Finally, there is a theoretical concern that normal cells with telomerase activity (germ cells and renewal tissues) would also be susceptible to telomerase inhibition. However, these cells generally have a longer than average telomere length and are much more slowly dividing than tumour cells, and thus would be expected to be relatively resistant to the consequences of telomerase inhibition.

The telomerase template RNA (hTR) and the catalytic core of the protein subunit (hTERT) are two obvious choices for drug design since both components are absolutely required for telomerase activity. Agents tested to date include antisense oligonucleotides and synthetic peptide nucleic acids targeted against the template region of telomerase (hTR). Using agents that fulfil these criteria, several groups demonstrated that telomerase inhibition in cultured cancer cells resulted in a marked (70–95%) decrease in telomerase activity, telomere shortening and widespread cell death after periods ranging from 2 to 3 months, depending on the initial telomere length. Discontinuation of the drug resulted in a rapid reactivation of telomerase and regrowth of telomeres to their initial length. In tumour cells that were inhibited until the point of cell death, there was no evidence for the development of resistant cells, suggesting that the alternative mechanism of telomere maintenance may not be readily adopted in these cells.

Overall, these initial observations suggest that sustained telomerase inhibition may be an effective and feasible anticancer strategy. However, given the delayed effect of such agents, the most appropriate setting for telomerase inhibition would appear to be the prevention of relapse due to small numbers of remaining cells or to cells resistant to initial conventional therapy. Other strategies would combine telomerase inhibitors with other agents, such as angiogenesis inhibitors to target tumour cells specifically and effectively. Additional novel and potentially effective approaches against telomerase are beginning to emerge, such as ribozymes directed against the template region of telomerase RNA, which cleave the RNA and render telomerase inactive. Molecules which couple the telomerase promoter to apoptotic genes are also in the early stages of development. Such approaches would be expected to have the dual benefit of optimizing tumour cell specificity while producing a more rapid biological effect. However, the effect of such agents on telomerase-competent stem cells and germ-line cells remains an important consideration.

TELOMERASE FOR THE DEVELOPMENT OF *IN VITRO* MODELS OF CANCER PROGRESSION

The introduction of telomerase into some normal human cells resulted in bypass of M1 and immortalization. After a doubling of their normal lifespan, these immortalized cells maintain a normal diploid karyotype and DNA damage and cell cycle checkpoints remained fully intact, suggesting that normal cells immortalized with telomerase do not develop additional cancer-associated changes. In cells expressing the Simian virus 40 Large T antigen for long periods of time followed by the introduction of an activated *ras* oncogene, the addition of telomerase appears to be sufficient for transformation into full tumorigenicity. Taken together, these results suggest that while telomerase expression *per se* does not result in genomic destabilization, in the context of underlying mutations, telomerase contributes to tumorigenicity by providing aberrant cells with an unlimited proliferative capacity.

The ability of telomerase to immortalize cells without altering the underlying genetic background has also been demonstrated in cells with inherited susceptibility syndromes, such as ataxia–telangiectasia, Bloom syndrome, xeroderma pigmentosum and premature ageing syndromes such as Werner syndrome and Hutchinson–Gilford progeria. Thus, telomerase may be an important tool for establishing premalignant cell lines which can be used for the development of *in vitro* models of cancer progression, for amassing large numbers of cells required for other assays or as standard cellular reagents for microarray analysis. Microarray is a novel technique which utilizes microchip technology to analyze cellular RNA for changes in patterns of genetic expression in mutated cells versus their normal counterparts.

The role of telomerase in cancer progression will undoubtedly represent a major continuing area of investigation in the field of cancer biology. The tight association between telomerase and cancer, the ability to generate immortalized human cell lines for studies of cancer progression and the development of telomerase inhibitors for use as anticancer agents all underscore the fundamental role of telomere maintenance as a major player in the development and continued unlimited growth of cancer cells.

REFERENCES

Allsopp, R. C., *et al*. (1992). Telomere length predicts the replicative capacity of human fibroblasts. *Proceedings of the National Academy of Sciences of the USA*, **85**, 10114–10118.

Allsopp, R. C., *et al*. (1995). Telomere shortening is associated with cell division *in vitro* and *in vivo*. *Experimental Cell Research*, **220**, 194–200.

Beattie, T. L., *et al*. (1998). Reconstitution of human telomerase activity *in vitro*. *Current Biology*, **8**, 177–180.

Harley, C. B., *et al*. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature*, **345**, 458–460.

Hastie, N. D., *et al*. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature*, **346**, 866–868.

Lindsey, J., *et al*. (1991). *In vivo* loss of telomeric repeats with age in humans. *Mutation Research*, **256**, 45–48.

Ramirez, R. D., *et al*. (2001). Putative telomere-independent mechanisms of replicative aging reflect inadequate growth conditions. *Genes and Development*, **15**, 398–403.

Shay, J. W. and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. *European Journal of Cancer*, **33**, 787–791.

Weinrich, S. L., *et al*. (1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature Genetics*, **17**, 498–502.

Wright, W. E., *et al*. (1996). Experimental elongation of telomeres extends the lifespan of immortal × normal cell hybrids. *EMBO Journal*, **15**, 1734–1741.

Zakian, V. A. (1989) The structure and function of telomeres. *Annual Reviews of Genetics*, **23**, 579–604.

FURTHER READING

Cech, T. R. (2000). Life at the end of the chromosome: telomeres and telomerase. *Angewandte Chemie International Edition in English*, **39**, 34–43.

Greider, C. W. (1999). Telomerase activation-one step on the road to cancer? *Trends in Genetics*, **15**, 109–112.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Wynford-Thomas, D. (1999). Cellular senescence and cancer. *Journal of Pathology*, **187**, 100–111.

Chapter 9

Apoptosis

John C. Reed
The Burnham Institute, La Jolla, CA, USA

CONTENTS

- Introduction
- Apoptosis is Caused by Proteases
- Inhibitor of Apoptosis (IAP) Family Proteins Function as Caspase Inhibitors
- Receptor-mediated Mechanisms of Caspase Activation
- Mitochondrial Pathways for Caspase Activation
- Bcl-2 Family Proteins—Regulators of Cytochrome *c* Release and More
- Interactions of Signal Transduction Pathways with Apoptosis Pathways
- Therapeutic Implications
- Acknowledgements

INTRODUCTION

Cell death is a part of normal physiology for most metazoan species. During development, redundant or unwanted cells are removed through programmed cell death, making important contributions to morphogenesis, organogenesis and other processes (Vaux and Korsmeyer, 1999). Programmed cell death typically occurs through a ritualistic process known as 'apoptosis,' a term adapted from the Greek and which has analogies with the falling away of leaves from trees in the autumn. Among the features of cells undergoing apoptosis are chromatin condensation, nuclear fragmentation, plasma membrane blebbing, cell shrinkage and ultimately shedding of membrane-delimited cell fragments known as apoptotic bodies (Wyllie, 1997).

In adult mammals, programmed cell death plays an essential role in tissue homeostasis, offsetting new cell production with cell death in all self-renewing tissues. Roughly one million cells commit suicide every second in the adult human, thus making room for new cells produced in tissues such as the bone marrow, skin and gut on a daily basis. So massive is the flux of cell birth and death through our bodies that the average person will produce and in parallel eradicate a mass of cells equivalent to his or her entire body weight each year. Consequently, defects in the cell death machinery which prevent the programmed turnover of cells can result in cell accumulation, thereby imparting a selective growth advantage to neoplastic cells without necessarily involving concomitant defects in the cell division cycle (Reed, 1999).

Defects in the pathways responsible for programmed cell death also play important roles in multiple aspects of tumour cell biology, besides cell accumulation. For example, because cancer requires the accumulation within a single clone of multiple genetic lesions, enhanced cell longevity as a result of defective apoptosis may indirectly promote cancer. The genetic instability that characterizes many cancers is also indirectly assisted by defects in apoptosis, since errors in DNA management typically disrupt cell cycle checkpoints, triggering a cell suicide response. Growth factor and hormone independence, hallmarks of many advanced cancers, can also be attributed in part to alternations in the cell death machinery, which permit cancer cells to thrive in the absence of these factors that cells normally require for maintenance of their survival. Metastasis is also assisted by defects in apoptosis, permitting tumour cells to survive in a suspended state (such as during circulation through the blood or lymph), whereas normal epithelial cells undergo apoptosis when detached from extracellular matrix and thus are confined to predefined locations in the body. Immune surveillance mechanisms are also thwarted by defects in apoptosis, since cytolytic T-cells and natural killer (NK) cells depend on components of cell death machinery to kill target cells. Finally, defects in apoptosis pathways contribute to resistance of cancer cells to chemotherapy and radiation, raising the threshold of drug- or radiation-induced damage necessary to trigger a cell suicide response (**Table 1**).

The core cell death machinery consists of families of genes and their encoded proteins, many of which are conserved throughout metazoan evolution (Metzstein *et al.*, 1998) . Several protein domains which are entirely or nearly unique to apoptosis pathways are found within apoptosis-suppressing and -inducing proteins, including caspase protease domains, caspase-recruitment domains (CARDs), death domains (DDs), death effector domains (DEDs), caspase-inducible DNA endonuclease (CIDE) domains,

Table 1 Relevance of defective apoptosis to cancer: *pathogenesis, progression, therapy resistance*

Cell accumulation (cell death < cell division)
Longevity (accumulation of genetic lesions)
Genomic instability (tolerate DNA mistakes)
Immune surveillance (resistance to immune attack)
Growth factor/hormone independence (survival without paracrine/endocrine growth factors)
Angiogenesis (resistance to hypoxia; hypoglycaemia)
Metastasis (survival without attachment)
Chemoresistance/radioresistance (increased threshold for cell death)

Table 2 Protein domains associated with apoptosis

Caspase (catalytic) domains
BIR domains (IAPs)
Caspase-recruitment domains (CARDs)
Death domains (DDs)
Death effector domains (DEDs)
Bcl-2 homology (BH) domains
CIDE domains

BIR domains and Bcl-2 homology (BH) domains (**Table 2**) (Reed, 2000). Three-dimensional structures have been obtained for at least one of each of these domains, providing the foundation of a clearer understanding of the molecular mechanisms of apoptosis regulation and, in some cases, ideas for how one might modulate apoptosis proteins with therapeutic intent (Fesik, 2000).

APOPTOSIS IS CAUSED BY PROTEASES

The biochemical events responsible for apoptosis can be linked to the activation in cells of a family of cysteine proteases, known as caspases, which cleave specific target proteins in cells at aspartic acid residues. It is these proteolytic cleavage events which directly or indirectly explain the morphological changes that we recognize as 'apoptosis.' As many as 14 caspase-encoding genes have been identified in humans and mice. Homologous genes are also found in lower organisms within the animal kingdom, such as the nematode *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*.

Caspases are initially synthesized as inactive zymogens in cells. How, then, are caspases activated? The answer is by proteolytic cleavage, occurring at conserved aspartyl residues, thus generating the large (~20 kDa) and small (~10 kDa) subunits that comprise the catalytic component of these enzymes. Structures of several caspases have been solved by X-ray crystallography, revealing a heterotetrameric assembly of two large and two small subunits, with two active sites per molecule. The observations that caspase zymogen activation involves cleavage at aspartyl residues and that active caspases also cut proteins at aspartyl residues have obvious implications. Namely, pro-caspases can become cleaved and activated either as a result of cleavage by other active caspases or through 'autoproteolytic' mechanisms which will be discussed below. Thus, once some caspase activation has occurred, the process can spread to other caspases through amplification steps in which one active caspase molecule cleaves and activates multiple caspase zymogens, as well as cascades of proteolytic activation of caspases through sequential stepwise mechanisms (Salvesen and Dixit, 1997; Thornberry and Lazebnik, 1998).

At least three mechanisms have been proposed for explaining how the 'first' caspase becomes activated: (1) the induced-proximity model; (2) the 'safety-catch' mechanism; and (3) introduction of an exogenous protease (**Figure 1**). In the induced-proximity model (**Figure 1**), two or more caspase zymogens are clustered together as a result of protein interactions, thus bringing them into close association. Because the zymogen forms of caspases are not entirely without proteolytic activity (some possessing about 1% of the activity of the processed 'active' enzymes), the close proximity allows these weakly active zymogens to trans-process each other, thereby generating the autonomously active proteases. This clustering of caspase zymogens typically is made possible by the presence in certain pro-caspases of an N-terminal prodomain which is located proximal to the portions of the pro-protein that give rise to the p20 (large) and p10 (small) catalytic subunits. These N-terminal prodomains serve as protein-interaction motifs that permit self-association of zymogens or that mediate interactions with adapter proteins which assemble multiprotein zymogen-activation complexes, sometimes referred to as 'apoptosomes' (Salvesen and Dixit, 1999).

Not all caspases possess significant N-terminal prodomains, thus creating a hierarchy of proteolytic cascades in which upstream caspases that possess large prodomains become activated as a result of protein interactions, and these initiator caspases then cleave and activate downstream effector caspases which possess only short N-terminal pro-peptides that are incapable for mediating protein interactions. This hierarchy of caspase proteolytic cascades is also supported by comparisons of the substrate preferences of upstream initiator and downstream target caspases, using combinatorial peptide library screening and other techniques. These studies have revealed the presence of tetrapeptide substrate cleavage sites preferred by upstream caspases within the sequences of the proforms of the downstream caspases, at sites corresponding to the authentic *in vivo* cleavage sites that separate the large and small subunits. The tetrapeptide cleavage sites preferred by the downstream caspases, in turn, are often found in several of the target proteins which are known to become cleaved at aspartyl residues during apoptosis, including

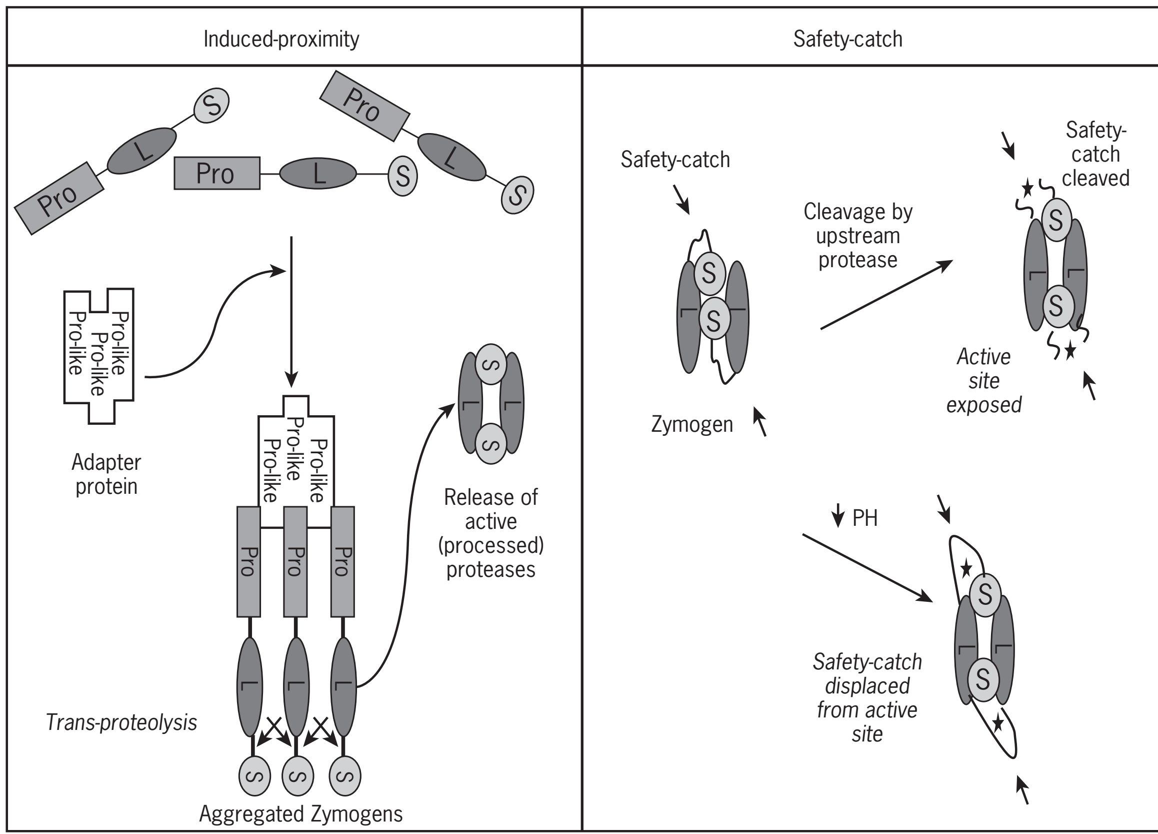


Figure 1 Mechanisms of caspase activation. Two of the proposed mechanisms of caspase activation are depicted, the induced-proximity model and the safety-catch model . In the induced-proximity model, pro-caspase zymogens are present as monomers that are single polypeptide chains with N-terminal prodomains (Pro) followed by the regions corresponding to the large (L) and small (S) catalytic subunits. Bringing these zymogens into close proximity, either by interactions with adapter proteins which typically share sequence and structural similarity with the prodomain regions of pro-caspases (shown) or by overexpression of pro-caspases resulting in self-aggregation (not-shown), allows these proenzymes to trans-process each other. This trans-processing is possible because the zymogens possess weak protease activity. Once proteolytically processed, separating the large and small subunit fragments and (often) removing the N-terminal prodomain, the active protease assembles, consisting of heterotetramer with two L and two S subunits. In the safety-catch mechanism, an autoinhibitory loop in the zymogen is envisioned as occupying or blocking the active site of the enzyme. Typically, cleavage at or near this autoinhibitory loop located between the L and S subunits exposes the active site (indicated by a star), creating the active protease (top). Alternatively, autorepression may be relieved by lowering pH, protonating acidic residues in the autoinhibitory loop and causing the loop to leave the active site of the enzyme (bottom).

the actin-regulatory protein gelsolvin, the chromatin-regulatory protein poly-ADP ribosyl polymerase (PARP), the endonuclease inhibitor ICAD and its relatives of the caspase-inducible DNA endonuclease (CIDE)-family proteins, nuclear lamins, certain protein kinases and various other substrate proteins (Thornberry and Lazebnik, 1998; Cryns and Yuan, 1999).

The safety-catch mechanism envisions an auto-repressing loop in the caspase zymogen which occupies the active site of the protease and maintains it in a highly constrained inactive state (**Figure 1**). Activation then is speculated to involve displacement of the auto-repressive loop from the active site, allowing the caspase possibly to auto-process and thereby auto-activate itself. This mechanism has been proposed for certain downstream effector caspases, such as caspase-3 and -7, which contain three adjacent acidic residues (aspartic or glutamic acid) in the candidate auto-repression loop. Neutralizing the negative charge of these acidic residues by lowering the pH or by site-directed mutagenesis (with conversion to uncharged alanine) lowers the barrier to caspase activation. It remains to be determined whether the proposed safety-catch mechanism can be extended to other caspases.

Finally, an example of caspase activation resulting from introduction of an exogenous protease has come from investigations of mechanisms of target cell killing by cytolytic T cells (Lowin *et al*., 1995). Here, perforin-mediated channels are used to inject granzyme B, a serine protease

with specificity (like caspases) for cleavage of substrates at aspartyl residues. Granzyme B is capable of directly cleaving and activating most caspase-family zymogens. Interestingly, however, intracellular trafficking of granzyme B appears to be under regulation by cellular factors that may limit its interactions with caspases in some scenarios.

Inactivation of caspases can occur through several mechanisms in cancers (**Table 3**). For example, structural alterations in caspase-family genes, such as deletions and loss of function mutations, have been documented in tumour specimens or cancer cell lines, including initiator and effector caspases, suggesting that inactivation or elimination of caspase-family genes represents one mechanism by which malignant cells may escape apoptotic elimination. In addition, some caspase-family genes, including caspase-2 and -9 in mammals, are capable of producing shorter protein isoforms through alternative mRNA splicing that function as trans-dominant inhibitors of their full-length counterparts, possibly by forming heterodimers composed of full-length and truncated pro-caspases or by competing for binding to upstream activators. Although little work has been performed to date, it is intriguing to speculate that tumours may overexpress the trans-dominant inhibitory isoforms of some caspases as an additional mechanism for subverting apoptosis. Protein phosphoryation is another mechanism capable of directly suppressing caspases in tumours. The only reported example of this so far is human caspase-9, which can become phosphorylated by the kinase Akt (protein kinase-B), thereby suppressing the active caspase-9 enzyme as well as suppressing activation of the pro-caspase-9 zymogen. Interestingly, elevated levels of Akt activity are observed in many tumours as a result of (1) amplification of the Akt-family genes, (2) increased signalling by upstream protein tyrosine kinases and Ras-family oncoproteins, (3) aberrant production of Akt co-activators and (4) inactivation of PTEN—the product of a tumour-suppressor gene which normally suppresses Akt activation by dephosphorylating second messenger polyphosphoinositol lipids (Datta *et al*., 1999). Theoretically, other post-translational modifications of caspases could also participate in their inactivation in cancers, including *S*-nitrosylation or glutathionylation of the active site cysteine.

Table 3 Mechanisms of caspase inactivation and suppression in cancers

Mechanism	Examples
Mutations in caspase genes	Caspase-3, -5, -8
Expression of dominant-negative caspase isoforms by alternative mRNA splicing	Caspase-2, -9
Phosphorylation	Caspase-9
Nitrosylation	Caspase-3
Overexpression of caspase-inhibiting IAPs	XIAP, cIAP-1, Livin

INHIBITOR OF APOPTOSIS (IAP) FAMILY PROTEINS FUNCTION AS CASPASE INHIBITORS

All protease networks studied to date include inhibitors which control flux through proteolytic cascades and which establish thresholds for protease activation which must be surpassed to trigger biological processes. The inhibitors of apoptosis proteins (IAPs) represent a family of anti-apoptotic proteins conserved throughout metazoan evolution that appear to serve this role. IAPs were first identified in the genomes of baculoviruses, where they suppress apoptosis induced by viral infection of host insect cells. Subsequently, cellular homologues of the baculovirus IAPs were discovered in humans, mice, flies and other animal species (Deveraux and Reed, 1999; Miller, 1999).

Membership in the IAP family requires two things: (1) an ability of the protein to suppress apoptosis, at least when overexpressed in cells, and (2) the presence of at least one copy of a conserved domain known as a BIR domain ('baculovirus IAP repeat'). The BIR domain represents a zinc-binding fold. One to three copies of the BIR domain are found in IAP family proteins, sometimes in association with other domains such as RING fingers, putative nucleotide-binding domains and caspase recruitment domains (CARDs) (described below).

A single BIR domain can be necessary and sufficient for inhibition of certain caspases by human IAP-family proteins. For example, the second of three BIR domains found in the human XIAP protein (where X indicates that the gene maps to the human X-chromosome) directly binds to and potently inhibits ($K_i < 1\,\mathrm{nmol\,L^{-1}}$) caspase-3 and -7. Thus, the concept has emerged that BIR domains represent caspase-inhibitory structures that bind active caspases and suppress them within cells. Interestingly, BIR domains do not inhibit all caspases, but rather exhibit clear selectivity. The second BIR domain of XIAP (in combination with adjacent residues), for instance, inhibits the downstream effector proteases, caspase-3 and -7, but does not bind or suppress the upstream initiator proteases, caspase-1, -8, -9 or -10. In contrast, the third BIR domain of XIAP suppresses caspase-9 but not other caspases tested. Although having caspase-inhibitory activity, at least some IAP family proteins may also participate in other processes, particularly signal transduction pathways involved in regulating kinases responsible ultimately for NF-κB induction and JNK activation. Contrasting the roles of IAPs as caspase inhibitors versus signal transduction modulators remains an active area of research.

Altered expression of IAP-family genes has been documented in cancers (LaCasse *et al*., 1998). For example, the *cIAP-2* gene becomes involved in chromosomal translocations in certain types of lymphomas, resulting in deregulation of this gene. The *cIAP-2* gene also is a target of the transcription factor NF-κB, a member of the

Rel family of oncoproteins which suppresses apoptosis. Moreover, a member of the IAP-family called survivin is inappropriately overexpressed in the majority of human cancers. Survivin is the smallest of the IAPs identified thus far, containing a single BIR domain in humans and mice. This particular IAP-family member is particularly intriguing because it appears to have a dual role in apoptosis suppression and in cell division. The survivin protein is physically associated with the mitotic spindle apparatus during M-phase and beyond into anaphase, evidently playing an essential role in cytokinesis. Homologues of survivin are found in yeast and in *C. elegans*, performing essential functions in cell division but lacking any clear role in cell death control. Thus, it is speculated that survivin's evolutionarily conserved core function is related to cytokinesis, and that an apoptosis checkpoint function for this protein may have evolved later in higher organisms as a way of linking defects in the late stages of cell division to an apoptotic response. Circumstantial evidence implicates survivin in caspase suppression in mammalian cells analogous to other IAPs, but a direct role for this BIR domain containing protein in caspase inhibition remains equivocal. Furthermore, not all BIR-containing proteins are involved in apoptosis regulation, indicating that this zinc-binding fold can serve alternative functions, despite a conserved arrangement of histidine and cysteine residues. Most likely, BIRs are protein-interaction domains that bind and inhibit caspases in some circumstances but not others.

Regardless of the mechanism, overexpression of survivin clearly reduces apoptosis in response to a variety of apoptogenic stimuli whereas antisense-mediated reductions in survivin expression or gene transfer-mediated expression of dominant-negative mutants of survivin sensitize tumour cells in culture to apoptosis induction by anticancer drugs. These findings, coupled with the observation that survivin is rarely expressed in normal adult tissues whereas survivin mRNA and protein levels are markedly increased in most cancers, have elevated the status of survivin as a potential drug-discovery target for cancer therapy (Altieri *et al*., 1999).

RECEPTOR-MEDIATED MECHANISMS OF CASPASE ACTIVATION

Several mechanisms of caspase activation have been elucidated, including the coupling of ligand binding at cell-surface receptors to caspase activation. Moreover, defects in at least some of these mechanisms have been uncovered in cancers.

The tumour necrosis factor (TNF) family of cytokine receptors includes at least six members that trigger apoptosis and which share a conserved protein-interaction domain in their cytosolic region which is essential for their cytotoxic activity (Ashkenazi and Dixit, 1998; Wallach *et al*., 1999). These TNF-family receptors include TNFR1, Fas (CD95), DR3 (weasle; tweak), DR4 (trail receptor-1; Apo2), DR5 (trail receptor-2) and DR6 (**Figure 2**). The conserved cytosolic domain is known as the 'death domain' (DD) and is comprised of a six α-helical bundle which forms trimers and possibly higher-order oligomers in response to ligation of the extracellular domains or simply when these receptors are overexpressed in the absence of their cognate ligands. Hereditary loss-of-function mutations or trans-dominant inhibitory mutations in the DD of Fas (CD95) have been associated with a lymphoproliferative disorder and autoimmunity in the *lpr*/*lpr* strain of mice and in humans with autoimmune lymphoproliferative syndrome (ALPS). Similarly, somatic mutations in the death domain of Fas have been detected in certain human neoplasms, including myelomas, lymphomas, leukaemias and carcinomas of the lung and bladder.

The DDs of TNF-family death receptors bind adaptor proteins which interact, in turn, with the N-terminal domains of specific initiator caspases (**Figure 2**). The assembly of these multiprotein complexes is thought then to activate the associated pro-caspases by the induced-proximity model described above. Fadd (Mort1) represents one such adaptor protein. This protein contains a DD which binds directly to the DD of Fas, as well as a protein interaction module known as a death effector domain (DED). The DED is similar to the DD in structure, comprised of six α-helices, but constitutes a separate domain family which can be differentiated by sequence homology. Two of the known human caspases contain DEDs in their N-terminal prodomain regions, pro-caspase-8 and -10. These pro-caspases each contain two DEDs upstream of the catalytic segments and are both capable of binding directly to the DED of Fadd. Studies of cells derived from caspase-8 knockout mice have revealed an obligatory role for this initiator caspase in apoptosis induction by Fas and TNFR1. Similarly, gene ablation studies indicate an essential role for Fadd in apoptosis induction by these TNF-family death receptors.

Another example of a caspase-activating adapter protein that associates with TNF-family receptor complexes is Raidd (Cradd). The Raidd protein contains a DD and another protein interaction module known as a CARD domain. The CARD domain of Raidd specifically binds a homologous domain found in the N-terminal prodomain region of pro-caspase-2. Structures of these CARD domains have been solved, revealing again a characteristic fold comprised of six α-helices and suggesting that complementary patches of acidic, basic and hydrophobic residues on the surfaces of these domains account for the selectivity of their interactions with each other but not with other CARD-family proteins. In most cases, interactions of Fadd and Raidd with TNF-family receptors is mediated by an intermediate DD-family protein, Tradd, via associations of their DDs.

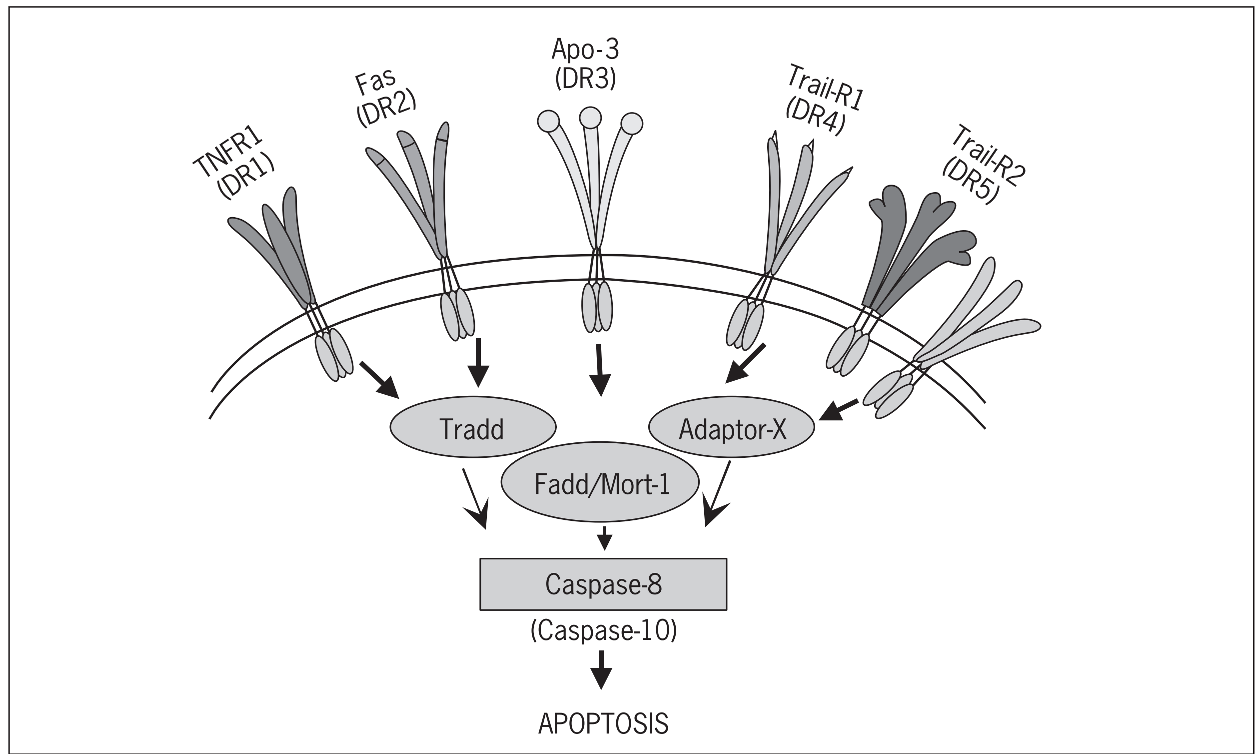


Figure 2 Caspase activation by TNF-family receptors. The currently known human members of the tumour necrosis factor (TNF) receptor family which contain cytosolic death domains (DDs) are depicted. The intracellular DDs of these receptors bind DD-containing adaptor proteins, such as Tradd, Fadd and others yet to be identified. These adaptor proteins then interact with initiator caspases which contain homologous protein interaction domains. The TNF family of receptors provided the first example of a receptor system which transduces signals into cells via proteolysis, thus representing a milestone in cell biology research.

Additional proteins which can participate indirectly in recruitment of pro-caspases to receptors include RIP and RIP-2 (Cardiac) which contain a conserved kinase domain of unknown functional significance in association with either a DD or CARD, respectively. Interestingly, whereas the CARD domain of RIP-2 binds the N-terminal CARD domain of pro-caspase-1, other portions of this protein reportedly interact with TRAF6, an adapter protein that associates with certain TNF-family receptors and Toll-family receptors (**Figure 3**). It should be noted, however, that caspase-1 and its close relatives, human caspase-4 and -5, may be only tangentially involved in apoptosis, and instead are principally involved in proteolytic processing of proinflammatory cytokine precursors, particularly pro-interleukin-1β and pro-IL-18. Caspase-1 knockout mice, for example, are grossly normal and exhibit only very modest defects in apoptosis induction, while manifesting extreme resistance to endotoxin-induced sepsis and displaying profound deficiencies in production of IL-1 and IL-18 *in vivo* (Reed, 1998a).

TNF-family death receptors play important roles in immune system interactions with tumours. One of the principal weapons used by cytolytic T cells for killing tumour targets, for example, is Fas-ligand (FasL). Consequently, defects in Fas-induced apoptosis can contribute to tumour avoidance of immune surveillance mechanisms. Moreover, upon achieving a Fas-resistant state, it has been shown that some tumours can then tolerate expressing FasL on their surfaces, thus using this death ligand as a weapon to kill neighbouring normal cells as well as activated lymphocytes.

Multiple mechanisms of tumour resistance to Fas and other TNF-family death ligands have been elucidated, including prereceptor, receptor and postreceptor defects (**Table 4**) (Tschopp *et al.*, 1998, 1999; Ashkenazi and Dixit, 1999). Prereceptor defects include the production of soluble 'decoy' receptors or fragments of receptors that compete with the transmembrane receptors for ligand binding. For instance, a soluble form of Fas can be produced by alternative mRNA splicing in which the exon encoding the transmembrane anchoring domain is skipped, resulting in production of a secreted version of this receptor. In the case of Trail (Apo2L), three genes have been identified which produce 'decoy' receptors that can compete with the death receptors, DR4 (TrailR1) and DR5 (TrailR1), for binding to Trail, thereby sparing cells from the apoptotic effects of this death ligand. Proteolytic removal of the extracellular domain of TNF-family receptor is also a possibility. Another potential resistance mechanism is found in the SODD (silencer of death domains) protein. SODD (also

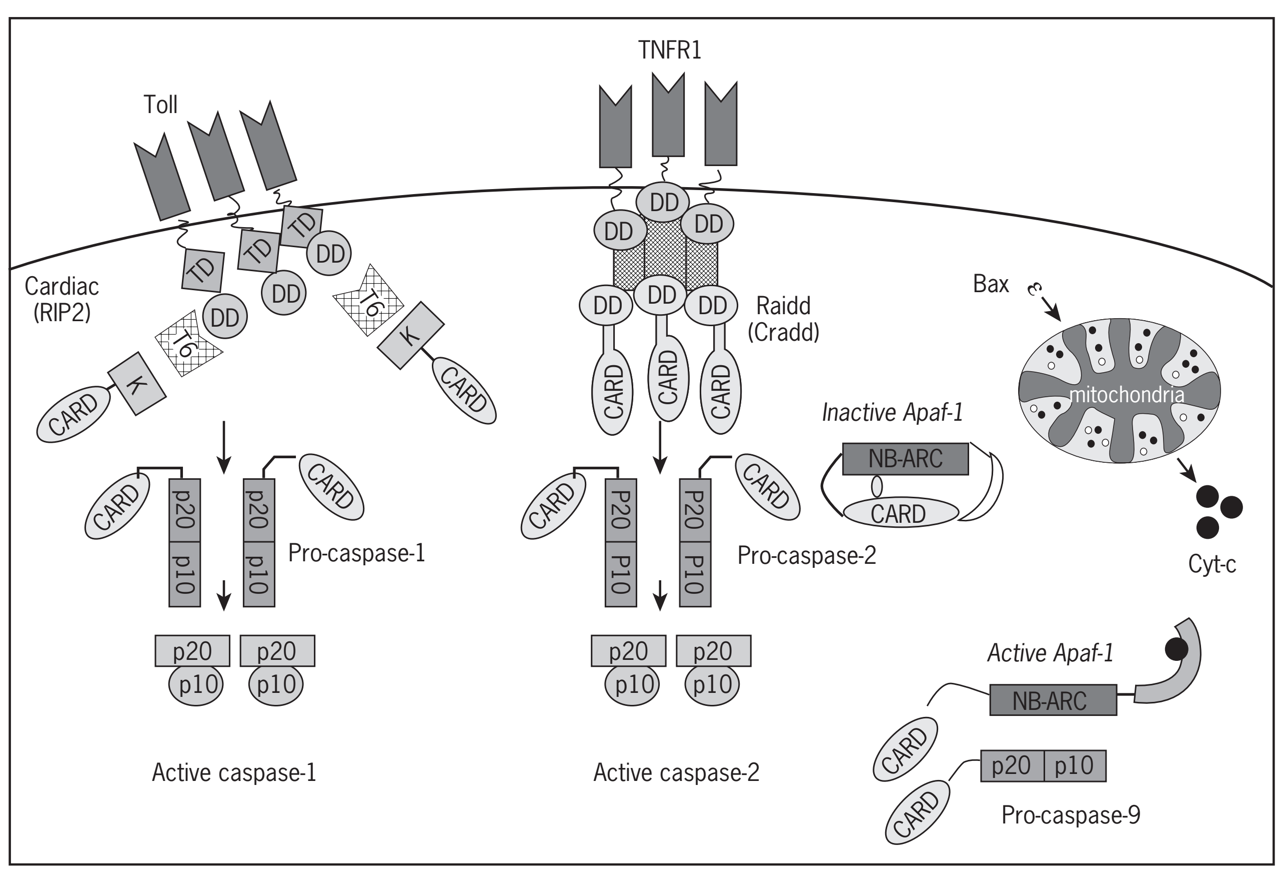


Figure 3 CARD domains and caspase activation. Examples of involvement of CARD-domain proteins in activation of CARD-containing caspases are presented, including examples from the Toll-receptor family (left), tumour necrosis factor receptor-1 (TNFR-1) (middle) and mitochondrial/cytochrome *c* (cyt-c) pathway (right). For each example, the target caspase contains an N-terminal prodomain consisting of a CARD. These CARD prodomains interact with homologous CARDs in various adapter proteins, such as Cardiac (RIP2; Rick) (left) which contains both a CARD and a kinase (K) domain of unknown significance, Raidd (Cradd) (middle) which contains both a CARD domain and a DD, the latter of which binds homologous DD-containing adapter proteins which associate with the TNF-receptor, and Apaf-1 (right), which contains a CARD, followed by a nucleotide-binding oligomerization domain (NB) and then WD repeats which mediates interactions with cyt-c. Release of cyt-c from mitochondria is required for converting Apaf-1 from a latent (inactive) conformation, resulting in activation of Apaf-1 so that it can bind pro-caspase-9 (see **Figure 5** for more details about Apaf-1).

known as BAG4) contains a domain that binds the DDs of TNFR1 and DR3, as well as a conserved Hsp70/Hsc70-binding domain. When overexpressed SODD prevents spontaneous aggregation of DDs, presumably by recruiting Hsp70/Hsc70-family molecular chaperones to these receptors and inducing conformational changes that prevent them from oligomerizing and signalling in the absence of ligand.

Table 4 Mechanisms for interfering with TNF-family death receptors

Mutations in death receptor or ligand genes
Decoy or soluble receptors
SODD (BAG4) suppression of receptor oligomerization
Antiapoptotic DED-family proteins (cFlip; BAR; Bap31)
NF-κB-mediated upregulation of caspase inhibitory proteins

Several postreceptor defects in death receptor signalling have also been uncovered in cancers. For example, FAP-1 is a Fas-binding protein phosphatase that, when overexpressed, can suppress Fas-induced apoptosis through an unidentified mechanism. Some tumours appear to overexpress FAP-1. Moreover, peptidyl antagonists that bind PDZ domains in FAP1 and that block interactions of FAP-1 with Fas have been shown to restore Fas sensitivity to Fas-resistant colon cancer cell lines *in vitro*. Additional antagonists of death receptor signalling include the Flip family of DED-containing proteins. These antiapoptotic proteins contain DEDs that permit them to bind either adapter proteins such as Fadd- or DED-containing pro-caspases such as pro-caspase-8 and -10. Overexpression of cellular Flip (also known as I-Flice, Cash, Casper, Usurpin, Mrit, Rick) or of its viral homologues can prevent assembly

of death receptor signalling complexes and interfere with caspase activation. Overexpression of c-Flip has been documented in some tumours, and has been associated in animal models with resistance to immune-mediated suppression of cancer. In addition to cytosolic c-Flip and related proteins, membrane-anchored proteins which contain DEDs capable of binding pro-caspase-8 and/or -10 have been identified. Proteins such as BAR and Bap31 localize to internal membranes, primarily endoplasmic reticulum and/or mitochondria, and can modulate apoptosis signalling by TNF-family death receptors. Finally, activation of the transcription factor NF-κB has been associated with resistance to apoptosis induction by several TNF-family death receptors. NF-κB transcriptionally upregulates the expression of several antiapoptotic proteins, including one or more of the IAP-family proteins, resulting in apoptosis resistance. This observation is directly relevant to mechanisms of TNF-induced apoptosis in that several TNF-family receptors recruit proteins that activate caspases as described above, thus promoting apoptosis, but simultaneously bind other proteins which trigger NF-κB induction, thus preventing apoptosis. The net outcome of engaging TNF-family receptors such as TNFR1, DR3 and DR6 is difficult to predict and can be extremely cell-type and cell-context dependent. Taken together, it is clear that malignant cells have many options for developing mechanisms that thwart apoptosis induction by TNF-family death receptors and ligands.

MITOCHONDRIAL PATHWAYS FOR CASPASE ACTIVATION

In addition to receptor-mediated mechanisms for coupling caspase activation to ligation of specific cell-surface receptors, a pathway has been elucidated which links mitochondrial damage to a mechanism for triggering caspase activation. This mitochondrial pathway for caspase activation is engaged in response to growth factor deprivation, genotoxic injury, hypoxia and many other insults, and is commonly referred to as the 'intrinsic' pathway, in contrast to the receptor-mediated caspase activation mechanisms which are sometimes referred to as the 'extrinsic' pathway (**Figure 4**) (Green and Reed, 1998; Kroemer and Reed, 2000).

What allows mitochondria to trigger caspase activation?—release of cytochrome *c* from these organelles into the cytosol. Cytochrome *c* is best known for its role in electron-chain transport, where it transfers electrons from complex III to IV in the respiratory chain. However, following exposure of cells to many apoptotic stimuli, the outer membranes of mitochondria undergo permeability changes that permit cytochrome *c* and other proteins normally sequestered in the space between the inner and outer membranes of these organelles to leak out and enter the cytosol. Once in the cytosol, cytochrome *c* binds a caspase-activating protein, called Apaf-1 (apoptotic protease activating factor-1). Apaf-1 is normally present in the cytosol in an inactive (latent) state in mammals and *Drosophila*. This protein consists of an N-terminal CARD domain, followed by a nucleotide-binding domain, ending with multiple WD repeat domains. The human Apaf-1 protein has been shown to oligomerize, apparently forming octamers, upon binding cytochrome *c*. This oligomerization is mediated by the nucleotide-binding domain and requires adenine triphosphate (ATP) or deoxy-ATP. Cytochrome *c* appears to interact with the WD repeat domains and nucleotide-binding domain, promoting a conformational change that results in oligomerization. Once oligomerized, activated oligomerized Apaf-1 molecules bind via their CARD domains to pro-caspase-9, a member of the caspase family which possesses an N-terminal CARD prodomain that has been shown by X-ray crystallography to possess a complementary interaction surface for association with the CARD of Apaf-1 (the interactions in this case are primarily electrostatic, involving associations of a surface patch of acidic residues on the CARD of Apaf-1 with basic residues on the CARD of pro-caspase-9). Activation of pro-caspase-9 then is achieved by the affinity-approximation model. However, rather than releasing processed caspase-9 to cleave and activate additional downstream effector caspases, instead caspase-9 must remain bound to Apaf-1 for maintaining its optimal protease activity. The 'free' processed caspase-9 molecule has only weak protease activity in the absence of Apaf-1, suggesting that the Apaf-1:caspase-9 complex represents the holoenzyme (**Figure 5**). Thus, the next caspase in the cascade, pro-caspase-3, is recruited to the Apaf-1:caspase-9 holoenzyme, where it becomes activated by cleavage at a conserved aspartyl residues separating the large and small subunits of the protease. Processed caspase-3 is then released from the Apaf-1:caspase-9 complex and autocatalytically cleaves and removes its own prodomain, which consists of a short ($\sim$20 amino-acid) segment, thereby generating the mature enzyme. Active caspase-3 then cleaves and activates additional pro-caspases, thus amplifying the proteolytic cascade. This and other downstream caspases also cleave a wide variety of specific substrates that ultimately commit the cell to an apoptotic demise, including CIDE-family proteins which control the activity of an endonuclease which appears to be largely responsible for the fragmentation of the nuclear DNA into oligo-nucleosomal length fragments—a hallmark (although non-obligatory event) of apoptosis.

Mechanisms for regulating the cytochrome *c*-dependent activation of Apaf-1 may exist, thus providing a means of fine tuning the coupling of mitochondria to caspases (Reed and Paternostro, 1999). For example, the efficiency of cytochrome *c*-mediated activation of caspases has been shown to be poor under conditions of isotonic salt $\sim$150 mmol L^{-1} (KCl) and neutral pH (7.4). Release of another protein from mitochondria, Smac (second mitochondrial activator

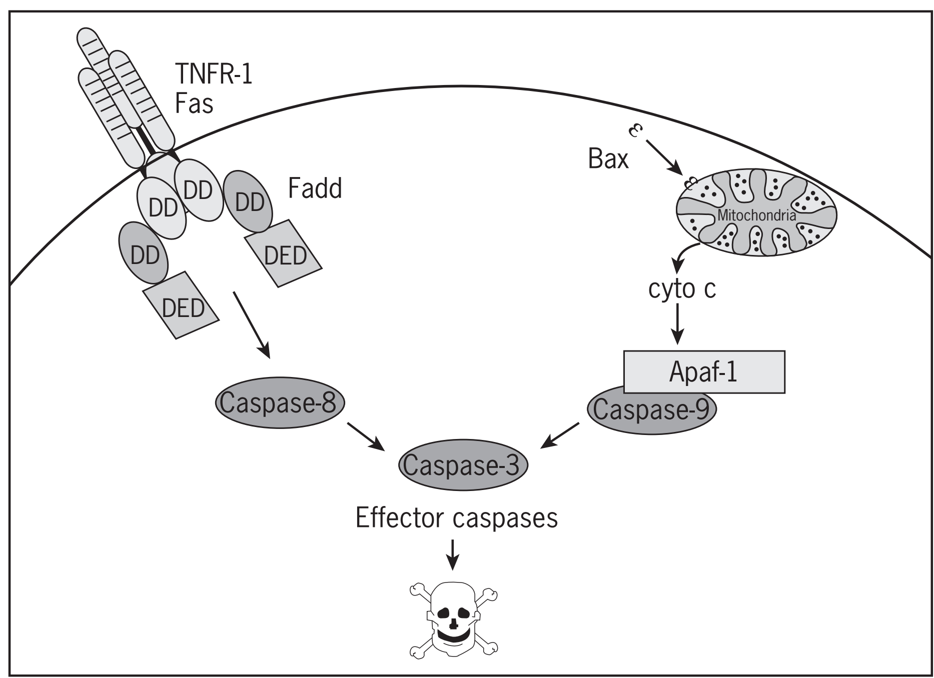


Figure 4 Intrinsic and extrinsic pathways for apoptosis. Two of the major caspase-activation pathways are depicted. The 'extrinsic' pathway is activated by extracellular ligands that bind death receptors such as TNFR-1 and Fas. These death receptors bind, in turn, to adaptor proteins such as Fadd (Mort1) which then bind the initiator pro-caspase, caspase-8. The 'intrinsic' pathway is activated by intracellular signals which induce cyt-c release from mitochondria. Cyt-c binding to the caspase-activator, Apaf-1, results in binding to and activation of pro-caspase-9. Gene ablation studies in mice have shown that pro-caspase-8 and pro-caspase-9 are absolutely required for apoptosis induction by TNF-family death receptors and cyt-c, respectively. Similarly, gene knockout studies in mice indicate that Fadd (Mort1) is critical for caspase activation by most (although perhaps not all) TNF-family death receptors and that Apaf-1 is absolutely required for caspase-activation by cyt-c. Many downstream effector caspases may be cleaved and activated by active caspase-8 and -9, with caspase-3 representing the best documented direct substrate of these upstream initiator caspases.

of caspases), can overcome the dependence on hypotonic salt. Furthermore, some apoptotic stimuli induce mitochondria to extrude protons (H^+), thus acidifying the cytosol and promoting cytochrome-mediated caspase activation. It may be of significance consequently that many growth and survival factors are known to induce cytosol alkalinization as a result of effects on the Na^+/H^+ antiporter, perhaps accounting for some of their antiapoptotic effects. Overproduction of Hsp70 also may interfere with cytochrome *c*-mediated activation of caspases, a finding of possible relevance to cancer given evidence of elevated Hsp70 levels in several kinds of cancer.

Also of relevance to activation of pro-caspase-9, the apical caspase in the mitochondrial pathway, at least two other proteins have been reported to bind pro-caspase-9 and to induce apoptosis through what appears to be cytochrome *c*-independent mechanisms, including an Apaf-1 homologue (CARD4/Nod) and Bcl-10 (hE10; CIPER). Both CARD4 (Nod) and Bcl-10 (hE10; CIPER) possess N-terminal CARD domains that interact with the CARD of pro-caspase-9. Interestingly, chromosomal translocations and somatic point mutations have been described in the *bcl-10* (hE10; CIPER) gene in lymphomas and solid tumours, respectively, which convert it from a pro-apoptotic to an anti-apoptotic protein.

Besides these caspase-9-binding proteins, another way of potentially modulating signalling through caspase-9-dependent pathways involves sequestration of this pro-protease in organellar compartments where it cannot interact with cytosolic proteins such as Apaf-1, CARD4 (Nod) and Bcl-10 (hE10; CIPER). Specifically, pro-caspase-9 has been shown to be stored in the intermembrane space of mitochondria, along with cytochrome *c*, in some types of terminally differentiated cells. Thus, changes in mitochondrial membrane permeability are required to release pro-caspase-9 into the cytosol where its interacting proteins are located, at least in some cell types. Finally, caspase-9 in humans is a substrate of the kinase Akt, with phosphorylation inhibiting its protease activity through undetermined mechanisms.

Loss of expression of Apaf-1 has been described in melanomas, due to gene hypermethylation. In this regard, in cancers, *in vitro* transformation studies using cells from Apaf-1 and caspase-9 knockout mice suggest that these

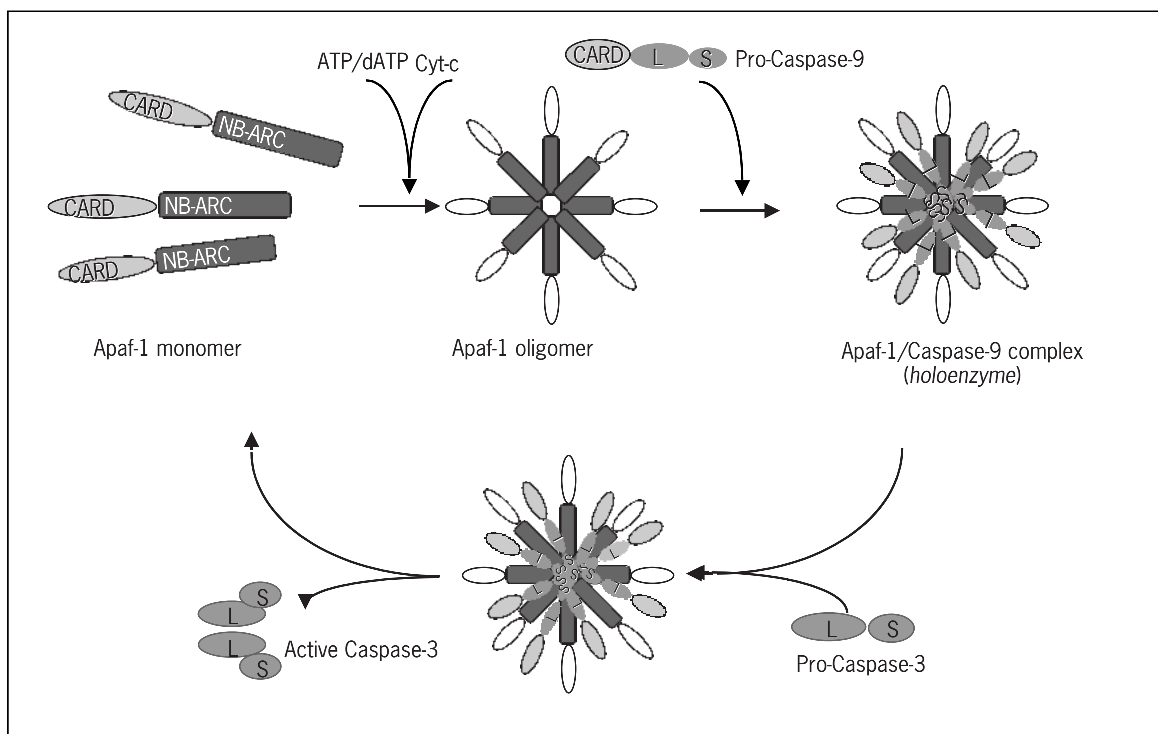


Figure 5 Dynamics of Apaf-1 'Apoptosome.' The sequence of events involved in cyt-c-mediated activation of Apaf-1 and pro-caspase-activation are depicted. Cyt-c plus either ATP or deoxy ATP (dATP) binding to Apaf-1 results in oligomerization mediated by the nucleotide-binding domain (NB-ARC) of this protein. Present estimates suggest that activated Apaf-1 may form an octamer. Oligomerized Apaf-1 binds pro-caspase-9, via a CARD–CARD interaction. Apaf-1-associated pro-caspase-9 molecules are then thought to transprocess each other, cleaving between the large (L) and small (S) subunits. Unlike most other caspases, caspase-9 does not remove its prodomain (CARD) by autoproteolysis. Also, unlike most caspases, after proteolytic processing, caspase-9 must remain associated with its activator Apaf-1 to maintain optimal protease activity. The Apaf-1/caspase-9 complex, therefore, represents the 'holoenzyme' complex. This complex directly recruits pro-caspase-3, which is cleaved by caspase-9, releasing active and processed caspase-3.

pro-apoptotic genes function as tumour suppressors within the p53 pathway. This finding is consistent with other data that suggest that p53 induces apoptosis primarily through a mitochondria-dependent mechanism (intrinsic pathway), although the death receptor (extrinsic pathway) may also be involved in some circumstances, as discussed below.

Bcl-2 FAMILY PROTEINS—REGULATORS OF CYTOCHROME *c* RELEASE AND MORE

A large family of evolutionarily conserved proteins has been identified which regulates the release of cytochrome *c* and other proteins from mitochondria (Adams and Cory, 1998; Reed, 1998b; Gross *et al*., 1999). The founding member of this family is known as Bcl-2, representing an acronym for B-cell lymphoma-2 and reflecting the original discovery of this gene because of its involvement in chromosomal translocations in B-cell lymphomas. In follicular non-Hodgkin lymphomas, the *Bcl-2* gene on the long arm of chromosome 18 at band q21 frequently becomes fused with the immunoglobulin (Ig) heavy-chain gene locus on the long-arm of chromosome 14 at band q32, creating t(14;18) (q32; q21) translocations. Because the Ig gene locus is highly active in B cells, the juxtaposed Bcl-2 gene become transcriptionally deregulated, resulting in continuous production of high levels of *Bcl-2* mRNAs and protein. Bcl-2 is an anti-apoptotic protein that uses a C-terminal membrane-anchoring domain to insert into the membranes of mitochondria as well as some other membranes inside cells. Elevated levels of Bcl-2 prevent cytochrome *c* release from mitochondria following exposure of cells to a wide variety of apoptotic agents and conditions, including growth factor deprivation, oxidants, Ca^{2+} overload, chemotherapeutic drugs and X-irradiation. The mechanism by which Bcl-2 family proteins control cytochrome *c* released from mitochondria remains unknown, although several theories have been advanced and are the subject of several reviews devoted entirely to this important topic (Adams and Cory, 1998; Green and Reed, 1998; Reed, 1998b; Gross *et al*., 1999; Vander Heiden and Thompson, 1999) (**Figure 6**).

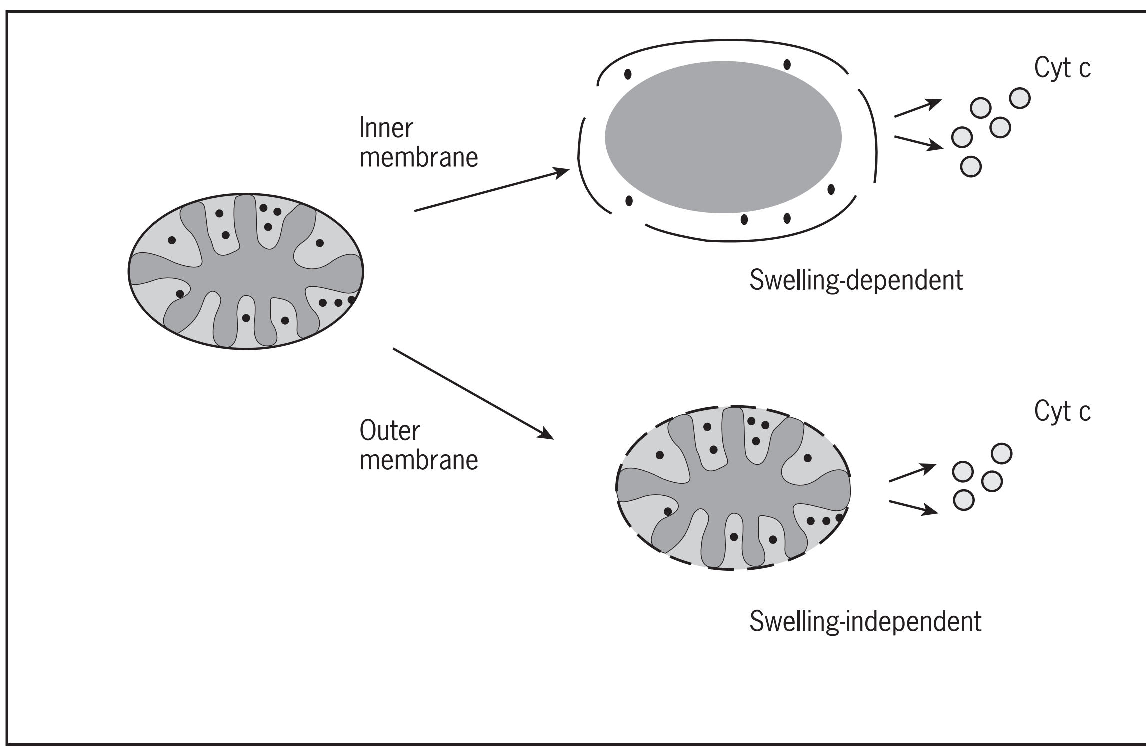


Figure 6 Alternative routes of cytochrome *c* release from mitochondria. Two alternative mechanisms are depicted for inducing release of cyt-c (shown) and other apoptosis-relevant proteins, such as AIF, Smac and intra-mitochondrial caspases (not shown) from the inter-membrane space of mitochondria. The swelling-dependent mechanism (top) involves alterations to the permeability of the inner membrane, causing osmotic disequilibrium and swelling of the matrix space. Because the surface area of the inner membrane, with its folded cristae, is greater than the area of the outer membrane, swelling results in rupture of the outer membrane and release of proteins stored in the inter-membrane space. In the swelling-independent model (bottom), a selective change in outer-membrane permeability occurs, allowing release of cyt-c and other proteins from the inter-membrane space, while inner-membrane integrity and volume homeostasis are preserved.

The Bcl-2 family consists of both anti-apoptotic members, which include in humans Bcl-X_L, Bcl-W, Bfl-1, Bcl-B and Mcl-1, as well as pro-apoptotic members, which include in humans Bax, Bak, Bok, Bad, Bid, Hrk, Bik, Bim, Nip3 and Nix. Homologues of some of these proteins are also found in lower organisms, including *Caenorhabditis elegans* and *Drosophila melanogaster*, where they often play important roles in developmental programmed cell death. Evolutionary conservation of function has been documented by cross-species gene transfer experiments, implying commonality in their mechanisms of action.

Antiapoptotic Bcl-2 family members may possess multiple mechanisms for suppressing cell death. In general, however, these mechanisms can be simplified into two general categories. First, based on determination of their three-dimensional structures or computer modelling predictions of structures, antiapoptotic Bcl-2 family proteins are recognized to be similar to certain types of α-helical ion-channel or pore-forming proteins. Specifically, these Bcl-2 family proteins share structural similarity with the pore-forming domains of certain bacterial toxins which have been implicated in transport of either ions (colicins) or proteins (diphtheria toxin) across membranes. At least *in vitro*, antiapoptotic Bcl-2 family proteins such as Bcl-2 and Bcl-X_L have been documented to form multiconductance ion channels in synthetic membranes, thus providing further experimental evidence in support of a role as channel/pore proteins. Second, antiapoptotic Bcl-2 family proteins also possess at least one hydrophobic pocket on their surface which mediates interactions with other proteins, thus altering the function or intracellular targeting of other proteins which may be relevant in some contexts to cell death. Proteins reported to interact directly or indirectly with antiapoptotic Bcl-2 family members such as Bcl-2 or Bcl-X_L include the caspase activator Apaf-1, the p53-binding protein p53BP2, the Ca^{2+}-dependent phosphatase calcineurin, the protein kinase Raf-1, the Hsp70/Hsc70 molecular chaperone regulators Bag1 and Bag3, the spinal muscular atrophy gene product Smn, the DED-containing proteins BAR and Bap31, and others. It is currently unknown which of these two major functions of antiapoptotic Bcl-2 family proteins is more important for their overall cytoprotective effects in cells. It is clear,

however, that Bcl-2 and Bcl-X_L possess cytoprotective functions which are independent of caspase-family cell death proteases and which can be manifested even in yeast, which lack the various Bcl-2/Bcl-X_L-interacting proteins described above. Hence it seems likely that these proteins possess an intrinsic biochemical function. Additional ancillary functions related to interactions with other protein may have been added over evolutionary time to this core-intrinsic function, as a means of integrating multiple pathways with cell life and death decision making (Reed, 1997).

Proapoptotic Bcl-2 family proteins are more structurally diverse than their antiapoptotic counterparts, probably reflecting differences in the mechanisms by which they promote cell death. A subgroup of proapoptotic Bcl-2 family proteins, which in humans includes Bax, Bak, Bok and Bid, appears to possess a similar pore/channel-like protein fold as the antiapoptotic members. Indeed, where tested, these pro-apoptotic Bcl-2 family members have been documented to form multiconductance, sometimes fairly large, channels in synthetic membranes *in vitro*. They also induce cytochrome *c* release from mitochondria when overexpressed in cells as well as when added (where tested) as recombinant proteins to isolated mitochondria *in vitro*. It remains controversial as to the mechanisms that explain why Bax, Bak, Bok and Bid are proapoptotic and induce cytochrome *c* release from mitochondria, whereas the structurally similar antiapoptotic proteins Bcl-2, Bcl-X_L, Bcl-W, Mcl-1, Bcl-B, and Bfl-1 are cytoprotective and block cytochrome *c* release.

By analogy with the bacterial proteins with which they share structural homology, it has been suggested that Bcl-2 family proteins may exist in two dramatically different conformations—one in which two of more α-helices of these proteins are integrated into membranes and another in which they are not integrated into membranes, although they are often tethered to membranes via their C-terminal membrane-anchoring domains. These two conformational states may then dictate whether these proteins are actively forming channels or perhaps interacting with other channel or pore-like proteins in mitochondrial membranes (inserted state) versus whether they assume conformations necessary for interacting with certain other proteins (noninserted state). Thus, the concept of on/off (active/inactive) conformations has emerged as an element in understanding the regulation of Bcl-2 family proteins—which may not correlate necessarily with the levels of expression of these proteins.

Apoptosis (caspase-dependent death) and necrosis (caspase-independent death) share some elements of overlap and the release of cytochrome *c* from mitochondria represents one of these. When cytochrome *c* exits mitochondria, not only can apoptosis be induced through the Apaf-1-dependent mechanisms described above, but also necrosis can ensue owing to failure of mitochondrial electron chain-transport, resulting in inadequate ATP production, increased generation of oxygen free radicals and other disturbances. Several Bcl-2 family proteins have been shown to regulate both apoptosis and necrosis, probably at least in part owing to their direct caspase-independent effects on mitochondrial membrane permeability. Although many mechanisms remain unclear, it can be deduced that these direct effects of Bcl-2, Bax and certain other Bcl-2 family proteins on mitochondria are also independent of CED-4/Apaf-1-family proteins, given that (1) caspases are required neither for induction of cytochrome *c* release by Bax, Bax, and Bok nor for prevention of cytochrome *c* release by Bcl-2 or Bcl-X_L and (2) the only known function of CED-4/Apaf-1 family proteins is regulation of caspases. Interestingly, a variety of data have provided evidence that clonogenic survival often correlates better with caspase-independent cell death than with caspase-dependent apoptosis, suggesting that a cell death commitment point which is regulated by Bcl-2 family proteins exists upstream of caspase activation in the mitochondrial pathway (Green and Reed, 1998).

Members of another subgroup of pro-apoptotic Bcl-2 family proteins, which in humans includes Bad, Bik, Bim and Hrk, lack similarity to pore-forming domains but do possess a short ($\sim$16–20 amino acid) domain found in nearly all Bcl-2 family proteins known as a Bcl-2 homology-3 (BH3) domain. The BH3 domain participates in protein dimerization among Bcl-2 family members, and therefore represents a functionally important region that permits many pro- and antiapoptotic Bcl-2 family proteins to antagonize each other's functions through physical hand-to-hand combat. The BH3 domain consists of a conserved amphipathic α-helix which inserts into the hydrophobic receptor-like pockets found on (or predicted to exist on) the antiapoptotic Bcl-2 family proteins. Bad, Bik, Bim and Hrk all bind selected antiapoptotic Bcl-2 family members via their BH3 domains, thus antagonizing the cytoprotective functions of Bcl-2, Bcl-X_L or similar proteins, but these killer proteins do not interact with themselves, nor do they bind to pro-apoptotic Bcl-2 family members such as Bax, Bak, Bok and Bid. Specificity exists in the preferences for dimerization partners within the network of Bcl-2 family protein interactions, thus creating opportunities for fine-tuning regulation through tissue-specific and temporally dynamic differences in the repertoires of family members expressed.

Several ways of regulating dimerization among Bcl-2 family proteins have been documented. One obvious way is by controlling the levels of various Bcl-2 family members at the transcriptional or post-transcriptional levels, as has been demonstrated for several members of the family. Another mechanism for controlling dimerization involves protein phosphorylation/dephosphorylation. For example, the proapoptotic protein BAD is a substrate for phosphorylation by several protein kinases, including Akt, Pak1, PKA and Raf-1. When phosphorylated on certain sites, BAD is unable to dimerize with Bcl-2 or Bcl-X_L,

thus abrogating its proapoptotic function. Elevations in Akt activity in cancers, therefore, suppress apoptosis at least in part through effects on BAD, at least in those cases where BAD is expressed. Conversely, sustained elevations in cytosolic Ca^{2+}, resulting in activation of the protein phosphatase calcineurin, have been shown to result in dephosphorylation and reactivation of BAD. In addition to phosphorylation/dephosphorylation, dimerization among some Bcl-2 family proteins can be regulated by proteolysis. For instance, the proapoptotic protein Bid is normally present in the cytosol in a latent state. Removal of an N-terminal 58 amino acid domain by caspase-mediated cleavage activates the Bid protein, exposing its BH3 domain, thus allowing it to bind other Bcl-2 family proteins including Bcl-2, Bcl-X_L, Bax and Bak. The caspase responsible for Bid cleavage is caspase-8, thus providing a mechanism for cross-talk between the death receptor ('extrinsic') and mitochondrial ('intrinsic') cell death pathways. Yet another mechanism of regulating dimerization among Bcl-2 family proteins can be found in the Bim protein. The longer isoforms of Bim are present in a complex with dynein light chain in association with microtubules, thus sequestering Bim proteins in a location where they cannot interact with other Bcl-2 family members. Release of Bim proteins from microtubule-associated protein complexes allows their translocation to the surfaces of mitochondria and other organelles where antiapoptotic Bcl-2 family dimerization partners reside. Release of Bim from microtubules may be relevant to some of the mechanisms by which microtubulin-interacting anticancer drugs induce apoptosis (Huang and Strasser, 2000).

Many examples of altered expression or structure of Bcl-2 family proteins have been documented in human cancers. In addition to its inappropriate overexpression in lymphomas containing the t(14;18) translocations described above, pathological overexpression of Bcl-2 (without attendant structural changes to the gene) has been estimated to occur in roughly half of all human cancers, including most advanced hormone-refractory prostate cancers, two-thirds of breast cancers and over half of all colon and lung cancers. Examples of upregulation of other antiapoptotic proteins such as Bcl-X_L or downregulation of proapoptotic proteins such as Bax and Bak also exist in human cancers. Moreover, mutations that inactivate the *BAX* gene have been detected. One of the more common mutations involves a homopolymeric stretch of eight guanosine residues, which occurs frequently in cancers that manifest the microsatellite instability phenotype due to errors in DNA mismatch repair enzymes or their regulators. Consistent with their important role in controlling apoptosis sensitivity to chemotherapeutic drugs and X-irradiation, the levels of Bcl-2, Bax or other Bcl-2 family proteins are of prognostic significance for some subgroups of patients with certain types of cancer.

INTERACTIONS OF SIGNAL TRANSDUCTION PATHWAYS WITH APOPTOSIS PATHWAYS

Many upstream inputs exist which link cellular responses to various stimuli with the core components of the apoptosis machinery. An exhaustive review of these connections is beyond the scope of this chapter, but a few deserve special mention.

The transcription factor NF-κB has been implicated in apoptosis suppression in many contexts and appears to be elevated in its activity in many types of cancer (Karin and Ben-Neriah, 2000). NF-κB can consist of various dimerizing pairs of Rel-family members, with the best studied representing a heterodimer of p50 and p65. NF-κB directly upregulates the transcription of several antiapoptotic proteins, including (1) the Bcl-2 family members Bcl-X_L and Bfl-1, (2) the IAP-family member cIAP-2, and possibly cIAP-1 and XIAP under some circumstances, (3) the TRAF-interacting protein A20, which displays antiapoptotic activity in some contexts, and (4) IEX-1L, an anti-apoptotic protein of unknown mechanism. Consequently, NF-κB has emerged as an attractive drug-discovery target for cancer therapy, with most efforts aimed at suppression of the activities of upstream kinases such as the IKKα and IKKβ, which are responsible for phosphorylating IκB-family proteins, the endogenous suppressors of NF-κB—resulting in targeting of these proteins for degradation via a ubiquitin/proteasome-dependent pathway.

As described above, the protein kinase Akt serves as an apoptosis-suppressing link between a variety of growth factor receptors and membrane-associated oncoproteins (Datta *et al*., 1999). Several apoptosis relevant substrates of Akt have been identified to date, including (1) the proapoptotic Bcl-2 family protein BAD, (2) the apical protease in the mitochondrial pathway caspase-9, (3) Forkhead-family transcription factors that control transcription of the promoters of apoptosis-inducing genes including Fas-L and (4) the NF-κB-inducing kinase IKKα. In each case, the net effect of phosphorylation of these substrates by Akt is increased protection from apoptosis. Multiple mechanisms for achieving abnormally elevated levels of Akt activity have been documented in tumour cells, perhaps the most common of which is mutations in the *PTEN* gene which result in decreased elimination of the second-messenger polyphosphoinositol phospholipids responsible for initiating Akt activation.

The tumour suppressor p53 is a well-known inducer of apoptosis which becomes inactivated in approximately half of all human tumours as a result of gene mutations and deletions, alterations in p53 kinases (ATM, CHK2), changes in the levels of p53 antagonist proteins and their regulators (MDM2, p19-ARF), ectopic expression of viral oncoproteins (HPV E6; SV40 large T antigen) and other mechanisms (Yan *et al*., 2000). Although p53 has other

functions relevant to cell cycle control, DNA repair responses and genetic instability, analysis of transgenic mice (when crossed with oncogene-bearing transgenic mice) suggest that the apoptosis-inducing activity of p53 is probably its most important attribute for suppression of tumorigenesis *in vivo*. How p53 induces apoptosis remains controversial. Three things however seem clear: (1) apoptosis induction by p53 derives from its function as a transcription factor; (2) p53 possesses multiple potential mechanisms for promoting apoptosis; and (3) the specific mechanisms employed are highly tissue-specific and may vary among clonal neoplasms even of the same lineage. Among the documented apoptosis-regulatory genes that make possible contributions to p53-induced apoptosis are the following: (1) the proapoptotic gene *BAX*, whose promoter in humans contains at least four consensus p53-binding sites and which is directly transcriptionally induced by p53; (2) Fas, which is a direct transcriptional target of p53; (3) DR5, another TNF-family member which may be a direct target of p53 trans-activation; and (4) Fas-L, which reportedly can be induced to translocate by p53 from a sequestered Golgi location to the plasma membrane where its receptor (Fas) primarily resides. Given the importance of p53 in inducing apoptosis of cancer cells (as well as certain types of normal cells) following genotoxic injury, this tumour suppressor has received extensive attention as a possible therapeutic target.

Multiple members of the nuclear receptor (NR) family of ligand-responsive transcription factors have been documented as regulators of the transcription of selected apoptosis genes. For example, retinoids which interact with and activate RAR- and/or RXR-family receptors are known to downregulate the expression of the antiapoptotic *BCL-2* and/or *BCL-*X_L genes in certain malignancies. Attempts to exploit this attribute of retinoids have been made in the clinic and are likely to continue into the future. Conversely, oestrogen is a positive regulator of *BCL-2* gene expression in mammary epithelial cells and in oestrogen receptor (ER)-positive breast cancers, possibly explaining some of the proapoptotic effects of antioestrogens such as tamoxifen on breast cancer cells *in vivo*. PPAR-γ ligands, which include certain prostaglandins produced by cyclooxygenases and other enzymes, as well as synthetic drugs that engage these receptors, have also been shown to either up- or downregulate transcription of *BCL-2*, depending on cellular context. These and other examples illustrate that many opportunities exist to regulate the output of apoptosis-relevant genes via effects on NR-family transcription factors in cancers.

THERAPEUTIC IMPLICATIONS

Knowledge about the core components of the apoptosis machinery and the various upstream inputs into apoptosis pathways has suggested a wide variety of new strategies for devising new therapies for cancer (Reed, 1999; Nicholson, 2000; Reed and Tomaselli, 2000). The full range of therapeutic modalities can be envisioned, including (1) small-molecule drugs that directly bind to and modulate the activities of specific protein targets; (2) antisense, DNAzyme and ribozyme nucleic acid-based therapeutics; (3) gene therapy using proapoptotic proteins; and (4) biologicals such as recombinant protein ligands or monoclonal antibodies, in the case of cell surface targets. Already proof of concept data have been obtained in animal models for many apoptosis-modulating agents, and some of these have advanced into clinical trials. At the time of this writing, for example, phase III trials are underway exploring the efficacy of antisense oligonucleotides directed against *BCL-2* mRNA for patients with a variety of types of refractory tumours. Recombinant Trail (Apo2L) protein is nearing its debut into clinical trials, as an attempt to trigger TNF-family death receptor pathways in tumour cells. Retinoids and PPAR-γ ligands are currently being examined as possible apoptosis sensitizers in leukaemias and solid tumours. Gene therapy trials are underway involving local delivery of p53- or *BAX*-expressing viral vectors, attempting directly to restore p53 or Bax expression in tumour cells. Monoclonal antibodies, recombinant proteins and synthetic peptidyl ligands that induce apoptosis of migrating endothelial cells by interfering with integrin-generated signals for cell survival (probably mediated largely by Akt) are also under investigation in clinical trials, as an approach for inhibiting angiogenesis. Preclinical analysis is also under way of multiple agents, ranging from small molecule organic compounds that restore activity to mutant p53 in malignant cells to synthetic peptides representing BH3 domains of pro-apoptotic Bcl-2 family proteins combined with membrane-penetrating peptides derived from viruses.

Our understanding is also rapidly evolving about the mechanisms by which currently available anticancer agents successfully trigger apoptosis of tumour cells and how resistance develops in all too many instances. A few examples include (1) binding of the triphosphate forms of purine nucleoside analogues to Apaf-1, resulting in improved catalytic efficiency of Apaf-1-mediated activation of caspases relative to endogenous dATP, (2) liberation of the BH3-only proapoptotic Bcl-2 family member Bim from microtubules by agents that disturb normal microtubule polymerization and (3) antioestrogen (tamoxifen)-mediated downregulation of Bcl-2 expression in breast cancers.

Exploiting apoptosis-based therapies for the treatment of cancer must be achieved with an acceptable therapeutic index, resulting in a selective killing of malignant cells. Fortunately, much evidence suggests that the inherent abnormalities found in cancer cells may also render them selectively more vulnerable compared with normal cells when deprived of their roadblocks to apoptosis. For example, several proto-oncogenes which become hyperactive in cancers, such as c-*Myc* and *cyclin-D1 (BCL-1)*,

drive rapid tumour cell division but also promote cell death unless apoptosis is concomitantly blocked (Evan and Littlewood, 1998). Genetically unstable cells also suffer errors in cell cycle checkpoint regulation and DNA/chromosome management which can be triggers for apoptosis when cell death pathways are intact. Finally, metastatic cells are potentially vulnerable because they often depend on defects in apoptosis pathways for avoiding cell death as a result of loss of survival signals from unoccupied cell adhesion receptors and from absence of local growth/survival factors. It seems clear, therefore, that apoptosis-based therapies will eventually find their place in the armentarium of weapons that will be used to wage and eventually to win the war on cancer.

ACKNOWLEDGEMENTS

The author acknowledges research funding from the National Cancer Institute, National Institutes of Health, Department of Defense, California Breast Cancer Research Program and CaP-CURE.

REFERENCES

Adams, J. and Cory, S. (1998). The Bcl-2 protein family: arbiters of cell survival. *Science*, **281**, 1322–1326.

Altieri, D. C., *et al*. (1999). Survivin apoptosis: an interloper between cell death and cell proliferation in cancer. *Laboratory Investigations*, **79**, 1327–1333.

Ashkenazi, A. and Dixit, V. (1998). Death receptors: signaling and modulation. *Science*, **281**, 1305–1308.

Ashkenazi, A. and Dixit, V. M. (1999). Apoptosis control by death and decoy receptors. *Current Opinions in Cell Biology*, **11**, 255–260.

Cryns, V. and Yuan, Y. (1999). Proteases to die for. *Genes and Development*, **12**, 1551–1570.

Datta, S., *et al*. (1999). Cellular survival: a play in three Akts. *Genes and Development*, **13**, 2905–2927.

Deveraux, Q. and Reed, J. (1999) IAP family proteins: suppressors of apoptosis. *Genes and Development*, **13**, 239–252.

Evan, G. and Littlewood, T. (1998) A matter of life and cell death. *Science*, **281**, 1317–1322.

Fesik, S. W. (2000) Insights into programmed cell death through structural biology. *Cell*, **103**, 273–282.

Green, D. R. and Reed, J. C. (1998) Mitochondria and apoptosis. *Science*, **281**, 1309–1312.

Gross, A., *et al*. (1999). BCL-2 family members and the mitochondria in apoptosis. *Genes and Development*, **13**, 1899–1911 (1999).

Huang, D. C. and Strasser, A. (2000). Bh3-only proteins—essential initiators of apoptotic cell death. *Cell*, **103**, 839–842.

Karin, M. and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-κB activity. *Annual Reviews of Immunology*, **18**, 621–663.

Kroemer, G. and Reed, J. C. (2000) Mitochondrial control of cell death. *Nature Medicine*, **6**, 513–519.

LaCasse, E. C., *et al*. (1998). The inhibitors of apoptosis (IAPs) and their emerging role in cancer. *Oncogene*, **17**, 3247–3259.

Lowin, B., *et al*. (1995). Perforin and granzymes: crucial effector molecules in cytolytic T lymphocyte and natural killer cell-mediated cytotoxicity. *Current Topics in Microbiology Immunology*, **198**, 1–24.

Metzstein, M., *et al*. (1998). Genetics of programmed cell death in *C. elegans*: past, present and future. *Trends in Genetics*, **14**, 410–416.

Miller, L. (1999). An exegesis of IAPs: salvation and surprises from BIR motifs. *Trends in Cell Biology*, **9**, 323–328.

Nicholson, D. W. (2000). From bench to clinic with apoptosis-based therapeutic agents. *Nature*, **407**, 810–816.

Reed, J. C. (1997). Double identity for proteins of the Bcl-2 family. *Nature*, **387**, 773–776.

Reed, J. (1998a). Caspases and cytokines: roles in inflammation and autoimmunity. *Advances in Immunology*, **73**, 265–287.

Reed, J. (1998b). Bcl-2 family proteins. *Oncogene*, **17**, 3225–3236.

Reed, J. (1999). Dysregulation of apoptosis in cancer. *Journal of Clinical Oncology*, **17**, 2941.

Reed, J. and Paternostro, G. (1999). Post-mitochondrial regulation of apoptosis during heart failure. *Proceedings of the National Academy of Sciences of the USA*, **96**, 7614–6.

Reed, J. C. (2000). Mechanisms of apoptosis. *American Journal of Pathology*, **157**, 1415–1430.

Reed, J. C. and Tomaselli, K. (2000). Drug discovery opportunities from apoptosis research. *Current Opinions in Biotechnology*, **11**, 586–592.

Salvesen, G. S. and Dixit, V. M. (1997). Caspases: intracellular signaling by proteolysis. *Cell*, **91**, 443–446.

Salvesen, G. S. and Dixit, V. M. (1999). Caspase activation: the induced-proximity model. *Proceedings of the National Academy of Sciences of the USA*, **96**, 10964–10967.

Thornberry, N. and Lazebnik, Y. (1998). Caspases: enemies within. *Science*, **281**, 1312–1316.

Tschopp, J., *et al*. (1998). Inhibition of Fas death signals by Flips. *Current Opinions in Immunology*, **10**, 552–558.

Tschopp, J., *et al*. (1999). Apoptosis: silencing the death receptors. *Current Biology*, **9**, R381–R384.

Vander Heiden, M. G. and Thompson, C. B. (1999). Bcl-2 proteins: regulators of apoptosis or of mitochondrial homeostasis? *Nature Cell Biology*, **1**, E209–E216.

Vaux, D. and Korsmeyer, S. (1999). Cell death in development. *Cell*, **96**, 245–254.

Wallach, D., *et al*. (1999). Tumor necrosis factor receptor and Fas signaling mechanisms. *Annual Reviews of Immunology*, **17**, 331–367.

Wyllie, A. H. (1997). Apoptosis: an overview. *British Medical Bulletin*, **53**, 451–465.

Yan, H., *et al*. (2000). Genetic testing—present and future. *Science*, **289**, 1890–1892.

FURTHER READING

Abrams, J. (1999). An emerging blueprint for apoptosis in drosophila. *Trends in Cell Biology*, **9**, 435–440.

Jacobson, M. D., *et al*. (1997). Programmed cell death in animal development. *Cell*, **88**, 347–354.

Metzstein, M., *et al*. (1998). Genetics of programmed cell death in *C. elegans*: past, present and future. *Trends in Genetics*, **14**, 410–416.

Wyllie, A. H., *et al*. (1980). Cell death, the significance of apoptosis. *International Review of Cytology*, **68**, 251–306.

Chapter 10

Signalling by Steroid Receptors

Torsten A. Hopp and Suzanne A. W. Fuqua
Breast Center at Baylor College of Medicine, Houston, TX, USA

CONTENTS

- Introduction
- Structure and Function of Oestrogen Receptors α and β
- The ER is a Ligand-dependent Transcription Factor
- The Importance of Oestrogen Receptor Expression in Cancer Initiation and Progression
- Conclusion

INTRODUCTION

Steroid hormones (androgens, oestrogens glucocorticoids and progestins) play a vital role in the development and maintenance of normal cellular function, as well as regulatory functions within the reproductive organs. These hormones mediate their activities through binding to specific intracellular receptor proteins, called the androgen receptor (AR), oestrogen receptor (ER), glucocorticoid receptor (GR) and progesterone receptor (PR), respectively. The steroid receptors (SRs) are members of a large superfamily of nuclear receptors, which modulate expression of target genes upon binding of their respective hormones (Evans, 1988). One of the important characteristics of this protein family is the highly conserved organisation of functional domains, implying that the underlying mechanism of transcriptional modulation might also be preserved. A highly simplified and schematized model of SR action can be seen in **Figure 1**: upon binding of the steroid hormone, the SR forms dimers and binds to a specific steroid response element (SRE) in the 5′ flanking region of hormone-responsive genes and stimulates/modulates their transcription.

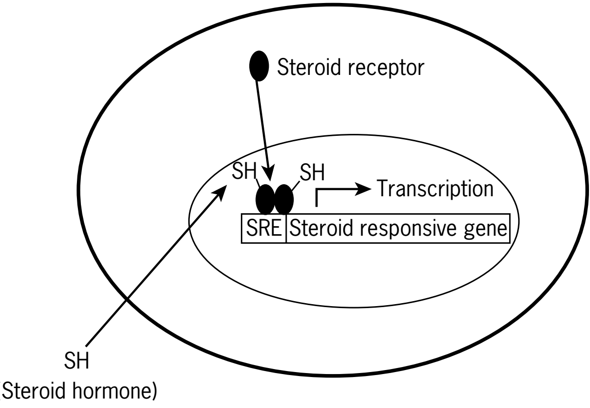


Figure 1 Simplified model of steroid action. Steroids enter a target cell by diffusion through the cell membrane and bind to steroid receptor (filled ovals). The hormone-bound receptor complex stimulates transcription of target genes via interaction with a steroid response element.

Recent research studying the molecular mechanism of transcriptional regulation of target genes by SRs has revealed a very complex network of protein–protein interactions in addition to protein–DNA interactions necessary for proper function. Disruption in this intricately regulated circuitry can perturb SR signalling. Specifically, mutations in the AR and ER, and also altered receptor expression, have been found in prostate and breast cancer, and have been associated with cancer progression and hormone resistance. However, our understanding of how these changes can affect the SR signalling pathway is still very limited in many cases. To comprehend the complex network of SR signalling, we will focus here on the ER, which has been extensively studied owing to its importance in the clinical management of breast cancer (Osborne, 1998).

STRUCTURE AND FUNCTION OF OESTROGEN RECEPTORS α AND β

Like many other members of the nuclear receptor family, the ER has more than one form encoded by separate genes. In the case of ER there are two, called *ER*α and -β. The human *ER*α gene resides on chromosome 6q25.1 and is transcribed in a single mRNA of 6.5 kb that encodes a protein of 595 amino acids with an approximate molecular mass of 66 kDa (Kumar *et al.*, 1987). Even though the *ER*α gene is transcribed from at least three different promoters in a cell- and tissue-specific manner, only a single open reading frame appears to exist. However, the three promoters transcribe mRNA isoforms which differ in their 5′ untranslated regions, but no biological differences have yet been reported.

The *ERβ* gene is located on chromosome 14q22–24 and encodes a protein of 530 amino acids with a molecular mass of 60 kDa (Mosselman *et al.*, 1996). Unlike *ERα*, several studies indicate that *ERβ* is transcribed as multiple mRNAs and also translated into proteins from at least two reading frames resulting in a second ERβ protein which lacks 53 amino acids of the N-terminus. Overall, ERα and ERβ are highly homologous (49% amino acid identity), but appear to be expressed in different organs and at different developmental stages. Interestingly, ERα is expressed in female organs such as the mammary gland, uterus, ovary, vagina and certain areas of the hypothalamus, whereas ERβ is found in the male in different areas of the hypothalamus when compared with ERα, and in the cerebral cortex. Additionally, *ERβ* mRNA has also been detected in prostate, ovary, uterus, lung, testis and artery.

In Vivo Function of the Oestrogen Receptors

The development of mice lacking the *ERα* (*αERKO*) or *ERβ* (*βERKO*) gene have proved to be valuable tools in evaluating the *in vivo* function of these receptors. The αERKO mice were generated in 1993, and the disruption of ERα expression not only caused infertility in both sexes, but also had profound effects on behaviour (Couse and Korach, 1999). Specifically, pre- and neonatal development of female reproductive organs such as uterus, ovary and mammary gland was almost normal, but maturation of these organs during and after puberty was severely impaired. The αERKO females also failed to display sexual receptivity when treated with the hormonal regime of oestrogen and progesterone that normally induces receptivity in wild-type mice. Surprisingly, adult αERKO males have significantly fewer epididymal sperm than heterozygous or wild-type males, caused by the disruption of spermatogenesis and degeneration of the seminiferous tubules, which becomes apparent 10 weeks after birth. Furthermore, these males develop obesity after sexual maturation, in addition to exhibiting decreased normal male-typical aggressive behaviour, including offensive attacks, and show a reduced number of mount attempts as compared with wild-type animals. Interestingly, both sexes of the αERKO mice also show a 20–25% reduction in bone density, implying that ERα is crucial for proper bone maturation and mineralization. However, the only described case of oestrogen insensitivity in a human male, which was normally masculinized, had incomplete epiphyseal closure with a history of continued growth into adulthood, and also osteoporosis probably induced by increased bone turnover.

More recently, the generation of βERKO mice revealed that ERβ does not affect normal development, and mice lacking ERβ are indistinguishable grossly and histologically as young adults from their littermates (Couse and Korach, 1999). Females are fertile and exhibit normal sexual behaviour, but have fewer and smaller litters owing to reduced ovarian efficiency, and multiple resorbed fetuses. Older males lacking ERβ display signs of prostate and bladder hyperplasia. In contrast to the αERKO animals, the βERKO females exhibit normal breast development and lactate normally, while all components of sexual behaviour in βERKO mice were found to be intact. These observations indicate that unlike ERα, ERβ is essential for normal ovulation efficiency but not for female or male sexual differentiation, fertility or lactation.

Structural and Functional Domains

The ERα and ERβ protein sequences encode all the above-mentioned *in vivo* functions, and can be divided into six functional domains, A–F (**Figure 2**) (Kumar *et al.*, 1987). The N-terminal A/B domain, which contains a hormone-independent transactivation function (AF-1), as well as the C-terminal F domain, demonstrate only weak homology (approximately 17%) between the two ERs. In contrast, the central C domain, which contains the DNA-binding domain (DBD) present in all SRs, possesses 97% homology between ERs α and β, suggesting that they recognize and bind to the same or similar oestrogen response elements (EREs) consisting of inverted repeats of the sequence GGTCA separated by three variable nucleotides (Kumar and Chambon, 1988). This DNA-binding domain consists of two functionally distinct zinc-finger motifs, forming a helix–loop–helix motif typical of many transcription factors. Following the C domain is the D domain, which appears to function as a hinge region between the DBD and the ligand-binding domain (LBD), and demonstrates only moderate homology (30%) between the two ERs. However, recent research demonstrates that this domain contains important binding sites for receptor coactivators/corepressors, as will be discussed later, and might even be involved in post-translational regulation of the ER. Finally, the C-terminal E domain contains the LBD and, even though ERα and ERβ possess only moderate homology (59%) in this domain, both receptors bind with nearly the same affinity to oestradiol and to other natural and synthetic ligands. The E domain also contains a ligand-dependent transactivation function (AF-2) and provides a crucial interface for the interaction with many coactivators, as will be discussed later. Recently, a third AF, AF-2a, was identified in the ER hinge and LBD domains (Norris *et al.*, 1997). AF-2a has been shown to activate transcription in a ligand-independent manner in the absence of both AF-1 and AF-2.

Structure of the Ligand-binding Domain

The LBD can be viewed as a molecular switch that, upon hormone binding, enables the receptor to activate transcription of target genes by direct interaction of the receptor with DNA in the promoter region of these genes

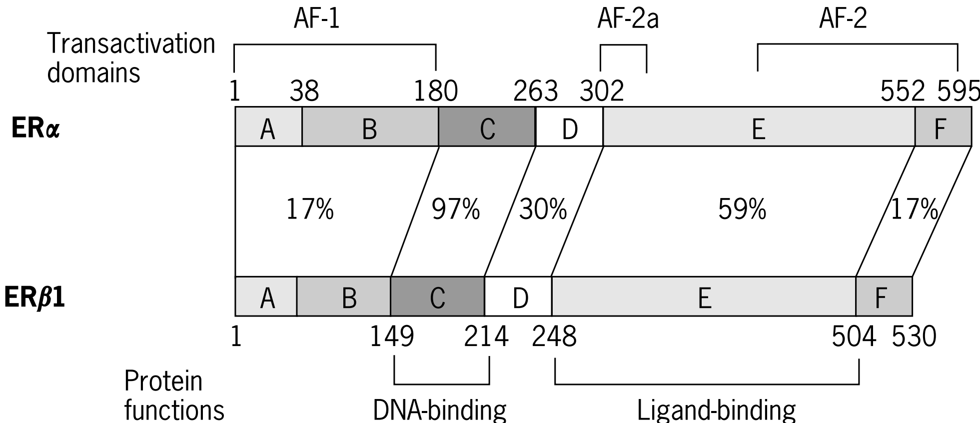


Figure 2 Comparison of ERα and ERβ1 functional domains. The amino acid residues of ERα are shown above domains A–F, and those of ERβ1 are shown below the domains. The degree of homology between them is shown as a percentage. Also shown are the regions for receptor function such as DNA binding, ligand binding and transcriptional activation functions.

and with components of the transcriptional machinery. In addition to the ligand-binding function, the LBD also contains signals necessary for nuclear localization, homo- and heterodimerization, in addition to the above-mentioned AF-2.

Recently, the crystallographic resolution of agonist (oestradiol) and antagonist (raloxifene) bound LBD of ERα revealed a compact structure consisting of 12 α-helices that form the ligand-binding pocket (Brzozowski *et al.*, 1997). This structure appears to be a common motif found also in other SR. The binding of oestradiol into the binding pocket induces important structural changes; in particular, helix 12, that prior to ligand binding extends away from the body of the domain, is repositioned over the ligand-binding pocket and seals this pocket like a 'lid.' This process is thought to trap the ligand in a hydrophobic environment, and also forms a coactivator binding surface on the LBD. Helix 12 is fixed in this active position by contact with both the hormone and amino acid residues in helices 3 and 4 on the surface of the LBD. In contrast, binding of the antioestrogen raloxifene into the binding pocket makes helix 12 extend away from the LBD, a conformation similar to one seen in unliganded receptor.

THE ER IS A LIGAND-DEPENDENT TRANSCRIPTION FACTOR

In the absence of hormone, the ER is thought to exist in a complex with chaperone proteins, like heat shock proteins Hsp90 and Hsp70, which may help to maintain the receptor in an appropriate conformation to respond rapidly to hormonal signals. Upon binding of oestrogen, this oligomeric complex dissociates, allowing the ER to homo- or heterodimerize and to interact with target gene promoters through two possible mechanisms. First, the ER can interact directly with DNA by binding to specific EREs, resulting in a bending of the DNA toward the major groove. This bending is thought to facilitate the interaction of ER with proteins of the transcription complex. Second, ERα is also able to interact indirectly with oestrogen-regulated gene promoters through its association with other transcription factors such as AP-1, NF-κB, C/EBPβ, GATA-1 and SP-1. It is currently thought that gene transcription depends on the formation of a preinitiation complex that consists of basal transcription factors. However, it has recently become clear that ER also recruits a host of ancillary factors, which are called coactivators if they enhance and corepressors if they inhibit receptor transcriptional activity.

The Interaction of ER with Basal Transcription Factors

The initiation of transcription by RNA polymerase II requires the assembly of basal transcription factors such as transcription factor IIA (TFIIA), TFIIB, TFIID, TFIIE and TFIIF. The binding of TFIID is the first and the rate-limiting step in this assembly process. TFIID consists of the TATA-box binding protein (TBP) and more than 10 other TBP- associated factors.

There is ample *in vitro* evidence of direct protein–protein interactions between ERα and basal transcription factors. In particular and consistent with the evidence that TBP recruitment is a rate-limiting step in transcriptional initiation, both the N-terminal AF-1 and the C-terminal AF-2 activation function of ERα have been reported to bind to this basal transcription factor. Another component

of TFIID, $TAF_{II}30$, also interacts with the AF-1 of ERα in a ligand-independent manner, and this interaction appears to be required for ERα-mediated transactivation. Following the binding of the TFIID complex to the promoter of target genes, the binding of TFIIB, RNA polymerase II and TFIIF is also required for the assembly of the minimal initiation complex. TFIIB not only contacts DNA sequences both upstream and downstream of the TATA box and other factors of the basal transcription machinery, but also interacts with AF-2 of ERα. However, the significance of all of these interactions is unclear at present since they are not significantly affected by mutations in the ER that are known to disrupt transcriptional activity.

The Interaction of ER with Coactivators

The observation that different classes of SR can interfere with each other's transcriptional activity, termed 'squelching', has indicated that SRs compete for limited amounts of assembly proteins called coactivators or corepressors, in addition to other factors of the basal transcriptional machinery. Over the last 7 years, a large number of potential coactivators have been identified, they have been extensively reviewed by Klinge (Klinge, 2000). Coactivators are generally defined as proteins that can interact in a ligand-dependent manner with DNA-bound SRs and are able to enhance their transcriptional activity. Furthermore, a coactivator should also be able to interact with components of the basal transcription machinery, but not to enhance basal transcriptional activity on their own, although they often contain an autonomous activation function. As examples, we will focus here on the mechanism of action of the intensively studied SRC family of coactivators and also the Creb binding protein (CBP)/p300 in regulating ER's function.

The SRC Family of Coactivators

Several coactivators have been identified as a family of related proteins, called the SRC family, which includes SRC-1 (also termed ERAP-160 or NcoA-1), SRC-2 (also called GRIP1, NcoA-2, or TIF-2) and SRC-3 (also known as ATCR, RAC3, p/CIP, AIB1 or TRAM-1). These coactivators are all able to stimulate oestradiol-mediated gene transcription and promote the interaction between AF-1 and AF-2 to produce full transcriptional activity. The SRC family shares a common domain structure with an overall sequence similarity of 40% between the three members (**Figure 3**). The highest degree of homology is observed in the N-terminal bHLH (for basic helix–loop–helix) and PAS (for Per/Arnt/Sim homology) domains.

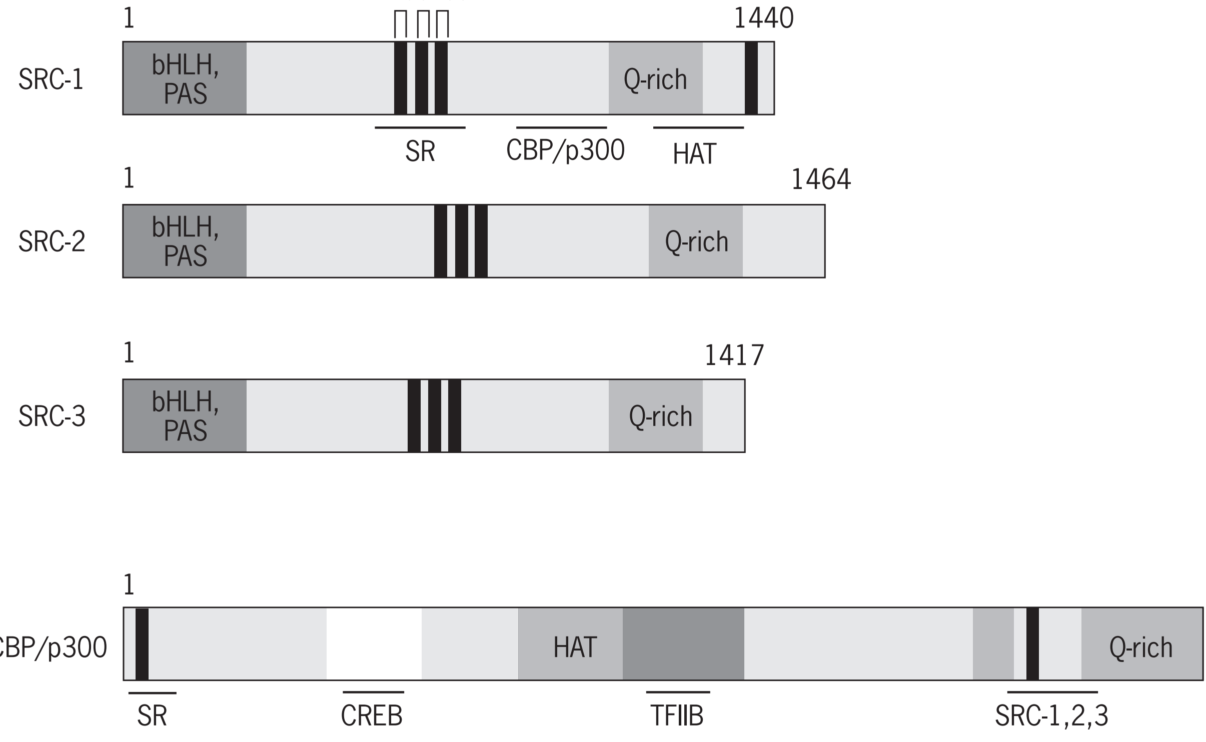


Figure 3 Structural features of the SRC family of coactivators and CBP. Regions to which specific functions of individual coactivators have been assigned are indicated. Below the coactivator representations are also shown the regions of interaction for various nuclear proteins such as steroid receptors (SR), CBP/p300, CREB, TFIIB and SRC1,2,3.

These particular domains mediate homodimeric and heterodimeric interactions between proteins containing these motifs, but their specific function in SRC coactivators remains unknown. It has been speculated that the presence of these motifs might indicate potential cross-talk between the SR pathway and other PAS-containing factors.

Most SR coactivators, including the SRC-family of coactivators, contain one or more copies of a short sequence motif, LXXLL (L stands for leucine and X is any amino acid), that are necessary and sufficient to mediate ligand-dependent direct interaction with the ER. This motif is called the nuclear receptor (NR) box or the receptor-interacting domian (RID). All of the SRC coactivators possess three NR boxes in the central portion of the protein, whereas only SRC-1 has an additional NR box in the extreme C-terminus. Protein structure prediction analyses of these motifs have indicated that they form amphipathic helices with the conserved leucine residues outlining a hydrophobic surface on the face of the helix. Cocrystallization of the LBD of ERα with NR box peptides, as well as systematic mutagenesis of ERα, have revealed that the helix formed by the NR boxes is able to interact with a hydrophobic groove in the AF-2 domain of ER consisting of ERα residues from helices 3, 4, 5 and 12. This hydrophobic groove is the result of conformational changes induced by hormone binding, as discussed earlier.

The SRC coactivators also contain intrinsic activation domains that retain their activity when tethering the coactivator to DNA. Additonally, SRC-1 and SRC-3 contain an intrinsic histone acetyltransferase (HAT) activity which is thought to modulate chromatin structure by histone acetylation, thus facilitating the access of other transcriptional regulators as well as the assembly of the preinitiation complex. Interestingly, these coactivators together with ER are believed to form a ternary complex with CBP/p300, CREB, and other proteins (**Figure 4**), but the function of this complex remains largely unclear. However, a recent study reporting oestradiol-dependent acetylation of SRC-3

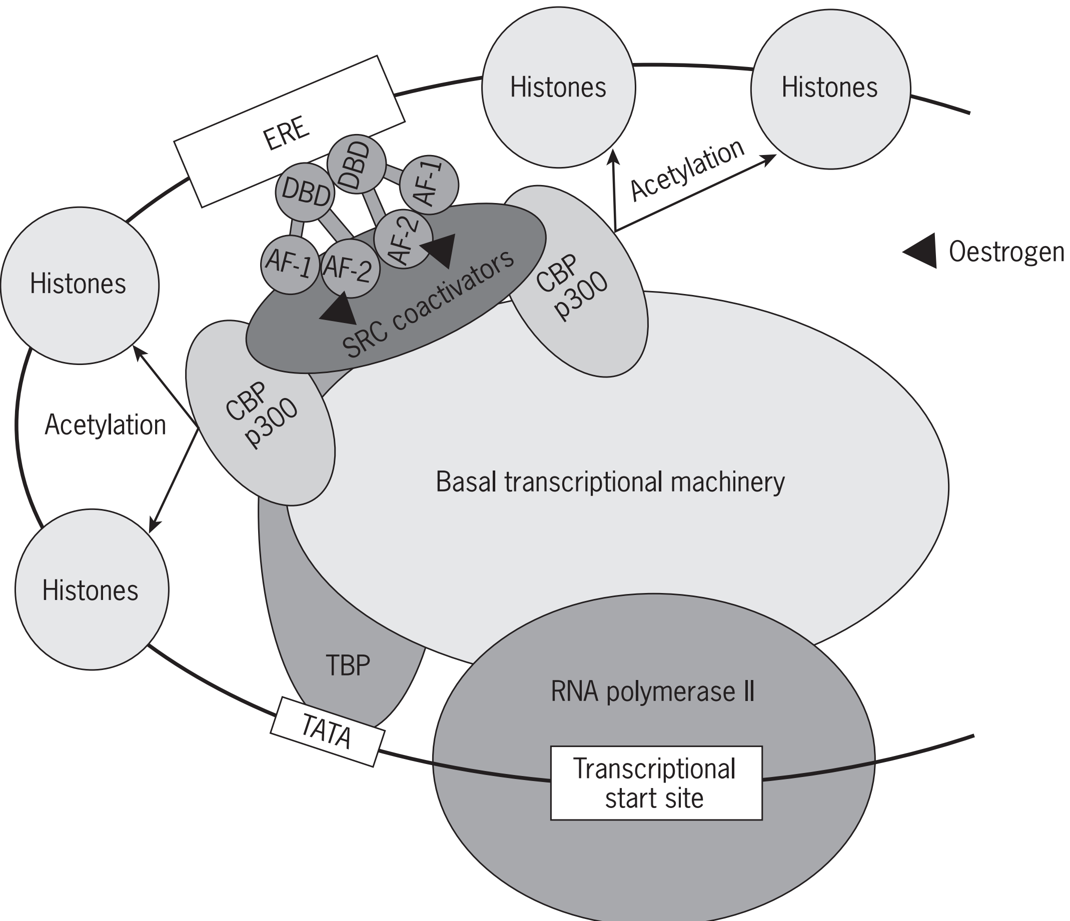


Figure 4 A schematic model of the transcription initiation complex formed at the ERE of an oestrogen-responsive gene. The ER is able to recruit an SRC coactivator upon oestrogen binding, which subsequently results in the recruitment of additional coactivators, such as CBP, and basal transcription factors. However, the precise *in vivo* composition of this complex is still under investigation. Additionally, histone acetylation by CBP/p300 is thought to facilitate the relaxation of chromatin at the target gene promoter, thereby enhancing transcriptional activation.

by CBP/p300 might reveal a new insight into the function of this complex. Surprisingly, acetylation of three lysine residues immediately upstream of the NR box of SRC-3 decreased SRC-3:ERα interactions *in vitro*, and disruption of this interaction appears to be a necessary event for the down-regulation of ER activity. These results indicate that hormone-induced transcription may be dynamically regulated by both histone and SRC-3 acetylation.

The recent development of mice lacking either SRC-1 or SRC-3 provides *in vivo* evidence of a partial functional redundancy between SRC-1 and SRC-2, while the physiological role of SRC-3 in development or disease appears to be different from that associated with SRC-1 expression. In particular, both SRC-1 and SRC-3 knockout mice are viable and fertile. However, in mice lacking SRC-1, oestrogen target organs such as the uterus, prostate, testis and mammary gland display decreased growth and development in response to steroid hormones, as well as increased expression of SRC-2 that is thought to compensate partially for the loss of SRC-1 function. On the other hand, disruption of SRC-3 expression in mice results in dwarfism, delayed puberty, reduced female reproductive function and blunted mammary gland development. This pleiotropic phenotype indicates that SRC-3 plays a critical role in overall growth and sexual maturation.

Interestingly, SRC-1 and SRC-2, but not SRC-3, interact also with ERβ and enhances its hormone-dependent transcriptional activity. In addition, SRC-3 is amplified and overexpressed in many ER-positive breast cancer cell lines, and a large study analysing 1157 clinical breast tumours and 122 ovarian tumours also found amplification of this coactivator gene in a small percentage of breast and ovarian cancers. Expression of SRC-3 seems to correlate with tumour size and with ERα and PR positivity. It has also been reported that SRC-1 and SRC-2 expression is relatively low in breast tumours when compared to normal tissues.

The Cointegrators CBP/p300

The cointegrator CBP and its related functional homologue p300 are thought to be responsible for the integration of numerous environmental stimuli on promoters containing multiple *cis*-acting elements (Goodman and Smolik, 2000). Even though CBP was initially characterized as a coactivator required for efficient activation of cAMP-regulated genes, several studies also implicate this protein as a coactivator for a broad range of transcription factors, including p53, NFκB and the SRs. CBP interacts with ERα in a hormone-dependent manner, and this interaction depends on a crucial NR box in the N-terminal domain of this CBP. Surprisingly, SRC-1 is also able to interact directly with carboxy terminus of CBP and p300, which synergistically enhances ERα-activated gene activity. Despite all the described potential interactions between CBP/p300 and transcription factors, coactivators and the basal transcription machinery, there is little biochemical evidence for the existence of these complexes *in vivo*. There is only one small study analysing the chromatographic elution of SRC-1-containing complexes from T47D breast cancer cells that indicates the existence of distinct coactivator complexes with different properties and activities.

The central portion of both CBP and p300 encodes a relatively large domain that possesses intrinsic HAT activity. Among the major substrates of acetylation are the core histones, particularly the N-terminal tails of histones H3 and H4 (**Figure 4**). The unmodified forms of these histones are thought to maintain DNA packing into highly organized chromatin structures, thus silencing transcription, while a high degree of histone acetylation correlates with increased promoter activity. More recently, it has been shown that nonhistone proteins such as p53, GATA-1, the AR and SRC-3 (that itself possesses HAT activity) are also substrates for CBP/p300 acetylation. However, the precise function of both histone and nonhistone protein modifications in transcriptional regulation is still largely unclear.

Interaction of ER with Corepressors

Unlike coactivators, only very few corepressors of ER action have been reported to date. These include NCoR (nuclear receptor corepressor; also called RIP13), SMRT (silencing mediator for retinoid and thyroid receptor, also termed TRAC2), REA (repressor of oestrogen receptor activity), SHP (short heterodimer partner) and BRCA-1 (breast cancer susceptibility gene). Repression of ER activity can occur in the absence of hormone or when an antagonist is bound to the receptor.

The corepressors NCoR and SMRT were originally identified by biochemical studies of cellular proteins associated with unliganded thyroid hormone receptor (TR) and retinoid acid receptor (RAR). These corepressors are thought to recruit another complex of proteins, including histone deacetylase (HDAC) activity, to DNA-bound TR and RAR (**Figure 5**). This complex is believed to repress gene expression by maintaining chromatin in a more condensed state which impairs the ready access of critical transcription factors. Upon ligand binding, corepressors are released from TR and RAR, and coactivators and basal transcription factors are then recruited to the receptors. In contrast, unliganded or hormone-bound ERα is unable to interact with either NCoR or SMRT, but the antioestrogen tamoxifen (an oestrogen antagonist which is commonly used in the treatment of ER-positive breast cancer) induces the interaction of ER with these corepressors. Interestingly, the antagonist tamoxifen can be switched into in agonist, much like oestrogen, when ERα's ligand-independent AF-1 is activated by the MAPK pathway (this kind of activation will be discussed later in the section

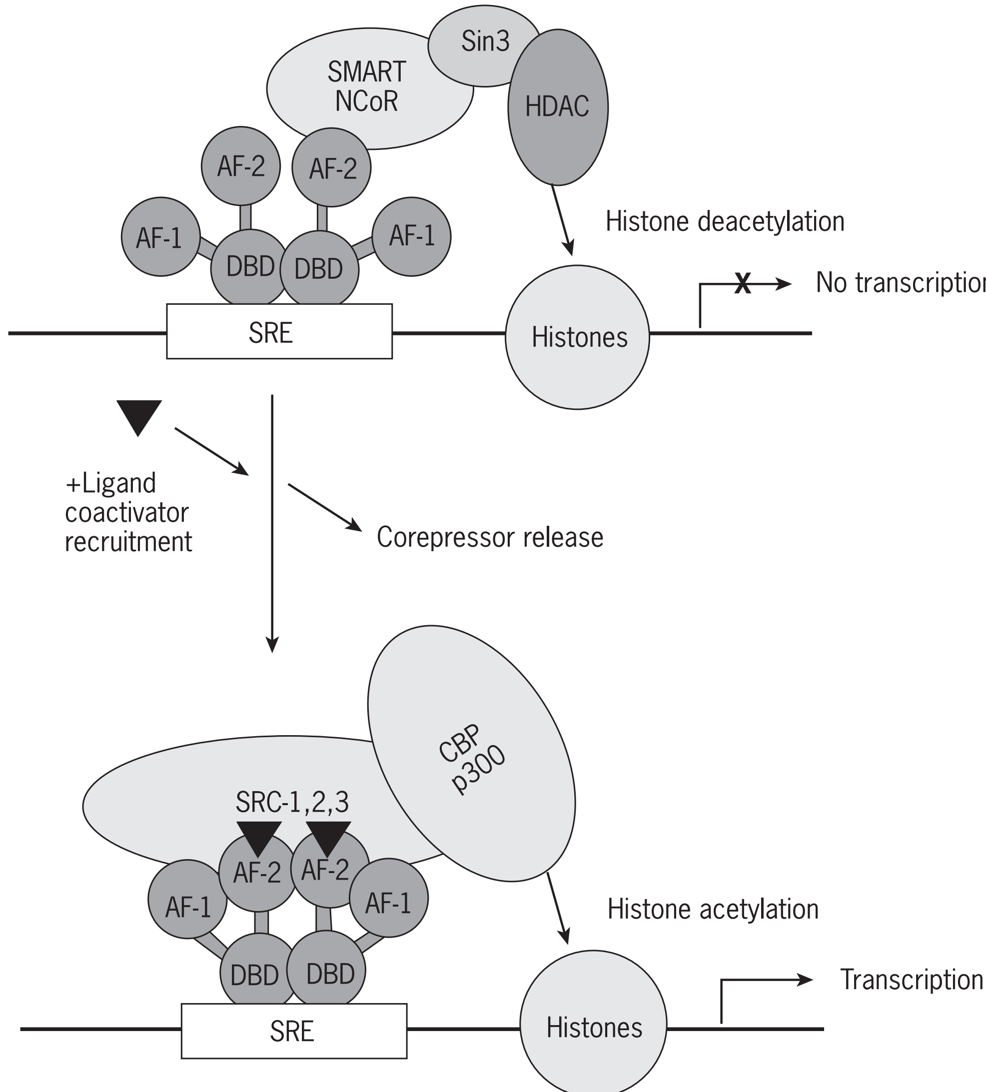


Figure 5 Ligand-dependent switch between an SR associated with either a corepressor or a coactivator complex. The SR is associated with a corepressor (SMRT or NCoR), which in turn recruits a histone deacetylase (HDAC) through its interaction with Sin3. Deacetylation of histones leads to transcriptional repression. Hormone binding disrupts the interaction between SR and repression complex in favour of the association of a coactivation complex. This complex consists in part of a SRC coactivator and CBP/p300, which possess histone acetyltransferase activity that modifies histones resulting in relaxation of chromatin structures and initiation of transcription.

Phosphorylation). This activation results in the release of corepressors from ERα and recruitment of coactivators to this receptor. This phenomenon may well explain why the majority of breast tumours become resistant to tamoxifen.

Regulation of ER Function

The activity of many transcription factors is regulated by post-translational modifications such as phosphorylation, proteasome-mediated degradation and cross-talk with other signal transduction pathways. We will focus upon these three aspects in our discussion of ER function.

Phosphorylation

Stimulation of a number of growth factor receptors and/or protein kinases leads to the phosphorylation of ER resulting in ligand-independent and/or a synergistic increase in transcriptional activation in response to hormone (Kato *et al*., 2000).

The AF-1 region of ERα contains phosphorylation sites for a number of kinases including MAPK, cyclin A/cdk2 and PI3/AKT. In particular, much work has been focused on phosphorylation of serine residue 118 within the A/B domain of ERα. Phosphorylation of this particular residue by growth factor activated MAPK leads to the

enhancement of the N-terminal AF-1 function. MAPK is activated by tyrosine kinase cell membrane receptors, that in turn are stimulated by growth factors such as insulin, IGF-1, EGF and TNF-α. Furthermore, phosphorylation at ER residue 118 also enhances the interaction of this receptor with the p68 RNA helicase, resulting in increased AF-1 but not AF-2 activity.

The phosphorylation of ERβ by MAPK in cells treated with exogenous EGF or via overexpression of activated Ras has been shown to enhance binding between SRC-1 and the ERβ AF-1 domain. This suggests that ligand-independent activation of ERβ also depends on phosphorylation of the N-terminal region, and that this event may be important for the recruitment of coactivators such as SRC-1.

Much less is known about the phosphorylation by the cyclin A/cdk2 or the PI3/Akt kinases and their effects on specific ER function. Specifically, phosphorylation of two serine residues at amino acid 104 and 106 of the ERα by the cyclin A/cdk2 enhances the ERα AF-1 function both in the absence of oestrogen or in the presence of tamoxifen. Furthermore, it was recently shown that PI3 kinase is able to increase both AF-1 and AF-2 activity, whereas Akt, a kinase downstream of PI3 kinase, increases only AF-1 activity. Phosphorylation of the ERα serine residue 167 by Akt also results in protection of breast cancer cells from tamoxifen-induced apoptosis, revealing an important potential mechanism for the onset of resistance to tamoxifen in breast cancer.

Another major ligand-independent phosphorylation site within ERα is the conserved tyrosine 537 residue and the homologous tyrosine 443 residue in ERβ, but the exact consequence of phosphorylation at this site is still controversial and remains under intense investigation. Replacement of this tyrosine residue with other amino acids suggests that phosphorylation at this site is important for hormone binding and transcriptional activation. Specifically, the substitution of tyrosine 537 with alanine, asparagine, or serine results in a mutant ERα that is constitutively active and binds to SRC-1 even in the absence of oestrogen. Another recent study also indicated that phosphorylation of the tyrosine 537 is critically involved in ligand-induced conformational changes of the ERα.

Proteasome-mediated Degradation of the ER

Recently, ubiquitin-dependent, proteasomal degradation of ligand-bound ERα was discovered as an additional mechanism involved in the regulation of hormone receptor-mediated gene transcription. The SRC coactivator family is also a target for degradation via the 26S proteasome. It is well known that the ubiquitin pathway is involved in the degradation of many short-lived proteins (Hershko and Ciechanover, 1998). Through a series of enzymatic reactions, ubiquitin is covalently linked to proteins targeted for degradation, marking them for recognition by the 26S proteasome, a large multisubunit protease. Abnormalities in ubiquitin-mediated degradation have been shown to cause several pathological conditions, including malignant transformation. In particular, it has recently been shown that ERα is ubiquitinated preferentially in the presence of hormone. It is thought that ERα protein degradation, which occurs through the 26S proteasome complex, is required for continued transcriptional activation by this receptor. ERα degradation could well be an important requisite to dissociate the preinitiation complex resulting in the release of the components necessary for another round of transcription. On the other hand, hormone-induced degradation may also serve as a negative feedback to down-regulate the transcription of oestrogen-responsive genes.

THE IMPORTANCE OF OESTROGEN RECEPTOR EXPRESSION IN CANCER INITIATION AND PROGRESSION

Oestrogen receptor gene expression in breast epithelium is an intricately regulated event, and is thought to play a central role in normal breast development, and also breast cancer evolution. ERα expression is significantly increased in both premalignant and malignant breast lesions, and many of these ERα-positive cells proliferate as compared with normal breast. Furthermore, normal breast epithelium, in addition to breast cancer tissue, contains alternatively spliced *ERα* and *ERβ* mRNA variants, but it is still unclear whether changes in the levels of these variants impact upon tumour development or the progression to hormone-independent tumour growth. Single amino acid mutations within the ERα are relatively rare, but may contribute to the progression of breast cancer or metastatic disease. We will next describe the potential role of ERα expression in premalignant disease, as well as the role of specific ER variants and mutations in breast cancer development and progression.

Oestrogen Receptor Expression in Normal Breast and Breast Cancer

In normal nonpregnant, premenopausal human breast, only about 5–10% of the total luminal epithelial cell population expresses ERα, and this expression tends to be highest in the follicular phase of the menstrual cycle. The highest percentage of ERα-expressing cells are found in undifferentiated lobules type 1 (Lob1), with a progressive reduction in the more differentiated Lob2 and Lob3 types. The highest level of cell proliferation is also observed in Lob1, but expression of ERα occurs in cells other than these

proliferating cells, indicating that they represent at least two separate cell populations. These data also suggest that oestrogen might stimulate ERα-positive normal cells to produce a growth factor that in turn stimulates neighbouring ERα-negative normal cells to proliferate. In premalignant and malignant breast lesions, however, ERα expression is significantly increased in the proliferating cell compartment, suggesting that ERα may be involved in the earliest changes to malignancy. Additionally, approximately two-thirds of breast tumours, at least initially, express abundant levels of ERα, and this expression is associated with lower risk of relapse and prolonged overall survival. Unfortunately, we still understand very little about the precise role of ERα expression in tumour progression.

Owing to its recent discovery, only limited data are available on the expression and function of ERβ in normal breast and its potential role in breast carcinogenesis. Studies of ERβ knockout mice suggest that ERβ plays a limited role in normal breast development and function. However, ERβ expression appears to be important for the growth control of urogenital tract epithelium, and may even afford a protective role against hyperproliferation and carcinogenesis in this particular tissue. This interesting hypothesis might also apply to the mammary gland, and is supported by a recent study reporting that ERβ expression in breast tumours is positively associated with ERα and progesterone receptor expression, as well as negative axillary nodes, DNA diploidy and low S phase fraction, all of which imply that ERβ-positive tumours may have a more favourable prognosis. On the other hand, two studies examining a relatively small number of tumours using RT-PCR determined that coexpression of ERβ and ERα is frequently associated with poor prognostic biomarkers, such as positive axillary nodes and higher tumour grade, and also that ER is significantly elevated in tumours resistant to tamoxifen treatment.

Oestrogen Receptor Variant Forms

Both of the *ER* genes undergo alternative splicing in normal and neoplastic oestrogen-responsive tissues. Alternative splicing results in *ER* mRNA variants with single or multiple exons skipped, and are usually coexpressed along with the wild-type receptor (Hopp and Fuqua, 1998). It is still unclear whether any or all of the *ER* splicing variants are indeed stably translated *in vivo*, and to what extent the formation of heterodimers of these splice variants with ERα and ERβ perturb the ER signalling pathway. *ERα* and *ERβ* variant forms fall into four major groups: (1) transcripts containing precise single or multiple exon deletions, (2) transcripts containing single nucleotide deletions and others in which several hundred nucleotides have been deleted within known exon sequences, (3) truncated transcripts and (4) transcripts containing insertions.

The most frequently observed *ERα* mRNA splice variants are those lacking exon 4, which has been detected in normal and neoplastic tissue, or exon 7, detected in many breast tumours regardless of their receptor expression status. The exon 4-deleted *ERα* variant is missing the nuclear localization signal and part of the hormone-binding domain, thus potentially encoding a protein whose cellular distribution and oestrogen-binding affinity may be different from those of the wild-type ERα. On the other hand, the exon 7-deleted splicing variant potentially translates into a receptor protein missing the C-terminal part of the hormone-binding domain, which includes the hormone-dependent AF-2 function and the F-domain. However, both *ERα* variants, when transfected into mammalian or yeast cells, can block normal ER signalling in certain cell types, but not in others. In addition, expression of the exon 4-deleted *ERα* variant is associated with two biological markers of good clinical outcome (PR positivity and low histological grade), and thus may prove useful as biological marker of good prognosis in clinical samples.

Another dominant-negative inhibitor of ERα signalling is the exon 3-deleted variant, which encodes a variant receptor lacking portions of the DNA-binding domain. This variant is found in both normal tissue and primary breast cancer, and the ratio of the exon 3-deleted variant to wild-type ERα is reduced about 30-fold in breast cancer cell lines as compared with normal tissue. Stable expression of the exon 3-deleted variant in MCF-7 breast cancer cell lines, which contain high levels of endogenous ERα, results in reduction of both invasiveness and anchorage-independent growth.

One of the best studied *ERα* mRNA splice variants is the exon 5-deleted receptor, which is the only variant so far detected at the protein level in breast cancer cell lines and breast tumours. This variant is a truncated 40-kDa protein missing most of the hormone-binding domain, but it retains AF-1 function and demonstrates variable strengths of hormone-independent transcriptional activity depending on the cell type. Expression of this variant was also significantly increased in cancers from patients relapsing after tamoxifen treatment as compared with the respective primary tumour, suggesting that tumours expressing high levels of the exon 5-deleted variant may acquire resistance to tamoxifen. However, it seems likely that other *ERα* splice variants could also be involved in acquired tamoxifen resistance, since multiple *ERα* variants can occur in the same tumour sample and tamoxifen resistance is thought to be multifactorial.

ER Mutations

In contrast to the abundant expression of *ERα* mRNA splice variants in both normal and neoplastic tissue, mutations of the *ERα* gene are seldom found in primary breast cancer. Changes in the *ER* nucleotide sequence

fall into at least two groups: (1) polymorphisms, which do not change the amino acid sequence, and (2) missense mutations, which do alter the amino acid sequence (Hopp and Fuqua, 1998).

Polymorphisms have been detected in both primary and metastatic breast cancer, but these silent changes do not appear to correlate with clinical parameters, such as tumour type, size, grade or stage. Like polymorphisms, missense mutations in the *ER*α, which potentially affect normal function, have been found in primary and metastatic breast cancers. However, functional characterization of most of these mutations is still missing.

Recently, a specific somatic mutation in the ERα has been found in many typical hyperplasias, a type of premalignant lesion that carries an increased risk of breast cancer development. The mutation substitutes a lysine with an arginine residue at amino acid 303, at the border between hinge domain and the beginning of hormone-binding domain. This mutant ERα shows much higher sensitivity to oestrogen than wild-type ERα, resulting in markedly increased proliferation at subphysiological levels of hormone. Additionally, the mutation enhances the ability of the SRC-2 coactivator to bind at physiological levels of hormone. These data suggest that this mutant receptor may promote or accelerate the development of cancer from premalignant breast lesions.

Another missense mutation, where tyrosine 537 is substituted by asparagine, was isolated from a metastatic lesion from a breast cancer patients. This mutant ERα exhibits a potent, oestradiol-independent transcriptional activity that is only weakly affected by oestrogen, and variably by antioestrogens. As mentioned earlier, phosphorylation of the tyrosine residue at codon 537 is thought to be required for efficient oestrogen binding, and substitution of this amino acid with asparagine may induce conformational changes mimicking hormone binding. Stable expression of this mutant ERα in an ERα-negative breast cancer cell line caused increased production of PTHrP, a known stimulator of osteoclastic bone resorption and a major mediator of the osteolytic process. Furthermore, TGFβ, which is abundant in bone marrow, significantly enhanced the transcriptional activity of this mutant receptor, resulting in further stimulation of PTHrP production. These data indicate a central role for the ERTyr537Asn mutant in the pathogenesis of osteolytic bone metastases from breast carcinoma.

The Role of Oestrogen Receptors in Hormone Resistance and Independence

As mentioned earlier, a large number of breast cancers express high levels of ERα, and the ER status of patients is highly predictive of response to long-term tamoxifen therapy, an antioestrogen frequently used in the treatment of breast cancer. Unfortunately, most breast cancers eventually acquire tamoxifen resistance, resulting in disease recurrence and the frequent emergence of more aggressive disease. Tamoxifen-resistant tumours often continue to express ERα, so that mechanisms for antioestrogen resistance other than the loss of ERα must exist. Interestingly, in some cases antioestrogen resistance is also reversible, e.g. tamoxifen-resistant patients who have been switched to a different type of therapy may later once again respond to tamoxifen.

Potential mechanisms of resistance include alterations in the expression levels of ERα and ERβ as well as *ER* splicing variants, ER mutations, interaction with other growth factor signal transduction pathways (such as erbB-2 and AP-1), abnormal expression or function of coactivators and corepressors and metabolic tolerance as a result of altered systemic antioestrogen metabolism.

CONCLUSION

The purpose of this chapter was to review ER signalling and its perturbation in cancer as an example of SR signalling in general. The ER is an important transcriptional activator for genes involved in many essential processes. Cloning and sequencing of the ERs and the resolution of the crystal structures of hormone- and antioestrogen-bound ERα, and also the development of mice lacking these receptors, have all increased our understanding of the structure–function relationships of these important transcription factors. However, the fairly recent discovery of receptor coactivators and corepressors has provided another level of complexity to models of oestrogen action. The understanding of the molecular mechanisms of action of the ER is beginning to provide an explanation for the function of clinically useful antiestrogens, and is also suggesting potential new therapeutic targets. Further progress in understanding the fine details of transcriptional activation may also provide new insights into mechanisms of hormone resistance.

ER expression is tightly regulated in normal breast epithelium, but increased ERα and expression of mutant ERs may drive abnormal proliferation in premalignant hyperplasias, providing a fertile environment for genetic alterations which, in turn, are associated with tumorigenesis. Unfortunately, very little is known about the precise role of ER expression during the transition of premalignant disease to cancer and the eventual development of hormone resistance. Continued exploration of the basic molecular mechanisms of ER signalling, as an example of steroid receptor action, will certainly enhance our understanding of underlying causes of some of the most prevalent human cancers, and may also provide new treatment approaches, as well as and new mechanisms to prevent these diseases.

REFERENCES

Brzozowski, A. M., *et al*. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, **389**, 753–758.

Couse, J. F. and Korach, K. S. (1999). Estrogen receptor null mice: what have we learned and where will they lead us? *Endocrinology Reviews*, **20**, 358–417; Erratum, **20**, 459.

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science*, **240**, 889–895.

Goodman, R. H. and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. *Genes and Development*, **14**, 1553–1577.

Hershko, A. and Ciechanover, A. (1998). The ubiquitin system. *Annuual Review of Biochemistry*, **67**, 425–479.

Hopp, T. A. and Fuqua, S. A. (1998). Estrogen receptor variants. *Journal of Mammary Gland Biology and Neoplasia*, **3**, 73–83.

Kato, S., *et al*. (2000). Molecular mechanism of a cross-talk between oestrogen and growth factor signalling pathways. *Genes and Cells*, **5**, 593–601.

Klinge, C. M. (2000). Estrogen receptor interaction with co-activators and co-repressors. *Steroids*, **65**, 227–251.

Kumar, V. and Chambon, P. (1988). The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell*, **55**, 145–156.

Kumar, V., *et al*. (1987). Functional domains of the human estrogen receptor. *Cell*, **51**, 941–951.

Mosselman, S., *et al*. (1996). ER beta: identification and characterization of a novel human estrogen receptor. *FEBS Letters*, **392**, 49–53.

Norris, J. D., *et al*. (1997). Identification of a third autonomous activation domain within the human estrogen receptor. *Molecular Endocrinology*, **11**, 747–754.

Osborne, C. K. (1998). Steroid hormone receptors in breast cancer management. *Breast Cancer Research and Treatment*, **51**, 227–238.

FURTHER READING

Elledge, R. M. and Fuqua, S. A. W. (2000). Estrogen and progesterone receptors. In: Harris, J. R., *et al*. (eds), *Diseases of the Breast*. 471–488 (Lippincott Williams & Wilkins, Philadelphia).

Freedman, L. P. (ed.) (1998). *Molecular Biology of Steroid and Nuclear Hormone Receptors* (Birkhauser, Boston).

Chapter 11

Signalling by Cytokines

William L. Farrar, Lihua Wang, Xiaoyi Yang, Weihua Xiao, O. M. Zack Howard
National Cancer Institute, National Institutes of Health, Frederick, MD, USA

Colin Duckett
National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

CONTENTS

- Structure and Function of γ_c-Dependent Cytokines and Receptors
- JAK–STAT Signal Pathway
- Diseases Associated with Perturbations in γ_c Receptor/JAK/STAT Signalling
- Interferon Signalling Pathway
- Tumour Necrosis Factor: Receptors and Signal Transduction Pathways
- Chemokine Receptor Signal Transduction

STRUCTURE AND FUNCTION OF γ_c-DEPENDENT CYTOKINES AND RECEPTORS

Members of the cytokine haematopoietic superfamily often share a common receptor subunit while retaining their own private receptor subunits. One well-documented example is that five cytokines, IL-2, IL-4, IL-7, IL-9 and IL-15, form one group, which is characterized by utilizing the common γ chain (γ_c) as a receptor subunit. The common γ_c subunit was initially cloned as the γ chain of the IL-2R complex. Soon it was discovered that this γ subunit also participates functionally in the receptors for IL-4, IL-7, IL-9 and IL-15 and, therefore, was designated γ_c, where c represents 'common.' This protein is constitutively expressed on essentially all cells of haematopoietic origin. It functions to enhance the binding of cytokines to their receptor, presumably by direct interaction with the ligand and to induce intracellular signals transduction events such as JAK–STAT signal pathway.

γ_c-Dependent Cytokines

γ_c-Dependent cytokines include IL-2, IL-4, IL-7, IL-9 and IL-15. They are peptides or glycoproteins with molecular masses of 14–20 kDa. Crystal structure analysis reveals that IL-2 is an α-helical protein, lacking β-sheet structure, with a four-fork core stabilized by a single intrachain disulfide bond. IL-4 contains six Cys residues that are all involved in intramolecular disulfide bridges. The secondary structure of IL-4 was shown to consist of a four-helix bundle with a unique up–up–down–down helix topology. IL-7, IL-9 and IL-15 contain a similar α-helical structure. Each cytokine is secreted by particular cell types in response to a variety of stimuli and produces a characteristic constellation of effects on the growth, motility, differentiation or function of its target cells. These cytokines exert multiple biological functions (**Table 1**). It is interesting that IL-2, IL-4 and IL-9, all produced by activated T cells, are important immune regulatory cytokines, whereas IL-7 and IL-15, which are primarily produced by nonlymphoid cells, have also been implicated in the regulation of lymphocyte development.

γ_c-Dependent Cytokine Receptors

With the exception of IL-2R and IL-15R, all γ_c-dependent cytokine receptor subunits are members of the cytokine receptor superfamily that contains an evolutionary-related extracellular region that results in a conserved structural fold for binding to helical cytokines. These receptor subunits are type I membrane glycoproteins with a single hydrophobic transmembrane domain. The extracellular domain contains two major regions of homology. The first is a region having four Cys residues located in the N-terminal half of that extracellular domain. The second region of homology is Trp-Ser-X-Trp-Ser(WSXWS), which is referred to as the 'WS motif.' This motif is close to the transmembrane region. The extracellular region also contains two fibronectin type III-like domains found in a series of cell surface molecules with adhesive properties. The functional significance of these domains remains to be clarified.

Table 1 Major properties of human γ_c-dependent cytokines

Cytokine	Mature protein (kDa)	Cellular source	Functional activities
IL-2	15	Activated T_H2 cells Tc cells NK cells T_HO cells	T cell growth Enhance B cell growth and Ig secretion Augment NK activity Induce LAK Programme T cells for apoptosis Reverse T cell anergy
IL-4	20	Activated T_H2 cells Mast cells Basophils MK1+ CD4+ T cells	T cell growth B cell growth IgG_1 and IgE class switch Enhance expression of MHC class II and CD23
IL-7	17	Bone marrow stroma Thymic stromal cells Intestinal epithelial cells Keratinocytes	T cell growth Proliferation of pre-B cells Viability of TN thymocytes Promotes development of CTL
IL-9	14	Activated T cells	Promotes the growth of mast cells Enhances mast cells secretion of IL-6 and expression of granzyme A and B and FcRε
IL-15	15	Placenta, epithelial cells Skeletal muscle Kidney, lung, fibroblasts Activated monocytes	T cell growth Enhanced NK activity Induce LAK Promote B cell growth and Ig secretion

JAK–STAT SIGNAL PATHWAY

All known γ_c-containing receptors signal through the associated Janus protein tyrosine kinases, JAK1 and JAK3 proteins, although not all γ_c-dependent cytokines activate the same STAT molecules. Phosphorylated tyrosines and flanking amino acid residues in the activated cytokine receptors determine this specificity by providing specific docking sites for the SH2 domains of STATs. Most likely, tyrosine phosphorylation of the receptor proteins is also directly mediated by JAKs. The JAK/STAT signal pathway, therefore, connects activation of the receptor complexes directly to transcription of genes. Upon receptor oligomerization, JAKs are activated, presumably by *trans*-'auto' phosphorylation on tyrosines. Subsequently, JAKs phosphorylate STAT proteins, which form homodimeric or heteromeric complexes via their SH2 domains. These complexes translocate to the nucleus, where they bind to specific targeting sequences and influence gene transcription (Horvath and Darnell, 1997).

Janus Kinases

The Janus kinases (JAKs) are cytoplasmic tyrosine kinases which mediate signalling from a number of cell surface receptors which lack intrinsic tyrosine kinase activity. Four mammalian members of the JAK family are known, JAKs 1–3, and TYK2 (Ihle, 1995). Whereas JAK1, JAK2 and TYK2 are expressed ubiquitously, expression of JAK3 is confined to haematopoietic and lymphoid cells. Characteristic of the structure of JAKs is the presence of two JAK homology (JH) domains, of which the C-terminal (JH1) domain has tyrosine kinase activity. Studies of knockout mice have provided important insights into the function of JAKs *in vivo*. The analysis of JAK3 knockout mice and JAK3-deficient humans has clearly demonstrated the essential, nonredundant role of JAK3 in several cytokine signalling pathways. The similarity with γ_c-deficient mice and humans strongly suggests that the major role of γ_c is the recruitment of JAK3 to each γ_c-receptor. In many cases, other JAKs, such as JAK1, that are found in association with the additional subunits of γ_c-containing cytokine receptors, are not sufficient to initiate signalling. JAK3-deficient mice are viable but exhibit severe defects in the development of lymphoid cells, the residual T cells being functionally deficient. Like JAK3-deficient mice, JAK1 deficiency leads to reduced numbers of T and B lymphocytes. Embryonic fibroblasts from JAK1 knockout mice do not respond to class II cytokine receptor ligands IFNγ and IFNα.

Signal Transducer and Activator of Transcription (STAT)

The STATs (signal transducers and activators of transcription) constitute a family of signal transduction proteins that are activated in the cytoplasm by the binding of extracellular polypeptides to transmembrane receptors and which then regulate the transcription of immediate-response genes. Following their obligatory tyrosine phosphorylation, induced by a cytokine ligand, STATs dimerize, translocate to the nucleus and bind directly to response elements present in the promoters of target genes in order to trigger induction of transcription. Thus far, six mammalian STAT proteins (plus several isoforms) have been identified (Darnell, 1997). Two homologues of STAT5 exist (STAT5A and STAT5B) that are encoded by different genes. Expression of STAT proteins is ubiquitous, except for STAT4, which is expressed in several tissues including spleen, heart, brain, peripheral blood cells and testis. Most STATs are activated by many different ligands. IL-2, IL-7, IL-9 and IL-15 activate STAT3 and STAT5, in contrast to IL-4, which activates STAT6. STAT knockout mice mainly show defects in a single or a few cytokine-dependent processes. Embryonic stem cells deficient for STAT3, or with dominant negative STAT3 proteins, fail to stay in an undifferentiated state in the presence of leucocyte inhibitory factors. STAT5A and STAT5B double knockout mice show loss of function with regard to prolactin and growth hormone receptors, i.e. disturbed ovary and mammary gland development and growth retardation. In addition, these mice lack NK cells, develop splenomegaly and have T cells with an activated phenotype, thus resembling IL-2 receptor β-chain-deficient mice. STAT6 knockout mice lack Th2 function as a consequence of impaired IL-4 and IL-13 signalling.

DISEASES ASSOCIATED WITH PERTURBATIONS IN γ_c RECEPTOR/ JAK/STAT SIGNALLING

Because γ_c-dependent cytokines orchestrate a variety of immune system responses via activating the γ_c receptor/ JAK/STAT signalling pathway, it is not surprising that most circumstances causing an inappropriate inhibition of this signalling pathway have generically immunosuppressive consequences. A number of pathological conditions have been identified with mutations or deregulation in γ_c cytokine receptors or associated signalling molecules.

SCID

Severe combined immunodeficiency (SCID) is a hereditary human disease characterized by functionally inactive T and B cells. The ensuing susceptibility to opportunistic infections is the prevalent cause of premature mortality in young patients suffering from this disease. More than 50% of SCID cases are X-linked. XSCID is commonly associated with mutations, which chromosomally map to Xq13, in the γ-chain of the IL-2, IL-4, IL-7, IL-9 and IL-15 receptors. Strikingly, a form of autosomal SCID exists with clinical symptoms identical with X-SCID, in which the gene encoding JAK3 is affected. A newly identified form of SCID with slightly different clinical features involves mutations in the interleukin 7 receptor chain. γ_c-Deficient mice were shown to have hypoplastic thymuses. Thymic cellularity was reduced by 10–25-fold when compared with normal littermates. CD4 and CD8 staining revealed the presence of all thymocyte subpopulations with a slightly increased proportion of CD4+ ‘single positive’ cells. In the bone marrow, B cell development was blocked at the pre-B cell stage. Although the cellularity of the spleen was reduced approximately 10-fold, mature T and B cells were detected. Both CD5+ and CD5− B-1 cells were identified in γ_c-deficient mice. The numbers of granulocytes, monocytes/macrophages and erythrocytes were normal or increased.

The phenotype of these γ_c-deficient mice indicated that signalling through γ_c is required for the development of multiple lymphoid lineages but not myeloid and erythroid lineages. When compared with human XSCID, a striking difference in B cell development was observed in mouse models of XSCID. In the mouse, B cell development was substantially inhibited at the pro-B cell stage, whereas in human XSCID, the production of B cells is outwardly normal. IL-7/IL-7R appears to represent the γ_c-dependent cytokine for mouse B cell development. The failure to block B cell development in human XSCID suggests a γ_c-independent pathway for the production of B cells in these patients.

Immunosuppressive Diseases and Suppression of JAK/STAT Signal Transduction

Cytokine receptor signalling substrates, in particular the JAKs and STATs, contribute to tumorigenesis. In *Drosophila*, a dominant mutant Jak kinase causes leukaemia-like abnormalities. In mammals, JAKs and STATs are known to be constitutively activated in haematopoietic cells transformed by diverse oncogenic tyrosine kinases and in a variety of lymphomas and leukaemias. Expression of a constitutively active STAT3 molecule in immortalized fibroblasts causes cellular transformation. Together these data are indicative of a role for constitutive activation of the JAK-STAT pathways in leukaemogenesis.

INTERFERON SIGNALLING PATHWAY

Interferons (IFNs) play a key role in mediating antiviral and antigrowth responses and in modulating the immune response (Seder, 1994; Trinchieri and Scott, 1995; Young and Hardy, 1995). Their signalling pathways provided the first evidence of, and have been used as the model for, the JAK–STAT pathway, which is utilized by many cytokines (Darnell *et al*., 1994). IFNs can be subdivided into two functional classes, and constitute the largest and most divergent subfamily of cytokines. There are more than 20 members in the type I IFN class (e.g. IFNαs, -β, and -τ) and one member in the type II IFN class, i.e. IFNγ. Type I and II IFNs function via related but distinct signal transduction pathways. In both classes of IFNs, signalling is initiated by the binding of the IFNs to their specific membrane receptors, that are expressed in many different cell types.

Interferon-γ Signalling Pathway

Type II, immune or IFNγ, is secreted by thymus-derived (T) cells under certain conditions of activation and by natural killer (NK) cells. The proximal events of IFNγ signalling require the obligatory participation of five distinct proteins, two IFNγ specific receptors, IFNGR1 and IFNGR2, two Janus kinase family members, JAK1 and JAK2, and one STAT family member, STAT1 (Schindler and Darnell, 1995).

IFNγ receptors contain a minimum of two peptide chains that are expressed in nearly all cell types, displaying a strict species specificity in their ability to bind IFNγ. The 90-kDa IFNGR1 consists of three domains, an extracellular (229aa), an intracellular (223aa) and a transmembrane (21aa). The IFNGR1 contains a constitutive and specific JAK1 binding site in the membrane-proximal region of the intracellular domain and a STAT1 binding and phosphorylation site that is activated upon phosphorylation by JAK1. The 62-kDa IFNGR2 differs from the 90-kDa IFNGR1 in that it contains a JAK2 binding site in its intracellular domain. Both IFNGR1 and IFNGR2 are required to activate functionally the IFNγ signalling pathway. IFNGR1 plays more important roles in mediating ligand binding, ligand trafficking through the cell and signal transduction, whereas IFNGR2 plays only a minor role in ligand binding.

Based on the available data and observations, Stark and colleagues have proposed a model, which has been broadly accepted for the IFNγ signalling pathway (**Figure 1**) (Stark *et al*., 1998). When a functional IFNγ homodimer binds two IFNGR1s on their extracellular domains, the IFNGR1 and IFNGR2 are brought into close proximity, forming more stable heterodimers together with their preassociated, inactive JAKs, JAK1 and JAK2. The

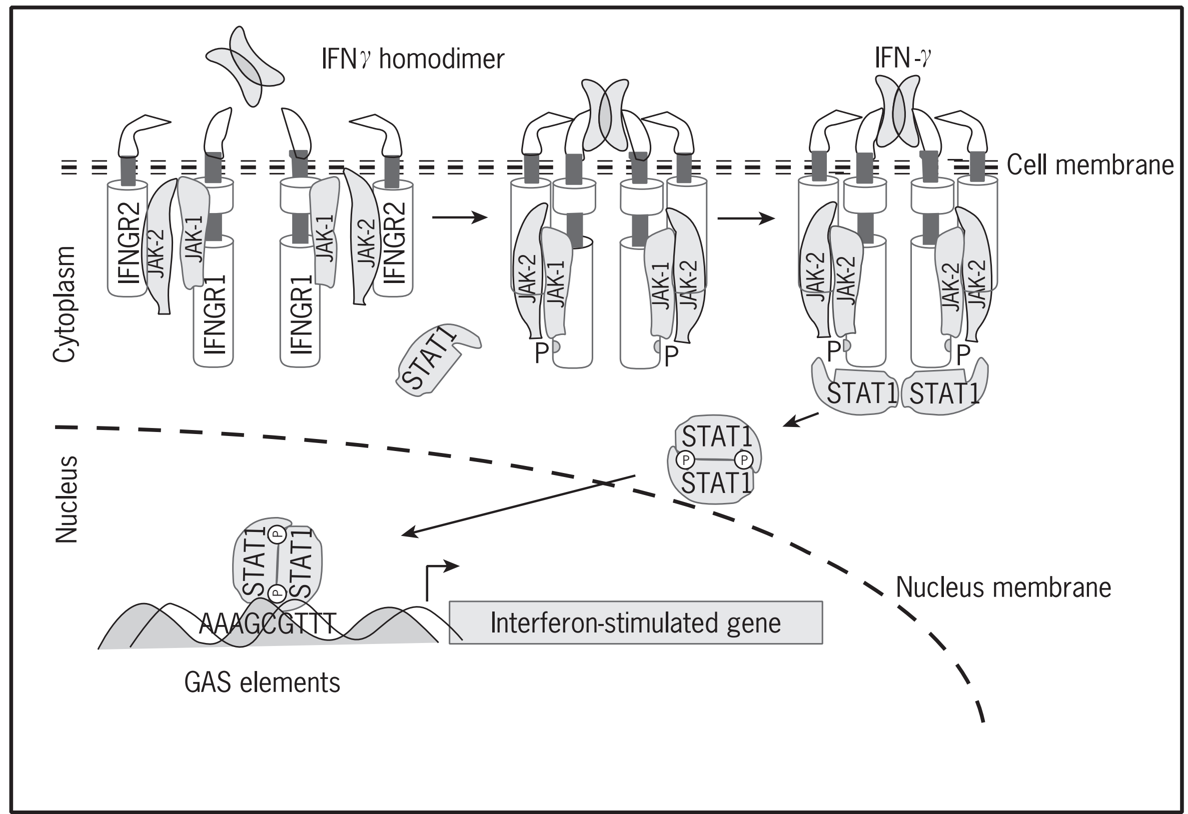


Figure 1 IFNγ signal pathway.

intracellular membrane-proximity domains of IFNGRs are then activated through auto- and transphosphorylation. The activation of the JAKs occurs in sequence, such that JAK2 activates first and is required for JAK1 activation.

Once activated, the JAKs phosphorylate a tyrosine-containing sequence near the C-terminus of IFNGR1, where paired ligand-induced docking sites for STAT1 are formed. Two inactive, monomeric STAT1 proteins then bind to these sites through their SH2 domain and are phosphorylated by the receptor-bound JAK kinases at tyrosine 701, near their C-terminus. After phosphorylation, the STAT1 proteins dissociate from the receptor and form a reciprocal homodimer, which then translocates to the nucleus, inducing the transcription of a set of IFN-stimulated genes (ISGs) via binding to the specific DNA elements residing within the ISGs promoters, and designated as either the interferon stimulated response element (ISRE) or the IFNγ activated site (GAS) (Darnell *et al*., 1994; Schindler and Darnell, 1995; Stark *et al*., 1998) (**Figure 1**).

A similar signalling pathway mechanism in which a ligand-induced, tyrosine-phosphorylated docking site on a receptor, and its association with the transcription factor STAT, has been found to mediate responses to other cytokines. As a result, this JAK–STAT signalling pathway mechanism is now the accepted paradigm that illustrates the important mechanism of how cytokine receptors are coupled to their specific STAT signalling systems.

The negative regulation of the IFNγ signalling pathway has also been defined recently by the discovery of a family of proteins known as SOCS/JAB/SSI, which are induced by IFN-γ (and also several other cytokines) and bind to and inhibit activated JAKs (Yasukawa *et al*., 2000). These discoveries have provided new insights into how JAK–STAT pathways are regulated in response to specific stimuli, and how they function in various tissues and environments.

Interferon-α/β Signalling Pathway

The common pathway for IFNα/β requires seven distinct proteins, which include two IFNα receptors, two JAKs, two STATs and the IRF-family transcription factor p48. The IFNα/β signalling pathways are comparatively illustrated in **Figure 2**, but the fine details of the mechanisms are lacking because information regarding the detailed interactions that play the crucial role for the pathway remain to be elucidated.

The IFNα/β receptors, designated IFNRAs, are composed of a multichain structure on both normal and malignant haematopoietic cells (Novick *et al*., 1994). All type I interferons, including IFNα, IFNβ, and IFNω, bind to the same receptors (Pestka *et al*., 1987). It has been clearly established that there are two distinct components comprising the type I IFNRAs, namely the IFNRA1 and

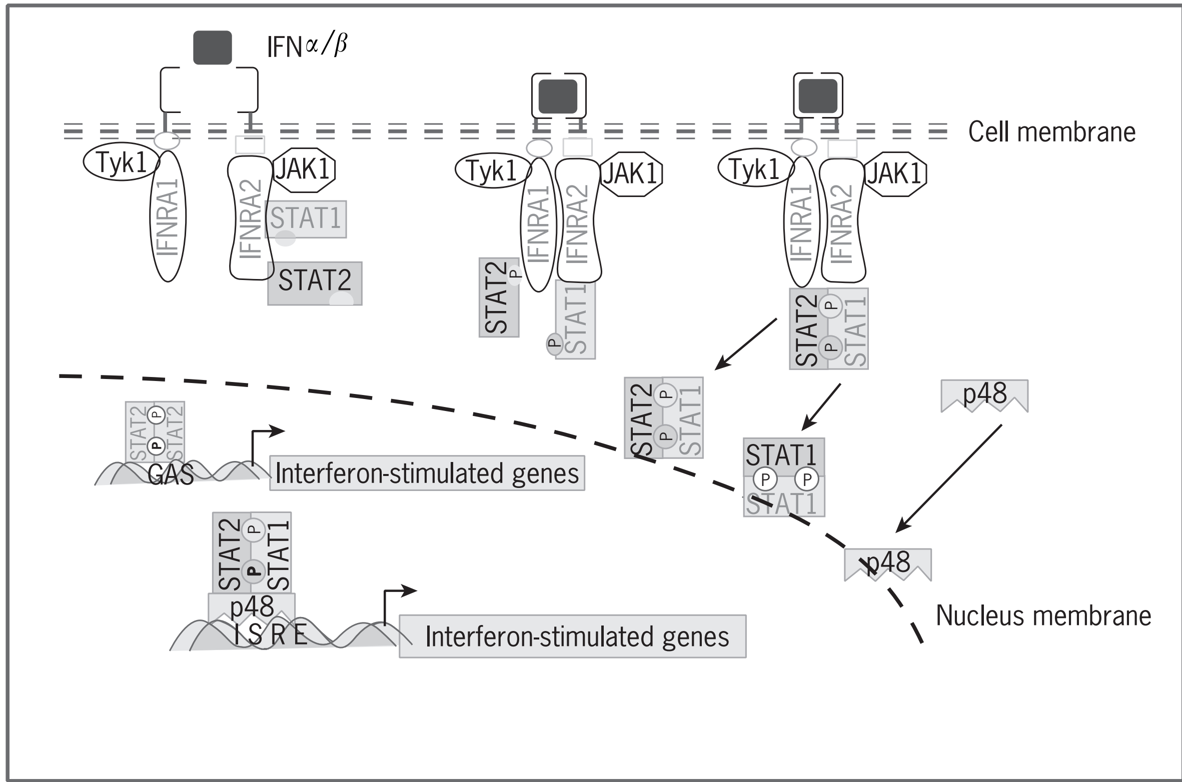


Figure 2 IFNα/β signal pathway.

the IFNRA2. IFNRA1 is a 110-kDa protein. The IFNRA2 subunit occurs in three different forms that are differentially spliced products of the same gene (Colamonici *et al*., 1994a; Novick *et al*., 1994), the soluble form of the extracellular domain, IFNRA2a; the alternatively spliced variant with a short cytoplasmic domain, IFNAR2b, which can have dominant negative activity and the only normal, fully functional form, IFNRA2c protein, with a relative molecular mass of 90–100 kDa. Neither IFNAR1 nor IFNAR2 alone binds to IFNα/β with the high affinity of the two-subunit combination (Cohen *et al*., 1995).

Two members of the Janus family of tyrosine kinases involved in the IFNα/β signalling pathway, Tyk2 and JAK1, are both constitutively associated with the IFNAR1 and IFNAR2c subunits, respectively (Colamonici *et al*., 1994c, 1995; Domanski *et al*., 1995, 1997). Upon the binding of IFNα/β to the IFNARs, the tyrosine kinases are activated and IFNAR1 and IFNAR2c are rapidly phosphorylated on tyrosines. Tyk2 also plays a role in stabilizing the IFNAR1 structure because the amount of IFNAR1 is low in Tyk2-null cells. However, the domains required for this role are different from those required to transduce the cytokine induced signal (Velazquez *et al*., 1995). The activation of tyrosine kinases results in tyrosine phosphorylation of two STAT proteins, STAT1 and STAT2, both preassociated with IFNAR2c in untreated cells (Fu, 1992; Darnell, 1997, 1998). When IFNAR1 is phosphorylated, the SH2 domain of STAT2 binds to it, followed by the phosphorylation of both STATs and subsequent dissociation of the phosphorylated heterodimer from the receptors. In addition to the STAT1–STAT2 heterodimer, much evidence also supports the IFNα/β-mediated induction of STAT1 homodimers (Darnell, 1997; Stark *et al*., 1998).

The phosphorylated STAT proteins form homo- and heterodimers and then translocate to the nucleus, associating with p48 to regulate gene transcription via binding to specific sequences present in the promoters of interferon-stimulated genes (ISGs) (Darnell, 1997; Ihle, 1996). The association of STAT1–STAT2 heterodimers with p48 to form the mature interferon-stimulated gene factor-3 (ISGF3) complex represents the major transcription factor complex formed in response to IFNα/β, and is required to drive the expression of most ISGs by binding to ISREs. While the STAT1–STAT2 heterodimer and STAT1 homodimer form in response to IFNα/β independently of p48, and can each drive the expression of a minority of ISGs, such as the IRF-1 gene, through GAS elements, the relative amount of STAT protein homo- and heterodimer and its complex formation with p48 depends on the level of p48, which can vary greatly among different cell types (Li *et al*., 1998). The functionality of p48 in the IFNα/β pathway seems to be that of an adaptor between STAT proteins and the DNA binding sites to redirect gene regulation and achieve the specificity of biological functions of IFNs.

TUMOUR NECROSIS FACTOR: RECEPTORS AND SIGNAL TRANSDUCTION PATHWAYS

The tumour necrosis factor (TNF) receptor superfamily comprises a group of cell surface receptors whose members generally bind ligands that are structurally related to TNF. The TNF ligand is structurally related to lymphotoxin-α (LTα; sometimes referred to as TNF-β), which is secreted from activated T cells, but binds the same receptors as TNF and has similar biological properties. TNF and LTα, however, are the prototype members of a large family of related proteins which includes CD30, CD40, Fas ligand and TRAIL ligand. TNF is a major physiological mediator of inflammation. It initiates the response to Gram-negative bacteria that produce lipopolysaccharide (LPS). TNF has been shown to induce fever, activate the coagulation system, induce hypoglycaemia, depress cardiac contractility, reduce vascular resistance, induce cachexia and activate the acute phase response in the liver.

Ironically, attempts to use TNF in the clinic actually predate its discovery and characterization. Towards the end of the nineteenth century, it was noticed that a small number of cancer patients experienced disease regression after suffering systemic bacterial infections. Subsequently a mixture of killed *Streptococcus pyogenes* and *Serratia marcescens* ('Coley's toxins') were administered to patients with advanced cancer, albeit with very occasional success. This approach became the treatment of choice for over three decades until superseded by advances in radiotherapy, chemotherapy and surgery. With hindsight, the most likely explanation for the results observed with Coley's toxins was the production of TNF, largely by macrophages in response to bacterial lipopolysaccharide present in the cell wall of Gram-negative bacteria such as *Serratia* sp.

TNF and Its Receptors

TNF and LTα are closely related homotrimeric proteins (32% identity). Human TNF is synthesized as a 233 amino acid glycoprotein, containing a long (76 residue) N-terminal leader sequence which anchors it to the cell membrane as a 25-kDa type II membrane protein. A secreted 17-kDa form of TNF is generated through the enzymatic cleavage of membrane-bound TNF by a metalloproteinase termed TNF-α-converting enzyme (TACE). Both soluble and membrane-bound forms of TNF are biologically active, although they have different affinities for the two TNF receptors, and probably as a consequence exhibit different biological properties (see below).

LTα differs from TNF in that it is synthesized as a secreted glycoprotein. Human LTα is synthesized as a 205 amino acid glycoprotein, which in native form exists as a 25-kDa homotrimer. As mentioned above, LTα can bind

both TNF receptors with affinities comparable to those of TNF, and has similar biological effects. However, a membrane-bound form of LT has been identified which consists of a heterotrimeric complex containing one LTα subunit noncovalently linked to two molecules of an LTα-related type II membrane protein termed LTβ. The LT$\alpha_1\beta_2$ heterotrimer (also known as mLT) is not cleaved by TACE and is thought to exist exclusively as a membrane-bound complex. mLT does not bind either of the two TNF receptors, but rather exerts its effects on another member of the TNF receptor superfamily, the lymphotoxin β receptor (LTβR). TNF and the two LT subunits are encoded by closely linked single copy genes, which are situated in the class III major histocompatibility locus, within a 25-kb region on the short arm of chromosome 6 in humans, at p21.

The two receptors for TNF (and LTα) are type I transmembrane glycoproteins designated TNFR1 (also termed p60 in humans, p55 in mice) and TNFR2 (also known as p80 in humans, p75 in mice). These receptors are characterized by cysteine-rich repeats of about 40 amino acids in their N-terminal extracellular domains. Each extracellular domain consists of three or four cysteine-rich regions containing four to six cysteines involved in intra-chain disulfide bonds. The cytoplasmic domains of these receptors have no obvious similarity to any known kinase and are thought to lack any intrinsic enzymatic activity. Signal transduction is therefore achieved by the recruitment and activation of adaptor proteins which recognize specific sequences in the cytoplasmic domains of these receptors. Recruitment of adaptor molecules activates a number of characteristic signalling pathways that can lead to a remarkably diverse set of cellular responses including differentiation, activation, release of inflammatory mediators and apoptosis. (See also chapter *Apoptosis*.)

Signal Transduction

Death Domains TRADD and FADD

The principal molecule thought to be involved in TNFR1 signal transduction is TNF receptor-associated death domain (TRADD), which is recruited to TNFR1 after activation by TNF. The interaction between TNFR1 and TRADD is mediated by the death domain, a motif found in both adaptor molecules such as TRADD and the cytoplasmic domains of the receptor itself (**Figure 3**). The binding of TRADD to TNFR1 leads to the recruitment and activation of numerous associated signalling molecules. TNF-induced apoptosis is generally thought to be achieved by the interaction of TRADD with FADD (Fas-associated death domain; also known as MORT1), a ~27-kDa protein which oligomerizes with TRADD through the death domains contained in both molecules. Recruitment of FADD is thought to activate a cascade of events which ultimately lead to apoptosis. This is brought about by the coordinate activation of several members of the caspase family. Caspases are cysteine aspartate proteases which are originally synthesized as zymogens, and are typically converted to their activated form by proteolytic cleavage, often by a distinct caspase upstream in the proteolytic cascade. Caspase-8, which is generally considered to be the apical caspase in the TNF and Fas pathways, is recruited to FADD in the activated complex, and is thought to be activated by self-cleavage induced by an increase in its local concentration. Cleaved caspase-8 can subsequently activate downstream caspases, notably caspase-3, and thereby induce apoptosis.

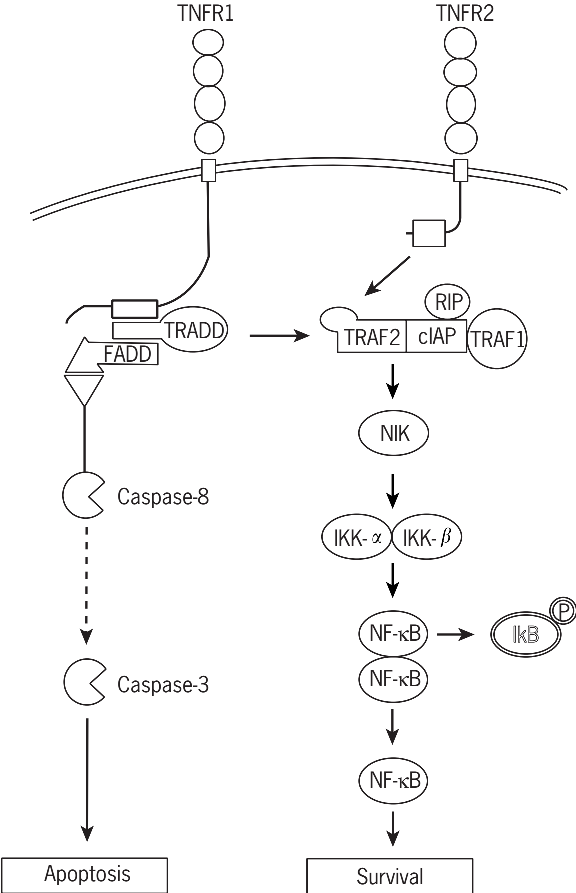


Figure 3 The role of the death domain- and death effector domain-containing molecules in signalling by TNFR1 and TNFR2.

TRAFs

The cytoplasmic domain of TNFR2 does not contain a death domain. In fact, under many circumstances, TNFR2 signalling can induce proliferation. This is thought to be mediated by the direct interaction of the intracellular signalling intermediate, TRAF2, with the cytoplasmic tail of TNFR2, which leads to the activation of NF-κB. In cells

which respond to TNFR2 signals by proliferating, no caspase activation is observed, and presumably no interaction of TRAF2 with TRADD. However, in cells which respond to TNFR2 signalling by undergoing apoptosis, there is mounting evidence that this is effected by a signal crosstalk mechanism with TNFR1, possibly by inducing expression of membrane-bound TNF or by affecting the stability of prosurvival proteins such as TRAF2 (Duckett and Thompson, 1997). As mentioned above, LTα has been shown to bind to both TNFR1 and TNFR2 with comparable affinities and biological outcomes. Membrane LT does not bind to either of these receptors, but binds exclusively to the LTβ receptor. In comparison with TNFR1 and TNFR2, the signal transduction pathways utilized by the LTβ receptor have been less well defined. Signalling by LTβR has been shown to induce apoptosis, although examination of the signalling, cytoplasmic tail has not revealed any obvious homology to the death domain. TRAF3 and TRAF5 have been shown to bind LTβR, but the role of TRAF3 in apoptosis induction is unclear.

TRAF2 is also a central component of the TNFR1 signalling complex, through a direct interaction with TRADD. Although TRAF2 lacks intrinsic enzymatic activity, it has been shown to bind several serine–threonine kinases, including NIK (NF-κB-inducing kinase), RIP (receptor interacting protein) and GCK (germinal centre kinase). Through the recruitment of these kinases, TRAF2 is thought to induce the activation of several transcription factors, particularly NF-κB (nuclear factor-κB), as well as downstream kinases involved in stress responses, notably c-Jun N-terminal kinase (JNK), which are crucial effectors of the TNFR1-mediated proinflammatory reaction. TRADD is also thought to associate with the serine–threonine kinase RIP, and RIP has been shown to interact with another protein, RAIDD (RIP associated Ich-1/CED3 homologous protein with death domain; also known as CRADD), causing the recruitment and activation of caspase-2 and the induction of apoptosis. However, the RAIDD/CRADD pathway is not thought to be the major signalling pathway utilized by TNFR1 to induce apoptosis. Despite the well-defined ability of TNFR1 signalling to induce cell death, the majority of normal cells do not respond to TNF by undergoing apoptosis. It has been proposed that this paradoxical situation can be accounted for by the activation of NF-κB, which has been shown to induce the expression of a number of antiapoptotic proteins. The best described of these are (i) A20, a zinc finger-containing molecule, (ii) A1/BflI, a Bcl-2 homologue, and (iii) c-IAP1/c-IAP2, members of the IAP (inhibitor of apoptosis) family.

Many key signalling intermediates responsible for the induction of NF-κB by the TNF receptors have been identified. NF-κB transcription factors are sequestered in the cytoplasm of cells by a protein called IκB (inhibitor of NF-κB). Phosphorylation of IκB leads to its degradation via ubiquitination by the 26S proteasome. The heterodimeric NF-κB subunits then translocate to the nucleus where they regulate expression of a wide variety of genes involved in inflammatory responses. A complex of two kinases and a regulatory protein are responsible for IκB phosphorylation. The kinases are termed IκB kinases α (IKKα) and β (IKKβ) and are constitutively associated with the regulatory protein IKKγ or NEMO. The serine–threonine kinase NIK, initially identified through its ability to associated with TRAF2, is thought to activate the IKK complex. A naturally occurring mouse mutation termed alymphoplasia (*aly*) is the result of a point mutation of NIK. *Aly/aly−/−* mice lack lymph nodes and Peyer patches, and also exhibit disorganized splenic and thymic structures.

Other Pathways

Alternative models have been proposed to account for the diverse range of outcomes following TNFR1 activation. Notably, SODD (silencer of death domains) was identifed in the basis of its ability to bind to the death domain of TNFR1. SODD is found in the TNFR1 receptor complex before receptor activation, but then dissociates from the receptor after ligand binding. It is thought that the SODD preassociation with TNFR1 may prevent spontaneous signalling by death domain-containing receptors. Neutral sphingomyelinase (N-Smase) activation is thought to mediate some of the inflammatory and proliferative responses to TNF through the activation of ERK and phospholipase A2. Other factors have been identifed which bind distinct sites in the cytoplasmic tail of TNFR1, such as FAN (factor associated with N-Smase activation), which is thought to couple the TNFR to activation of neutral sphingomyelinase activity, which in turn results in the production of ceramide and is thereby thought to lead to the activation of MAP kinases.

CHEMOKINE RECEPTOR SIGNAL TRANSDUCTION

Chemoattractant cytokines are typically <15-kDa proteins which are secreted by many tissue and cell types. Chemokines were recently reviewed (Zlotnik *et al*., 1999). The classical chemokine-induced biological activity is leucocyte migration, but as the field has matured other chemokine-mediated physiological functions have been identified, including regulation of vascularization in embryogenesis, lymphocyte maturation, cellular activation, regulation of angiogenesis and apoptosis. Although these additional functions have been demonstrated, the components of the distinct signalling cascades are poorly characterized, so this section will focus on chemokine-induced chemotaxis, cell activation and apoptosis.

Seven Transmembrane G-protein-coupled Receptors (GPCRs)

Chemokines bind to and activate seven pass transmembrane G-protein-coupled receptors (GPCRs) that are structurally similar to the rhodopsin (type A) subfamily. The N-terminus of the receptor, which is also the first extracellular domain, is essential for high-affinity ligand binding. The closely packed position of the seven transmembrane domains are maintained by disulfide bonds between the extracellular domains. The disulfide bonding is needed for efficient chemokine-induced signalling but is not necessary for HIV-1 coreceptor activity. The chemokine receptor signal is dependent on a ligand-induced dynamic change in the receptor that results in an increased affinity for Gi and Gq heterotrimeric G-proteins. Fine regulation of the GPCR signal occurs at the membrane where increased phosphatidylcholine in the lipid bilayer enhances GTPase activity of the G-proteins and is essential for the activity of G-protein coupled receptor kinase(s) (GRKs).

Heterotrimeric G-proteins

Chemokine-induced chemotaxis is inhibited by pertussis toxin, indicating that chemokine receptors activate trimeric G proteins in the Gi subfamily. However, activation of phospholipase C (PLC), intracellular calcium (Ca^{2+}) mobilization and cellular exocytosis can be mediated through pertussis toxin-insensitive Gq proteins. The type of cell expressing the GPCR and its activation state regulates which G-protein couples to the receptor, such that chemokine receptors have been reported to couple to several classes of G-proteins.

Heterotrimeric G-proteins are composed of α-, β- and γ-subunits. In the resting G-protein, the α-subunit is bound to GDP. Once activated, GDP is exchanged for GTP and the α-subunit separates from the β–γ-subunits. The α-subunits are apparently necessary for regulation of the GPCR function. Regulators of G-protein signal (RGS) are widely expressed GTPase-activating proteins that contain a 130 amino acid domain that binds Gα-GTP subunits accelerating the hydrolysis to Gα-GDP and blocking Gα interaction with PLC. Additionally, some Gα subunits are substrates for protein kinase C (PKC) resulting in autoregulation of this G-protein-mediated signal. Gαi subunits were shown to be nonessential in chemokine-induced chemotaxis. These data indicated that any Gα linked to β–γ-subunits, which are essential, could transmit a chemotactic signal. In addition to transmitting the chemotactic signal, β–γ-subunits participate in signal component receptor docking and cell activation.

Heterotrimeric G-proteins interact with several intracellular domains of GPCRs found in the cytoplasmic tail and second and third intracellular loops. The β–γ-subunits were shown to guide GRK2 to its phosphorylation site on a GPCR, suggesting that both the correct membrane lipid composition and G-protein components are needed for GRK regulation of GPCR signal. The G-protein β–γ-subunits directly bind to and activate Raf-1, phosphoinositide 3-kinase gamma (PI3Kγ) and some small GTPases. Raf-1 is a serine/threonine kinase that links the mitogen-activated protein kinase cascade to tyrosine kinase-dependent growth factor receptors. Mutagenesis studies showed that β–γ-subunits bind to Raf-1 with an affinity similar to that between β–γ-subunits and GRKs, suggesting that there maybe a competition between receptor inactivation by GRKs and the mitogenic signal. Stromal-derived factor-1 (SDF-1, also known as CXCL12) is a CXC chemokine that activates the CXCR4 receptor. Recently it was shown that SDF-1-induced chemotaxis and extracellular signal-related kinase (ERK) activation was inhibited by PI3K inhibitor treatment of T lymphocytes. These data show a link between G-protein β–γ-subunits and PI3K and suggest an order of activation.

Mutagenesis of G-protein β-subunits resulted in inappropriate organisation of cellular cytoskeleton (Peracino *et al*., 1998), and additional studies have shown that small GTPases of the Rho family are essential for chemotaxis and sequestration of β–γ-subunits leads to rearrangement of the actin cytoskeleton, suggesting a link between β-subunits and these small GTPases.

Phospholipase

An approach to demonstrate the role of PLC in the GPCR signal cascade used mice lacking PLC-β2 and -β3. Neutrophils from animals lacking both PLC-β2 and -β3 did not produce inositol triphosphate (IP_3), flux calcium or superoxide in response to chemokines or chemoattractants. Animals lacking only PLC-β2 clearly showed reduced both IP_3 and Ca^{2+} flux in response to interleukin-8 (IL-8 also known as CXCL8) and macrophage inflammatory protein 1 beta (MIP1β, also known as CCL4), but not to the same extent, suggesting that both PLC isoforms participate in signal transduction in neutrophils. In contrast to the calcium flux signal, the chemotaxic response of the neutrophils from animals lacking both PLC-β2 and -β3 was not reduced; rather, there was an enhanced chemotactic response to IL-8. The PLC-β2- and -β3-deficient animals failed to activate PKC in response to chemoattractants, indicating that the PKC pathway is linked to the Gα and not the Gβ–γ signal. Additionally, PLC-deficient animals failed to phosphorylate the mitogen-activated protein kinase (MAPK) c-JUN N-terminal kinase (JNK). However, PLC deficiency had no effect on Rac activation, suggesting that cytoskeletal modification is a separate signal in neutrophils.

In addition to activation of PLC, there is strong evidence that phospholipase D (PLD) is also activated by chemokines. The position of PLD in the chemokine signal

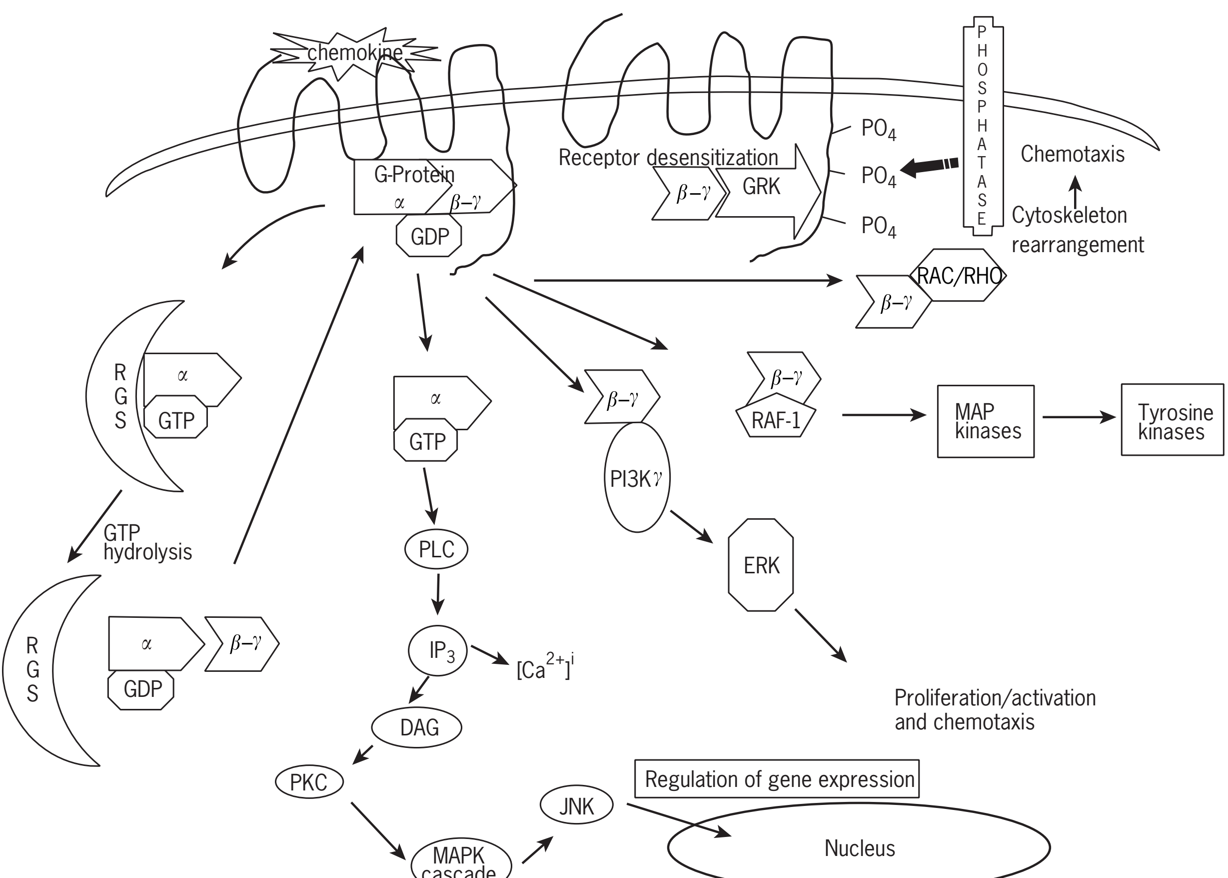


Figure 4 G-protein-coupled chemokine signalling.

cascade is ambiguous because diacylglycerol (DAG) can be interconverted to the lipid hydrolysis product of PLD, phosphatidic acid (PA), suggesting that PLC activity may regulate PLD activity. Additionally, PLD is activated by Ras and Rho family members and PKC. A function for PA in chemokine-induced cell activation has not been identified, but PA is strongly associated with cell vesicle transport, suggesting that PLD may act late in the chemokine-induced cascade by regulating receptor localization to the membrane (**Figure 4**).

Further, PA is in the pathway that leads to respiratory burst (Lennartz, 1999), suggesting that PLD activation by chemokines maybe participate in chemokine-induced NADPH production.

Phosphoinositide 3-Kinase Gamma (PI3Kγ)

PI3Ks have been implicated in many cellular responses, including, proliferation, apoptosis, adhesion and chemotaxis. Recently, a number of groups generated mice deficient in PI3Kγ. All three groups observed a severe reduction ($\leq$85%) in chemokine-induced chemotaxis, but there was some activity left, indicating that Gβ–γ may link to other intracellular components and induce chemotaxis. The least effected was MIP-5-induced chemotaxis (Hirsch *et al*., 2000). MIP-5 (also known as HCC-2 or CCL-15) has a unique six-cysteine structure. The effect this chemokine has on other intracellular signalling components has not been evaluated, but mutational analysis indicated that the third set of disulfide bonds were not needed to induce chemotaxis. PI3Kγ deficiency had no effect on chemoattractant-induced Rac activation, actin polymerization or calcium flux. PI3Kγ deficiency had a profound inhibitory effect on chemoattractant-induced activation of PKB, ERK1 or ERK2. The activation of ERK1 and ERK2 by PI3Kγ directly links the chemokine GPCR signal to both proliferation and activation signals. Taken together, these data indicate that PI3Kγ is a major component in chemokine-induced chemotaxis and cell activation, but not the only component.

Focal Adhesion Kinases and Tyrosine Phosphatases

Several tyrosine kinases are activated during chemotaxis, but whether the effect is mediated through a GPCR pathway or by an adhesion molecule-activated signal pathway

is still unclear. Focal adhesion kinase (FAK) activity is clearly stimulated by RANTES (regulated on activation of normal T cell expression and secreted, also known as CCL5) binding to CC chemokine receptor-5 (CCR5).

SDF-1 and stem cell factor appear to activate related adhesion focal tyrosine kinase (RAFTK), a FAK homology expressed in some leucocytes. The FAK family has been shown to activate or be activated by some MAPK and Rho family members. Thus, although there is no disagreement that focal adhesion kinases are activated in cell migration, their role in chemokine-induced cell migration and activation is unclear.

Studies using tyrosine phosphatase, SHIP-1 and SHP-1-deficient mice showed enhanced chemotaxis and reduced suppression of proliferation in response to chemokines. Activation of CD45, a cell surface glycoprotein with tyrosine phosphatase activity, decreased the cell surface expression of IL-8 receptors and IL-8-induced calcium flux. While it is tempting to theorize a role for these phosphatases in G-protein-mediated chemokine signalling, more research is needed to connect these signals clearly. Recent reviews suggest that the components of adhesion, selectins and integrin receptors, have bidirectional signalling cascades, and that the activation of focal adhesion kinases and tyrosine phosphatases may be an 'outside in' signal.

Apoptosis

Studies of cell death in HIV-1-infected cells showed that CXCR4 is required for this apoptotic signal and the signal is G-protein independent. Pertussis toxin, PI3K inhibitor (wortmannin), ERK pathway inhibitors and MAPK pathway inhibitors had no effect on CXCR4-mediated HIV-1-induced lymphocyte apoptosis. CCR5 does signal through ERK and MAPK when activated by HIV-1. Gp120-independent virus–cell fusion and recombinant gp120 did not induce apoptosis, but, cell surface-expressed gp120 did.

Deletion of the cytoplasmic tail of CXCR4 blocked ligand-induced receptor internalization, but did not block HIV-1-induced apoptosis. Since none of the well-characterized chemokine signal components appear to be necessary for the CXCR4-mediated apoptosis signal, what is? Caspase 3 activity correlates with CXCR4-mediated cell death. These data indicate that HIV-1 induces CXCR4-dependent apoptosis by activating caspase 3.

Musings

The chemokine-induced cellular signal is complex and still poorly characterized. The activation of G-protein subunit-mediated signals is clear, but other non-G-protein-mediated signals still require clarification. Receptor desensitization and resensitization are a great mystery, with few well-characterized components, other than serine–threonine phosphorylation of the GPCR carboxyl tail. Several groups have shown tyrosine kinase activation in response to chemokines; however, because the cell shape change and activation of these kinases parallel each other, an integrin receptor-mediated signal cannot be ruled out. Further, activation of Janus kinases (JAKs) by chemokines has been demonstrated in human embryonic kidney and lymphocyte cell lines. These studies took place in membrane raft microdomains, so that other components of the rafts may have contributed to the activation. Characterization of chemokine-induced activation of signal transducers and activators of transcription (STATs) has begun. SDF-1, RANTES and MIP1α have been shown to stimulate STAT phosphorylation. However, it remains to determine the mediator, be it binding to a domain on the chemokine receptor or JAK activation. Further structural analysis of chemokine receptors is needed to link conclusive tyrosine kinases and STAT activation to chemokine receptors. After 12 years, our understanding of chemokine-mediated signal transduction is almost ready to enter adolescence, but a long way from a mature field.

REFERENCES

Ahmad, M., *et al*. (1997). CRADD, a novel human apoptotic adaptor molecule for caspase-2, and FasL/tumor necrosis factor receptor-interacting protein RIP. *Cancer Research*, **57**, 615–619.

Arai, H. and Charo, I. F. (1996). Differential regulation of G-protein-mediated signaling by chemokine receptors. *Journal of Biological Chemistry*, **271**, 21814–21819.

Bach, E. A., *et al*. (1997). The IFN gamma receptor a paradigm for cytokine receptor signaling. *Annual Review of Immunology*, **15**, 563–591.

Bacon, K. B., *et al*. (1996). RANTES induces tyrosine kinase activity of stably complexed p125FAK and ZAP-70 in human T cells. *Journal of Experimental Medicine*, **184**, 873–882.

Beutler, B. A. (1999). The role of tumor necrosis factor in health and disease. *Journal of Rheumatology*, **26**, Suppl. 57, 16–21.

Blanco, J., *et al*. (1999). The implication of the chemokine receptor CXCR4 in HIV-1 envelope protein-induced apoptosis is independent of the G protein-mediated signalling. *Aids*, **13**, 909–917.

Bluyssen, H. A., *et al*. (1995). Combinatorial association and abundance of components of interferon-stimulated gene factor 3 dictate the selectivity of interferon responses. *Proceedings of the National Academy of Sciences of the USA*, **92**, 5645–5649.

Cohen, B., *et al*. (1995). Ligand-induced association of the type I interferon receptor components. *Molecular and Cellular Biology*, **15**, 4208–4214.

Colamonici, O. R., *et al*. (1994a). Interferon alpha (IFN alpha) signaling in cells expressing the variant form of the type I IFN receptor. *Journal of Biological Chemistry*, **269**, 5660–5665.

Colamonici, O. R., *et al*. (1994b). Ligand-independent anti-oncogenic activity of the alpha subunit of the type I interferon receptor. *Journal of Biological Chemistry*, **269**, 27275–27279.

Colamonici, O. R., *et al*. (1994c). p135tyk2, an interferon-alpha-activated tyrosine kinase, is physically associated with an interferon-alpha receptor. *Journal of Biological Chemistry*, **269**, 3518–3522.

Colamonici, O. R., *et al*. (1995). Transmembrane signaling by the alpha subunit of the type I interferon receptor is essential for activation of the JAK kinases and the transcriptional factor ISGF3. *Journal of Biological Chemistry*, **270**, 8188–8193.

Coppolino, M. G. and Dedhar, S. (2000). Bi-directional signal transduction by integrin receptors. *International Journal of Biochemical Cell Biology*, **32**, 171–188.

Darnell, J. E., Jr, *et al*. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, **264**, 1415–1421.

Darnell, J. E., Jr (1997). STATs and gene regulation. *Science*, **277**, 1630–1635.

Darnell, J. E., Jr (1998). Studies of IFN-induced transcriptional activation uncover the Jak–Stat pathway. *Journal of Interferon and Cytokine Research*, **18**, 549–554.

Domanski, P., *et al*. (1997). A region of the beta subunit of the interferon alpha receptor different from box 1 interacts with Jak1 and is sufficient to activate the Jak–Stat pathway and induce an antiviral state. *Journal of Biological Chemistry*, **272**, 26388–26393.

Domanski, P., *et al*. (1995). Homodimerization and intermolecular tyrosine phosphorylation of the Tyk-2 tyrosine kinase. *FEBS Letters*, **374**, 317–322.

Duckett, C. S. and Thompson, C. B. (1997). CD30-dependent degradation of TRAF2 implications for negative regulation of TRAF signaling and the control of cell survival. *Genes and Development*, **11**, 2810–2821.

Fu, X. Y. (1992). A transcription factor with SH2 and SH3 domains is directly activated by an interferon alpha-induced cytoplasmic protein tyrosine kinase(s). *Cell*, **70**, 323–335.

Gonzalez-Amaro, R. and Sanchez-Madrid, F. (1999). Cell adhesion molecules: selectins and integrins. *Critical Reviews in Immunology*, **19**, 389–429.

Gouilleux-Gruart, V., *et al*. (1997). Activated Stat related transcription factors in acute leukemia. *Leukemia Lymphoma*, **28**, 83–88.

Greenlund, A. C., *et al*. (1995). Stat recruitment by tyrosine-phosphorylated cytokine receptorsan ordered reversible affinity-driven process. *Immunity*, **2**, 677–687.

Haque, S. J. and Williams, B. R. (1994). Identification and characterization of an interferon (IFN)-stimulated response element-IFN-stimulated gene factor 3-independent signaling pathway for IFN-alpha. *Journal of Biological Chemistry*, **269**, 19523–19529.

Harada, H., *et al*. (1996). Regulation of IFN-alpha/beta genes-evidence for a dual function of the transcription factor complex ISGF3 in the production and action of IFN-alpha/beta. *Genes and Cells*, **1**, 995–1005.

Hayakawa, F., *et al*. (1998). Differential constitutive activation between STAT-related proteins and MAP kinase in primary acute myelogenous leukaemia. *British Journal of Haematology*, **101**, 521–528.

Hemmi, S., *et al*. (1994). A novel member of the interferon receptor family complements functionality of the murine interferon gamma receptor in human cells. *Cell*, **76**, 803–810.

Hirsch, E., *et al*. (2000). Central role for G protein-coupled phosphoinositide 3-kinase gamma in inflammation. *Science*, **287**, 1049–1053.

Horvath, C. M., *et al*. (1995). A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. *Genes and Development*, **9**, 984–994.

Howard, O. M., *et al*. (1999). Naturally occurring CCR5 extracellular and transmembrane domain variants affect HIV-1 Co-receptor and ligand binding function. *Journal of Biological Chemistry*, **274**, 16228–16234.

Ihle, J. N. (1996). STATs signal transducers and activators of transcription. *Cell*, **84**, 331–334.

Ihle, J. N. (1995). The Janus protein tyrosine kinase family and its role in cytokine signaling. *Advances in Immunology*, **60**, 1–35.

John, J., *et al*. (1991). Isolation and characterization of a new mutant human cell line unresponsive to alpha and beta interferons. *Molecular and Cellular Biology*, **11**, 4189–4195.

Kaul, M. and Lipton, S. A. (1999). Chemokines and activated macrophages in HIV gp120-induced neuronal apoptosis. *Proceedings of the National Academy of Sciences of the USA*, **96**, 8212–8216.

Kotenko, S. V., *et al*. (1995). Interaction between the components of the interferon gamma receptor complex. *Journal of Biological Chemistry*, **270**, 20915–20921.

Lennartz, M. R. (1999). Phospholipases and phagocytosis: the role of phospholipid-derived second messengers in phagocytosis. *International Journal of Biochemical Cell Biology*, **31**, 415–430.

Li, X., *et al*. (1998). Cooperative binding of Stat1–2 heterodimers and ISGF3 to tandem DNA elements. *Biochimie*, **80**, 703–710.

Li, X., *et al*. (1996). Formation of STAT1–STAT2 heterodimers and their role in the activation of IRF-1 gene transcription by interferon-alpha. *Journal of Biological Chemistry*, **271**, 5790–5794.

Lim, L., *et al*. (1996). Regulation of phosphorylation pathways by p21 GTPases. The p21 Ras-related Rho subfamily and its role in phosphorylation signalling pathways. *European Journal of Biochemistry*, **242**, 171–185.

Lopez-Ilasaca, M. (1998). Signaling from G-protein-coupled receptors to mitogen-activated protein (MAP)-kinase cascades. *Biochemical Pharmacology*, **56**, 269–277.

Ma, Q., *et al*. (1998). Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proceedings of the National Academy of Sciences of the USA*, **95**, 9448–9453.

Matsumoto, M., *et al*. (1999). Activation of the transcription factor ISGF3 by interferon-gamma. *Biological Chemistry*, **380**, 699–703.

Monteclaro, F. S. and Charo, I. F. (1997). The amino-terminal domain of CCR2 is both necessary and sufficient for high affinity binding of monocyte chemoattractant protein 1. Receptor activation by a pseudo-tethered ligand. *Journal of Biological Chemistry*, **272**, 23186–23190.

Moriggl, R., *et al*. (1999). Stat5 is required for IL-2-induced cell cycle progression of peripheral T cells. *Immunity*, **10**, 249–259.

Noguchi, M., *et al*. (1993). Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. *Cell*, **73**, 147–157.

Nosaka, T., *et al*. (1995). Defective lymphoid development in mice lacking Jak3. *Science*, **270**, 800–802.

Novick, D., *et al*. (1994). The human interferon alpha/beta receptor characterization and molecular cloning. *Cell*, **77**, 391–400.

Pease, J. E., *et al*. (1998). The N-terminal extracellular segments of the chemokine receptors CCR1 and CCR3 are determinants for MIP-1alpha and eotaxin binding, respectively, but a second domain is essential for efficient receptor activation. *Journal of Biological Chemistry*, **273**, 19972–19976.

Pellegrini, S., *et al*. (1989). Use of a selectable marker regulated by alpha interferon to obtain mutations in the signaling pathway. *Molecular and Cellular Biology*, **9**, 4605–4612.

Peracino, B., *et al*. (1998). G protein beta subunit-null mutants are impaired in phagocytosis and chemotaxis due to inappropriate regulation of the actin cytoskeleton. *Journal of Cell Biology*, **141**, 1529–1537.

Pestka, S., *et al*. (1987). Interferons and their actions. *Annual Review of Biochemistry*, **56**, 727–777.

Pfeffer, L. M., *et al*. (1997). The short form of the interferon alpha/beta receptor chain 2 acts as a dominant negative for type I interferon action. *Journal of Biological Chemistry*, **272**, 11002–11005.

Platanias, L. C. and Fish, E. N. (1999). Signaling pathways activated by interferons. *Experimental Hematology*, **27**, 1583–1592.

Platanias, L. C., *et al*. (1992). Expression of the IFN alpha receptor in hairy cell leukaemia. *British Journal of Haematology*, **82**, 541–546.

Porter, A. C., *et al*. (1988). Interferon response element of the human gene 6–16. *EMBO Journal*, **7**, 85–92.

Puck, J. M., *et al*. (1993). The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X-linked severe combined immunodeficiency, SCIDX1. *Human Molecular Genetics*, **2**, 1099–1104.

Ram, P. A. and Waxman, D. J. (1997). Interaction of growth hormone-activated STATs with SH2-containing phosphotyrosine phosphatase SHP-1 and nuclear JAK2 tyrosine kinase. *Journal of Biological Chemistry*, **272**, 17694–17702.

Rane, S. G. and Reddy, E. P. (1994). JAK3: a novel JAK kinase associated with terminal differentiation of hematopoietic cells. *Oncogene*, **9**, 2415–2423.

Ray, M., *et al*. (2000). Inhibition of interferon-gamma signaling by *Leishmania donovani*. *Journal of Infectious Diseases*, **181**, 1121–1128.

Richardson, R. M., *et al*. (1998). Multiple signaling pathways of human interleukin-8 receptor A. Independent regulation by phosphorylation. *Journal of Biological Chemistry*, **273**, 10690–10695.

Russell-Harde, D., *et al*. (1995). Reconstitution of a high affinity binding site for type I interferons. *Journal of Biological Chemistry*, **270**, 26033–26036.

Russell-Harde, D., *et al*. (2000). Role of the intracellular domain of the human type I interferon receptor 2 chain (IFNAR2c) in interferon signaling. Expression of IFNAR2c truncation mutants in U5A cells. *Journal of Biological Chemistry*, **275**, 23981–23985.

Schindler, C. and Darnell, J. E., Jr (1995). Transcriptional responses to polypeptide ligands in the JAK–STAT pathway. *Annual Review of Biochemistry*, **64**, 621–651.

Seder, R. A. (1994). Acquisition of lymphokine-producing phenotype by CD4 + T cells. *Journal of Allergy and Clinical Immunology*, **94**, 1195–1202.

Shuai, K., *et al*. (1996). Constitutive activation of STAT5 by the BCR-ABL oncogene in chronic myelogenous leukemia. *Oncogene*, **13**, 247–254.

Singer, W. D., Brown, H. A. and Sternweis, P. C. (1997). Regulation of eukaryotic phosphatidylinositol-specific phospholipase C and phospholipase D. *Annual Review of Biochemistry*, **66**, 475–509.

Soh, J., *et al*. (1994). Identification and sequence of an accessory factor required for activation of the human interferon gamma receptor. *Cell*, **76**, 793–802.

Stark, G. R., *et al*. (1998). How cells respond to interferons. *Annual Review of Biochemistry*, **67**, 227–264.

Takahashi, M., *et al*. (1997). Mechanotransduction in endothelial cells: temporal signaling events in response to shear stress. *Journal of Vascular Research*, **34**, 212–219.

Takeda, K., *et al*. (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proceedings of the National Academy of Sciences of the USA*, **94**, 3801–3804.

Takeda, K., *et al*. (1996). Essential role of Stat6 in IL-4 signalling. *Nature*, **380**, 627–630.

Tau, G. and Rothman, P. (1999). Biologic functions of the IFN-gamma receptors. *Allergy*, **54**, 1233–1251.

Tournamille, C., *et al*. (1997). Close association of the first and fourth extracellular domains of the Duffy antigen/receptor for chemokines by a disulfide bond is required for ligand binding. *Journal of Biological Chemistry*, **272**, 16274–16280.

Trinchieri, G. and Scott, P. (1995). Interleukin-12a proinflammatory cytokine with immunoregulatory functions [editorial]. *Research in Immunology*, **146**, 423–431.

Uze, G., *et al*. (1990). Genetic transfer of a functional human interferon alpha receptor into mouse cells cloning and expression of its cDNA. *Cell*, **60**, 225–234.

Uze, G., *et al*. (1995). Alpha and beta interferons and their receptor and their friends and relations. *Journal of Interferon and Cytokine Research*, **15**, 3–26.

Velazquez, L., *et al*. (1995). Distinct domains of the protein tyrosine kinase tyk2 required for binding of interferon-alpha/beta and for signal transduction. *Journal of Biological Chemistry*, **270**, 3327–3334.

Wong, M. and Fish, E. N. (1998). RANTES and MIP-1alpha activate stats in T cells. *Journal of Biological Chemistry*, **273**, 309–314.

Wu, G., *et al*. (1998). Receptor docking sites for G-protein beta-gamma subunits. Implications for signal regulation. *Journal of Biological Chemistry*, **273**, 7197–7200.

Wu, M. Y., *et al*. (1999). The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell death in HeLa cells. *Journal of Biological Chemistry*, **274**, 11868–11873.

Xia, Z., *et al*. (1998). Expression of signal transducers and activators of transcription proteins in acute myeloid leukemia blasts. *Cancer Research*, **58**, 3173–3180.

Yamamoto, K., *et al*. (1997). cDNA cloning, expression and chromosome mapping of the human STAT4 gene: both STAT4 and STAT1 genes are mapped to 2q32.2q32.3. *Cytogenetics and Cell Genetics*, **77**, 207–210.

Yasukawa, H., *et al*. (2000). Negative regulation of cytokine signaling pathways. *Annual Review of Immunology*, **18**, 143–164.

Young, H. A. and Hardy, K. J. (1995). Role of interferon-gamma in immune cell regulation. *Journal of Leukocyte Biology*, **58**, 373–381.

Zheng, B., *et al*. (1999a). Divergence of RGS proteins: evidence for the existence of six mammalian RGS subfamilies. *Trends in Biochemical Sciences*, **24**, 411–414.

Zheng, J., *et al*. (1999b). Intracellular CXCR4 signaling, neuronal apoptosis and neuropathogenic mechanisms of HIV-1-associated dementia. *Journal of Neuroimmunology*, **98**, 185–200.

Zhou, N., *et al*. (2000). Molecular modeling and site-directed mutagenesis of CCR5 reveal residues critical for chemokine binding and signal transduction. *European Journal of Immunology*, **30**, 164–173.

Zlotnik, A., *et al*. (1999). Recent advances in chemokines and chemokine receptors. *Critical Reviews in Immunology*, **19**, 1–47.

Chapter 12

Signalling by Tyrosine Kinases

Kermit L. Carraway and Coralie A. Carothers Carraway
University of Miami School of Medicine, Miami, FL, USA

CONTENTS

- Introduction
- Receptor Tyrosine Kinases
- Binary Receptors
- Nonreceptor Tyrosine Kinase Mechanisms
- Signalling Pathways and Cellular Functions of Tyrosine Kinases
- Tyrosine Kinases and Cancer
- Overview

INTRODUCTION

Historical Perspective

Few molecules have been more closely linked to cancer than the protein tyrosine kinases. Tyrosine phosphorylation and the tyrosine kinase enzymes which create phosphotyrosine residues in their substrate proteins were discovered during studies of the oncogenic factors in tumour viruses. The identification of v-Src as the transforming factor of the Rous sarcoma virus and the recognition of its activity as a protein tyrosine kinase conceptualized the oncogene theory of tumorigenesis, creating a model which has dominated much of cancer research over the past two decades. The importance of the work on viral oncogenes in the early research on tyrosine phosphorylation is shown by an accounting of landmark events in this area (**Table 1**) (Hunter, 1998). The discovery of a cellular counterpart (c-Src) of the viral Src (v-Src) tyrosine kinase indicated a much broader role for these enzymes and suggested a major role for c-Src and tyrosine phosphorylation in mediating normal cell behaviours. For example, c-Src has been implicated in such diverse cell functions as platelet aggregation, cell cycle control and cell motility. The diversity of cellular functions of tyrosine kinases was further indicated by early observations that both the epidermal growth factor (EGF) receptor, important in epithelial cell growth, and the insulin receptor, a key component in metabolic regulation, are tyrosine kinases. Another tyrosine kinase receptor, called Sevenless for its role in determining cell fate of a particular cell type, was found to be necessary for normal eye development in the fruit fly (*Drosophila*). These and many other discoveries have demonstrated the critical role that the tyrosine kinases and tyrosine phosphorylation play in normal cell regulation and communication and in developmental biology.

Evolutionarily, tyrosine kinases are primarily, if not exclusively, a product of eukaryotic cells. The emergence of protein tyrosine kinases with the appearance of multicellular

Table 1 Important discoveries in history of tyrosine phosphorylation

Year	Discovery
1979	Tyrosine phosphorylation of polyoma tumour virus middle T antigen
1980	Protein tyrosine kinase activities of v-Src, c-Src, v-Abl and EGF receptor
1981	Insulin-stimulated protein tyrosine kinase activity of insulin receptor
1982	Sequence similarity of v-Src to cAMP-dependent protein kinase catalytic subunit
	c-Abl gene rearrangement in chronic myelogenous leukaemia
1983	Polyoma middle T association with and activation of c-Src
1984	v-ErbB oncogene derived from EGF receptor
1985	Negative regulation of c-Src by tyrosine phosphorylation
1986	Neu oncogene as EGF receptor family member with activating point mutation
1987	*Drosophila* Sevenless is receptor tyrosine kinase
1988	Acetylcholine receptor regulated by tyrosine phosphorylation
1989	Cell cycle regulatory kinase negatively regulated by Tyr phosphorylation
1990	Src homology-2 domains bind phosphotyrosines
1992	Individual receptor Tyr phosphorylation sites bind distinct SH2-containing proteins
	STAT transcription factors activated by Tyr phosphorylation

(Adapted from Hunter, 1998.)

(metazoan) organisms is indicative of their importance in cell communication and organisation. Sequence information from the Human Genome Project indicates that there are 90 tyrosine kinases encoded in the human genome.

The Kinase Superfamily

Tyrosine kinases are members of a much larger family of protein kinases (Hunter, 1998), which can be categorized by two classifications, one based on specificity for the target amino acid and the other on structure and cellular localization. The major specificity classes are the serine or threonine-specific (Ser/Thr) kinases and the tyrosine-specific (Tyr) kinases which catalyse the phosphorylation of serine and threonine or tyrosine residues, respectively. In addition, a few mixed function kinases, which catalyse both Ser/Thr and Tyr phosphorylation, have been described. The discovery of serine/threonine kinases and their importance in regulating metabolic pathways preceded the discovery of tyrosine kinases by about two decades. The regulation of gluconeogenesis by phosphorylation–dephosphorylation established as a paradigm the reversible regulation of enzymes by the covalent addition and removal of phosphate groups. Further, the discovery of cAMP-dependent kinases was a catalyst for the study of cAMP as a second messenger in the transduction of signals through certain types of receptors. Structural analyses of serine/threonine kinases, particularly the cyclic AMP-stimulated kinase (protein kinase A), have provided models for the catalytic domains and enzymic mechanisms of all of the kinases (Taylor *et al*., 1995).

Structurally, there are two major classes of kinases, receptor (**Figure 1a**) and nonreceptor (**Figure 1b**), based on elements of their primary sequences which determine their cellular localization. Almost all receptor kinases characterized to date are tyrosine kinases, with the exception of the TGF-β receptor, a Ser/Thr kinase (see chapter on *Signalling by TGF-β*). Receptor kinases are defined by a hydrophobic transmembrane domain, which passes through the plasma membrane, an extracellular ligand-binding domain and a cytoplasmically oriented kinase domain (**Figure 1a**). In contrast, nonreceptor kinases have no transmembrane or extracellular domains, although they may be associated with the cytoplasmic surfaces of cellular membranes by one of two mechanisms. The first is constitutive membrane localization via a lipid modification which anchors the protein to the phospholipid bilayer. The lipid-linked kinases identified to date have been tyrosine kinases, and Src is the prototype of this kinase type, having an N-terminal consensus site for myristoylation (**Figure 2a**). The second membrane localization mechanism involves binding to a nonenzymic membrane receptor to give a binary tyrosine kinase (**Figure 2b**).

Tyrosine phosphorylation is a relatively rare event in normal, unactivated cells, representing $<0.1\%$ of total protein phosphate groups (serine, 90%; threonine, 10%). However, tyrosine phosphorylation is transiently increased during normal cell activation and often substantially increased, sometimes constitutively, in tumour cells. The functions of tyrosine kinases are significantly different from those of the serine/threonine kinases. Conceptually, the primary role of tyrosine kinases is to provide a mechanism for transmitting information from a factor outside a cell to the interior of the cell without requiring that the factor cross the cell's exterior membrane barrier. That role is exemplified by the ability of a growth factor or polypeptide hormone in the extracellular milieu, which cannot pass through the membrane, to activate specific gene transcription in the nucleus of a target cell (see **Figure 3a**). This example is only one of the many cell functions mediated by tyrosine kinases (**Table 2**), and these functions must be performed in a variety of cell and organismal contexts. Thus, this versatility of cellular tyrosine kinases requires that they be complex proteins and carry a variety of ancillary domains in addition to their tyrosine kinase domain to be able to interact with various cellular proteins in the performance of these various functions. Src is a prototype for these multiple interactions because it has two domains in addition to that encoding its tyrosine kinase (**Figure 1b**), Src homology 2 (SH2) and Src homology 3 (SH3), as noted below, which are frequently found in associations among signalling components.

Mechanistic

The transfer of a signal from an extracellular factor through the membrane via a tyrosine kinase is usually performed by one of two types of mechanisms, exemplified in **Figures 2b** and **3a**. These two mechanisms use receptor and nonreceptor tyrosine kinases. In the first mechanism (**Figure 3a**) the receptor tyrosine kinase is solely responsible for the transfer of signal across the membrane. In the second mechanism (**Figure 2b**) a nonreceptor tyrosine kinase is coupled to a transmembrane receptor cytoplasmic domain in a binary receptor. In either case it is the activation of the tyrosine kinase accessible to the cytoplasm which is the key step in transferring the signal across the membrane. These mechanisms also define the two types of tyrosine kinases which have evolved (**Figure 1**) and the minimal domains necessary for each type of tyrosine kinase. The receptor tyrosine kinase must have an extracellular ligand-binding domain(s), a transmembrane domain, a kinase domain and sites for docking the cytoplasmic molecules to which the signal is transferred. The cytoplasmic tyrosine kinases must have the kinase domain, a binding domain for attaching to the receptor and sites for docking the cytoplasmic molecules to which the signal is transferred.

The transmission of signal from the kinase at the membrane to the nucleus for activation of transcription can

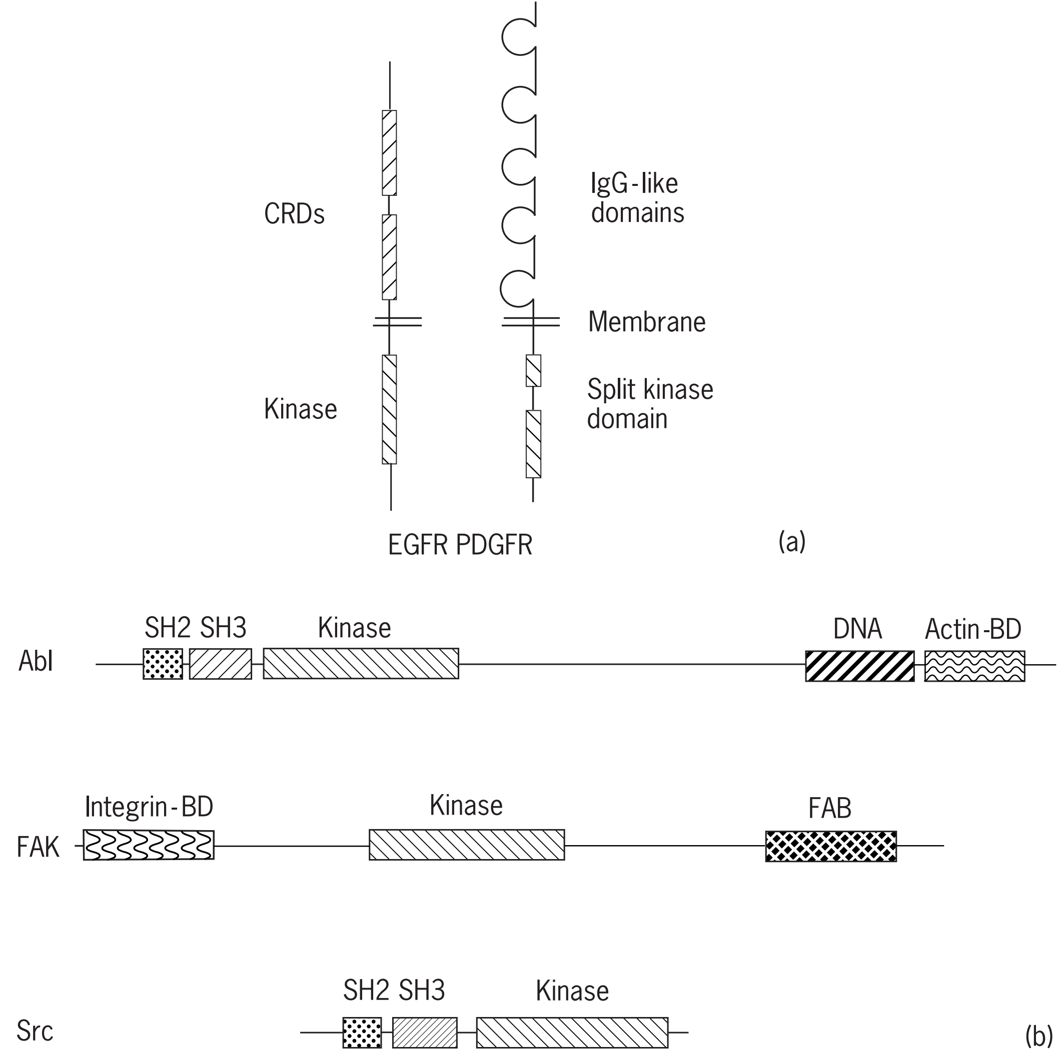


Figure 1 Schematic structures for receptor and nonreceptor protein tyrosine kinases, showing variations in the two classes. (a) Receptor tyrosine kinases ErbB1 (EGF receptor) and PDGF receptor, showing common variations in structure (see also **Table 6**). Ligand-binding sites are contained in the extracellular domains, which are highly variable between different receptor tyrosine kinase families (**Table 6**). CRD, cysteine-rich domain; IgG, immunoglobulin. (b) Nonreceptor (cytoplasmic) tyrosine kinases Abl, FAK and Src. Note the multiple binding domains in the nonreceptor kinases which link them to other functional components in the cell as part of their signalling functions. BD, binding domain; FAB, focal adhesion binding domain. (Adapted from Hunter, 1998.)

also occur by multiple pathways. Two of the most important are illustrated in **Figure 3b**, Ras–MAP kinase and JAK–STAT. The receptor–Ras-MAP kinase pathway consists of four main elements (**Figure 3a** and **b**; **Table 3**): the receptor which receives and transmits the signal through the membrane; an adaptor system (Grb2-SOS) which couples the receptor to a membrane switch (Ras); the Ras switch, which is activated by that coupling mechanism, then activates downstream kinases; and the MAP kinase cascade (Raf–MEK–Erk in **Figure 3**), which ultimately transmits the signal to the nucleus. The central feature of this pathway is the small G protein/GTPase Ras. It is switched on and off in response to its binding of the nucleotide GTP. The binding is controlled by three different types of regulatory proteins: activators which promote the Ras GTPase activity to convert bound GTP to GDP; exchange proteins which replace GDP with GTP; and inhibitors which reduce the GTPase activity. In addition, Ras is not simply an on/off switch. Depending on the cellular context, it can also switch the signal to a different pathway, activating different cell functions. The versatility of this mechanism is illustrated in **Table 4**, which shows some of the multiple downstream components to which Ras can be coupled, again illustrating the multiple cellular functions in which the tyrosine kinases can participate (**Table 2**). Obviously, the complexity of the Ras pathway provides many sites for regulation and for integration of signals by interactions with other pathways.

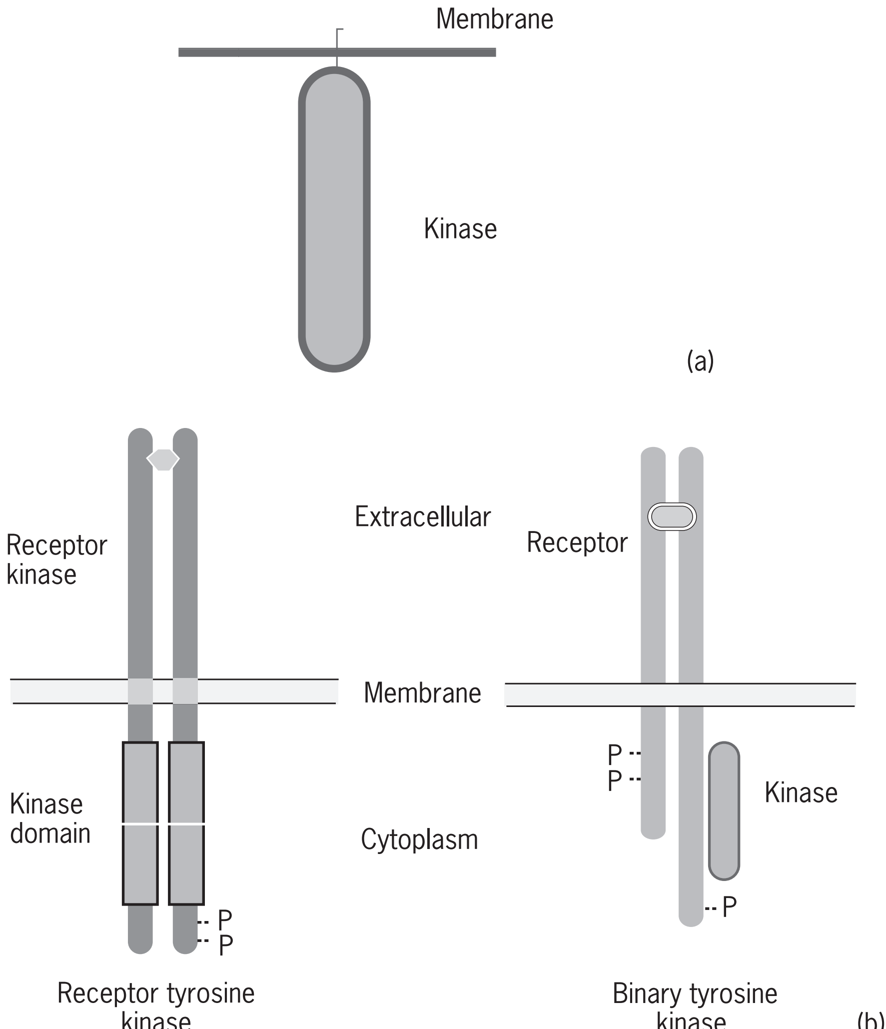


Figure 2 Comparison of mechanisms for associating tyrosine kinases with membrane. (a) Lipid-linked tyrosine kinase; (b) Comparison of structures for transmembrane receptor tyrosine kinases and binary receptor tyrosine kinases. The latter consists of a nonreceptor tyrosine kinase noncovalently linked to a nonenzymic transmembrane receptor.

In the JAK–STAT pathway (**Figure 3b**, also described in the previous chapter, Signalling by Cytokines) the tyrosine kinase JAK phosphorylates a STAT molecule, which translocates to the nucleus. This pathway has the advantage of being much simpler (**Figure 3**), but also has fewer opportunities for regulation.

RECEPTOR TYROSINE KINASES

Mechanism of Signal Transduction

The receptor tyrosine kinase provides a single transmembrane molecule which can transfer information across the membrane from the extracellular milieu to the cytoplasm. This transfer involves two related events: stimulation of the kinase activity and phosphorylation of tyrosine residues accessible to the cytoplasm of the cell. In the most common model for this process, ligand binding to the extracellular domain results in dimerization of the receptor in the membrane (**Figures 2b**, **3a** and **4**). This mechanism has been demonstrated by high-resolution X-ray crystallographic analysis for the growth hormone receptor activated by its dimeric ligand (Wells, 1996). In some cases the receptors may form larger complexes (e.g. tetramers). Regardless, the receptor association leads to potentiation of the kinase activity, probably through structural (conformational) changes in the kinase domain. The second phase of the information transfer involves the phosphorylation of one or more tyrosine residues of the cytoplasmic domain of the receptor (**Figure 4**). Available evidence suggests that this occurs primarily through cross-phosphorylation (*trans*-phosphorylation). In a receptor

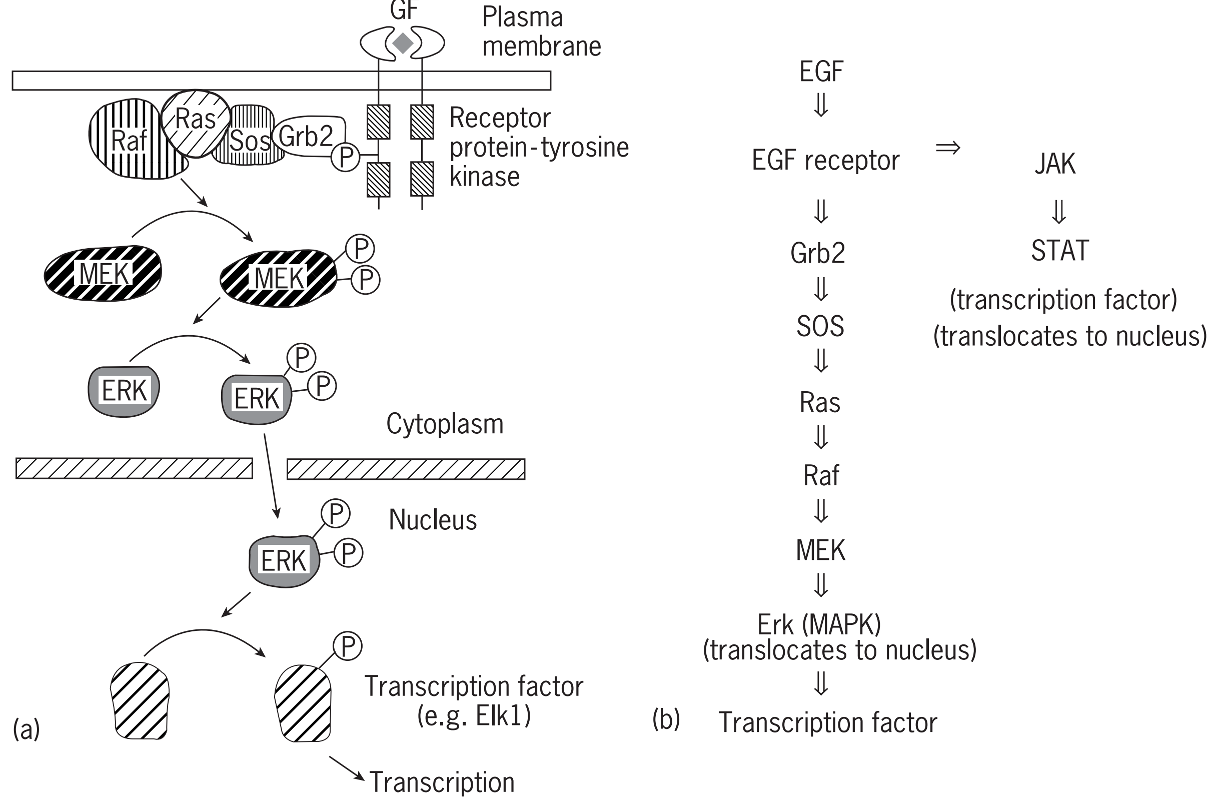


Figure 3 Signal transmission from the extracellular space to the nucleus to activate transcription via the Ras-MAP kinase pathway. (a) Signal translocation; GF, growth factor (b) Comparison of Ras-MAP kinase pathway with JAK–STAT pathway. (Adapted from Hunter, 1998.)

Table 2 Examples of cellular functions regulated by tyrosine kinases

Cell function	Receptor[a]	Downstream effectors[a]
Proliferation	**ErbB**	Ras–Erk (MAPK)
Cell–matrix adhesion	Integrin	**FAK–Src**
Cell–cell adhesion	Cadherin	Catenin-**Src** family
Movement	**PDGFR**	Rac–Rho
Apoptosis control	**IGFR**	PI3K–Akt
Transcription	Cytokine R	**JAK**–STAT
Membrane transport	Channel	**Src** family

[a]Tyrosine kinases are indicated in bold. FAK, focal adhesion kinase; PDGFR, platelet-derived growth factor receptor; IGFR, insulin-like growth factor receptor; PI3K, phosphoinositide 3-kinase; STAT, signal transduction and transcription.

dimer the kinase on one receptor (half-dimer) would phosphorylate tyrosines on the associated receptor (half-dimer) and vice versa. Thus, the ligand-induced association not only activates the enzyme, but also brings the enzyme and its substrate tyrosine residues into proximity. The significance of the tyrosine phosphorylation is that the phosphotyrosine residues created form sites at which cytoplasmic proteins can bind to initiate intracellular signalling (see **Figure 3a**), thus passing a signal across the membrane without the passage of the activating ligand.

The primary function of the phosphorylated tyrosine residues on the activated receptors is to recruit signalling components from the cytoplasm or cytoplasmic surface of the plasma membrane to initiate signalling pathways (**Figures 3a** and **4**) (Panayotou and Waterfield, 1993). The particular pathway initiated is determined by two complementary factors: the site of the tyrosine on the receptor polypeptide chain and the specificity of the phosphotyrosine-binding domains on the cytoplasmic proteins. The sites of tyrosine phosphorylation on a particular receptor are determined by the ligand–receptor and receptor–receptor interactions, which regulate the accessibility of any particular tyrosine to the kinase catalytic site. Two different types of domains have been found in signalling components which bind to phosphotyrosines to initiate their recruitment to the receptor: SH2 (see **Figures 1b** and **5**) and PTB (**Figure 5**). The SH2 domain is a compact globular unit of about 100 amino acids, which was originally discovered in Src and is found in a large number of signalling components (**Table 5**). Its binding specificity is determined primarily by the phosphotyrosine and 1–5 amino acid residues following it (C-terminal direction) in the receptor amino acid sequence. The PTB domain is also globular and has about 150 amino acids; its specificity is determined by the phosphotyrosine and 1–8 amino acid residues preceding it (N-terminal direction) in the receptor

Table 3 Components of pathways regulating transcription

Receptor–Ras–MAPK pathway		**Receptor–JAK–STAT pathway**	
Component	**Function**	**Component**	**Function**
EGFR	Receptor	EGFR	Receptor
Grb2	Adaptor	JAK	Nonreceptor tyrosine kinase
SOS	GTP exchange protein	STAT	Transcription factor
Ras	G protein switch		
Raf	Ser/Thr kinase		
MEK	Dual-function Tyr/Thr kinase		
Erk (MAPK)	Ser/Thr kinase		
Transcription factor	Transcription factor		

Table 4 Examples of signalling elements and cellular functions coupled to Ras activation

Signalling component	**Potential cell function**[a]
Rac	Shape, movement
RalGDS	Shape, movement
P120 GAP	Shape, movement
Raf	Transcription, proliferation
PI3K	Proliferation, apoptosis resistance
MEKK	Transcription, stress reaction

[a] Listed functions are illustrative, not inclusive.

sequence. Other types of phosphotyrosine-binding domains have been suggested, but have not been well characterized.

These observations indicate that most components initiating signalling pathways from receptor tyrosine kinases should contain either SH2 or PTB domains. Indeed, identification of proteins with these domains has provided significant insights into signalling. As shown in **Table 5** and **Figure 5**, these proteins include three classes of signalling pathway initiators: enzymes, docking proteins and adaptors. Each of these types of molecules facilitates signalling by a different mechanism. Enzymes actively participate in downstream signalling events. Docking proteins provide additional or surrogate tyrosine phosphorylation sites for further diversification of receptor sites to initiate signalling. Adaptors are involved in assembling complexes of signalling components for initiating and regulating downstream signalling pathways. However, in all three types of mechanisms the key aspect of the pathway is the recruitment and relocalization of the signalling component to the site of the receptor (exemplified by Grb2 in **Figure 3a**) (Panayotou and Waterfield, 1993; Carraway and Carraway, 1995).

Four different rationales can be envisioned for the recruitment of enzymes to initiate signalling pathways. First, recruitment can permit the receptor tyrosine kinase to phosphorylate the recruited molecule, thus changing its activity or binding function. An example of this is phospholipase Cγ, whose recruitment and phosphorylation activates a phospholipid-cleaving enzyme to stimulate a pathway leading to calcium influx into the cytoplasm (**Figure 6**). The calcium plus lipid hydrolysis product diacylglycerol combine to stimulate a potent calcium-dependent serine/threonine protein kinase, protein kinase C, which then has multiple cellular effects. Second, recruitment brings the recruited molecule to the membrane, where it may act upon membrane components as substrates. This mechanism applies to phospholipase Cγ, which cleaves phospholipid molecules found in membranes, and to phosphoinositide 3-kinase, one of the signalling enzymes implicated in oncogenesis, which phosphorylates specific phospholipid molecules in membranes (**Figure 6**). These enzymatic modifications of membrane components may change the organisation of the membranes or produce new signalling molecules, such as the lipid phosphate PIP_2, to initiate and perpetuate downstream signalling pathways. It is important to note that the same signalling molecule may initiate different signalling pathways, depending on the cellular context. Both PI3 kinase and phospholipase Cγ can be involved in multiple cellular functions, as previously indicated for Ras (**Table 4**). Third, binding of a molecule to the receptor phosphotyrosine residues may induce its activation by a conformational change (allosteric effect), a mechanism which contributes to the signalling effects of phosphoinositide 3-kinase (**Tables 2** and **4**; **Figure 6**). Fourth, some of the recruited enzymes are protein tyrosine kinases or protein tyrosine phosphatases. Since the receptor tyrosine kinases are often involved in large, multimeric signalling complexes (Carraway and Carraway, 1995), such as focal adhesion sites (see below), this recruitment brings the additional kinases and phosphatases into proximity of their substrates in these complexes and facilitates regulation of the signalling components involved. An important aspect of such complexes is that they are often located at sites influencing cell behaviour, such as sites for membrane-microfilament interactions involved in cell shape and movement (Carraway *et al*., 1997).

Recruitment of proteins can also amplify signalling potential through the docking protein mechanism. The

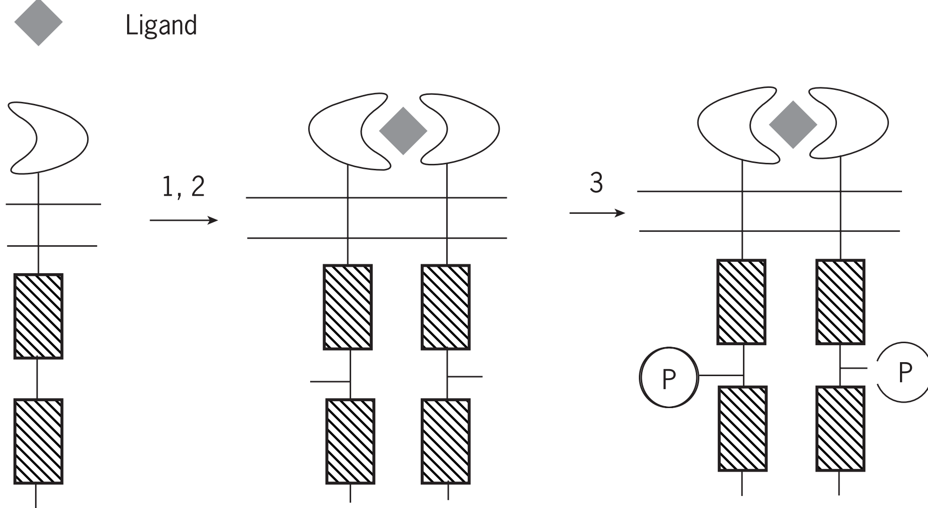


Figure 4 Mechanism for transmission of signal across the membrane by binding of ligand, dimerization of receptor and transphosphorylation of cytoplasmic domain to create sites for recruitment of cytoplasmic signalling components to initiate signalling pathways.

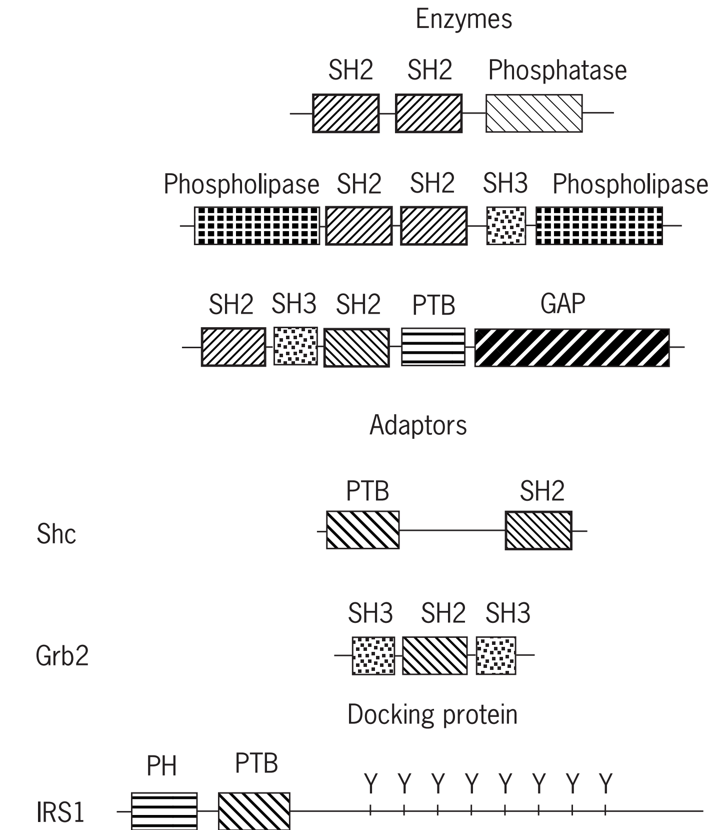


Figure 5 Schematic structures for enzymes, adaptors and docking proteins involved in initiating cellular signalling pathways by binding to tyrosine-phosphorylated receptors or binary receptor complexes. GAP, GTPase activating protein. (Adapted from Hunter, 1998.)

best characterized docking proteins are the insulin receptor substrates (IRSs, **Figure 5**), which regulate a variety of cellular functions, including membrane transport, gene expression, protein synthesis and lipid and carbohydrate metabolism. IRSs contain a PTB domain (**Figure 5**) which allows them to bind the activated insulin receptor. Formation of this insulin–IRS complex increases the number of tyrosines available for phosphorylation and increases the insulin receptor signalling potential. Moreover, IRSs act via multiple receptors, including the insulin-like growth factor (IGF) receptor and various cytokine receptors. Phosphorylated tyrosines on IRSs can interact with a variety of SH2-containing components to activate multiple downstream signalling pathways. Among these are the adaptors Grb2, Crk and Nck (**Figure 5**), the phospholipid kinase PI3 kinase, the tyrosine kinase Fyn and the phosphotyrosine phosphatase SHP2 (**Figure 5**).

Adaptor molecules participate in the recruitment and organisation of signalling components from the cytoplasm into membrane complexes. The best studied example is the initiation of the Ras–MAP kinase mitogenic pathway by the adaptor Grb2 (**Figure 3a**). Grb2 contains both SH2 and SH3 domains, allowing it to link a tyrosine-phosphorylated receptor to a cytoplasmic SH3-binding protein containing a proline-rich motif. In the cytoplasm of unactivated cells Grb2 is present as a complex through its SH3 domain with a second protein SOS, a GTP exchange protein which activates Ras. Activation of ErbB1 (EGF receptor) by its ligand EGF creates a phosphotyrosine binding site for the Grb2 SH2 domain on the receptor, recruiting the Grb2–SOS complex from the cytoplasm to the plasma membrane. When associated with the membrane, SOS can bind to and activate the proto-oncogene product Ras. Activated Ras stimulates a series of serine/threonine protein kinases, culminating in the activation of a specific MAP kinase (Erk), which can migrate to the nucleus as a signal for inducing gene transcription (**Figure 3a** and **b**). The key step in initiation is the linkage of the Ras switch protein to the ErbB receptor via the adaptor Grb2. As noted previously, Ras also has the potential to activate other signalling pathways (**Table 4**).

Table 5 Phosphotyrosine-binding proteins which could activate or diversify signalling pathways by binding phosphorylated receptors

Protein	Class/function	Domains
PLC-γ	Phospholipase	SH2 (2), SH3
GAP	GTPase activator	SH2, SH3
Src family	PTK	SH2, SH3
ZAP/SYK	PTK	SH2 (2)
Shp1/2	Tyrosine phosphatase	SH2 (2)
PI3K	Lipid kinase	SH2 (2), SH3
Ship	Lipid phosphatase	SH2
Vav	GTP exhange factor	SH2, SH3
Shc	Adaptor	SH2, PTB
Nck	Adaptor	SH2, SH3 (3)
Grb2, Crk	Adaptor	SH2, SH3 (2)
IRS1/2	Docking protein	PTB
Talin	Focal adhesion protein	SH2
STAT	Transcription factor	SH2, SH3

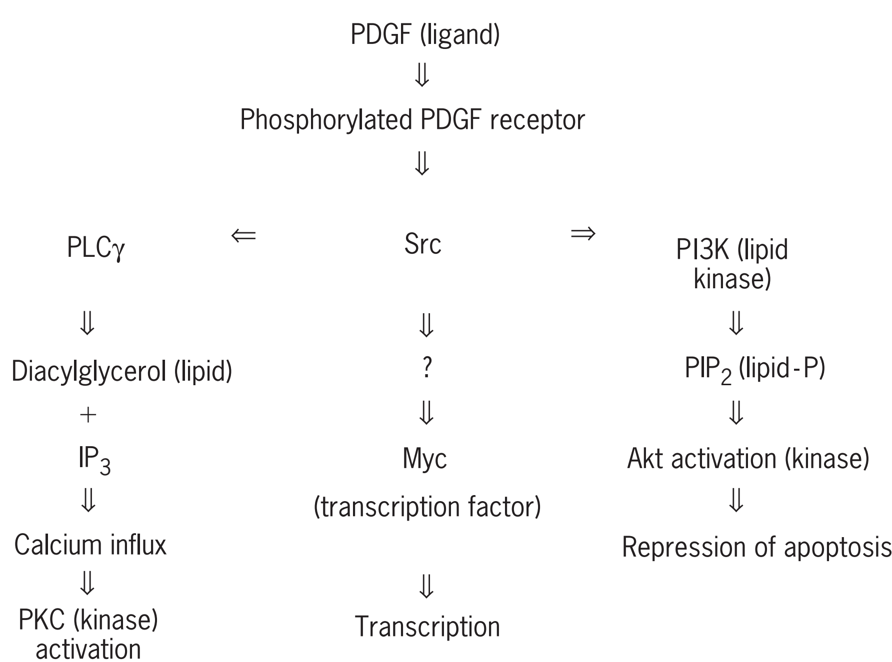


Figure 6 Examples of multiple downstream pathways coupled to receptor phosphorylation of PDGF receptor.

Versatility of Receptor Tyrosine Kinase Signalling

One of the hallmarks of the receptor tyrosine kinase mechanism is its versatility, mediating signal transduction via multiple pathways in different organisms, tissues and cells. This versatility allows these enzymes to regulate many of the vast number of cellular processes and interactions required in the development and function of complex multicellular organisms. The bases for the versatility are several-fold, the first being the multiplicity of receptors. There are a minimum of 20 families of mammalian receptor tyrosine kinases, each with different ligand-binding domains and most with multiple members (**Table 6**). Thus, part of the versatility of the receptor tyrosine kinase mechanism arises from the diversity of ligand–receptor interactions available in different tissues of different organisms at different times. Second, many of these individual receptor kinases can bind more than one ligand, as exemplified by the ligand binding patterns of the class I (ErbB) family (**Table 7**), leading to signal diversification. Third, the distribution of receptors in the organism is tissue and cell dependent. Some receptors are widely distributed; others are specifically localized to a small number of sites. Tumours may have aberrant distributions of receptors. Fourth, the distribution of many receptor tyrosine kinases and their ligands is developmentally regulated, being found only during specific stages of the life history of the organism. Fifth, ligands are produced at different sites from the target cells to exert different levels of control, as shown by examples in **Table 8**. Endocrine ligands are produced in different organs and transmitted through the circulation to the site of action. Paracrine ligands are produced by a different cell type or tissue from the target cells, but near the target. The juxtacrine mechanism is a special case of the paracrine ligand in which the ligand and receptor are expressed on the plasma membranes of adjacent cells. Autocrine responses arise from ligands produced by the same cell (cell type) as the receptor-bearing target cell. The intracrine and intramembrane mechanisms are special cases of the autocrine response, in which ligand and receptor are produced in the same cell.

The receptor structure itself contributes to the diversification of signals. Each receptor contains multiple

Table 6 Multiplicity of receptor tyrosine kinases

Receptor family	Extracellular domain[a]	Number in family
ErbB	Cysteine rich	4
PDGF	Ig	5
FLT	Ig	3
Insulin	Cysteine rich, FN III	2
TRK	Ig, leucine rich	3
FGF	Ig, acid box	4
Eph	FN III, CR, Ig	14
Axl	Ig, FN III	3
TIE	Ig, EGF, FN III	2
RET	Cadherin	1
MET	Sema	2
ROR	Ig, CR, kringle	2
MCK10	Factor VIII	2
MuSK	Ig	1
CCK4	Ig	1
ROS	FNIII	1

[a] Ig, immunoglobulin-like; FNIII, fibronectin III-like; CR, cysteine rich.

Table 7 Ligand–receptor couples for the ErbB family

Receptor	Soluble ligands
ErbB1	EGF, TGFα, amphiregulin, HB-EGF, betacellulin, epiregulin
ErbB2	None known
ErbB3	Neuregulins
ErbB4	Neuregulins, HB-EGF, betacellulin, epiregulin

Table 8 Examples of ligand sources and binding mechanisms

Mechanism	Ligand	Receptor	Definition	Example
Endocrine	Insulin	Insulin R	Ligand reaches receptor through circulation	Pancreatic cell insulin; fat cell receptor
Paracrine	Neuregulin	ErbB3	Ligand produced by adjacent cell or tissue	Mesenchymal NRG; epithelial ErbB3
Autocrine	TGF-α	ErbB1	Ligand secreted by same cell bearing receptor	TGF-α secretion by tumour cells with ErbB1
Intracrine	v-sis	PDGF R	Ligand and receptor interact inside cell	v-sis binding to PDGF R in cells
Juxtacrine	Bride of Sevenless (BOSS)	Sevenless	Ligand and receptor in plasma membranes of adjacent cells	BOSS-Sevenless interaction in control of *Drosophila* eye development
Intramembrane	MUC4	ErbB2	Ligand and receptor in same membrane	MUC4/ErbB2 complexes in tumour cells and epithelia

Table 9 Phosphorylated tyrosine-binding sites for signalling components in PDGF

Sequence position	Binding protein	Class[a]
579	Src, STAT	PTK, TF
581	STAT	TF
716	Grb2	Adaptor
740	P13K	Lipid kinase
751	Nck, PI3K	Adaptor, LK
771	GAP	GTPase activator
775	STAT	TF
778	Grb7	Adaptor
1009	SHP-2	Tyrosine phosphatase
1021	PLC-γ	Phospholipase

[a] PTK, protein tyrosine kinase; TF, transcription factor; LK, lipid kinase.

tyrosine residues in its cytoplasmic domains which can be phosphorylated to form different binding sites for cytoplasmic signalling proteins. For example, the PDGF receptor contains at least 10 tyrosine residues which have been shown to be phosphorylated in response to ligand activation (**Table 9**). Each different phosphorylated tyrosine potentially represents a different signalling pathway which could be initiated (Claesson-Welsh, 1994), some of which are shown in **Figure 6**, though overlaps and redundancies inevitably reduce that number. By comparing **Table 9** with **Tables 2** and **4**, it is clear that the PDGF receptor has the potential for participating in a large number of cellular functions. Diversity and specificity of signalling are achieved in part by phosphorylation of different combinations of the tyrosine residues in response to different extracellular ligands, an example of a combinatorial mechanism for regulation. A second diversification and specificity mechanism arises from the type of receptor association during receptor activation. Formation of a heterodimer by two different molecules of a receptor family obviously can yield more different phosphorylated tyrosine residues than formation of a homodimer by two identical molecules from the same family. Thus, the number of potential signals is increased. In the case of the ErbB family of receptors (**Table 7**), the four receptors can form 10 different combinations of homodimers and heterodimers in response to different ligands, all of which have been observed and have potentially different signalling capabilities (Riese and Stern, 1998). For this family of receptors, heterodimerization appears to be the preferred mechanism of activation in many physiological contexts (Riese and Stern, 1998).

Regulation of Receptor Tyrosine Kinase Signalling

The complex signalling pathway from the receptor to the nucleus (**Figure 3**, ErbB1 through MAP kinase) involves at least seven different components and provides multiple levels of control of the signal. In any phosphorylation-dependent system, one obvious control mechanism is the removal of the phosphate(s). In a complex chain involving multiple phosphorylations, reversal of any phosphorylation event can potentially break the chain and block the signal. Cells contain a large variety of phosphatases to hydrolyse phosphotyrosine and phosphoserine/threonine residues (Streuli,1996; Cohen, 1997). Contrary to early expectations that phosphatases would provide nonspecific 'off' switches for kinase signalling, both serine/threonine and tyrosine phosphatases exhibit considerable specificity. They even participate directly in the activation of protein tyrosine kinases such as Src family members, as described below. As with the tyrosine kinases, both membrane and nonmembrane forms of tyrosine phosphatases have been observed. Both also contain multiple domains which regulate their associations and locations in cells. Although membrane tyrosine phosphatases contain multiple types of extracellular domains (Streuli, 1996), it is not clear whether they can act as true receptors, since no ligand activation mechanisms for their enzyme activities are known.

Ligand-activated receptor tyrosine kinases can also be regulated by controlling the availability of the ligand to the receptor. Ligands are often synthesized from high molecular weight membrane precursors, which are cleaved proteolytically to release the ligand. Thus, ligand release and availability are determined in part by the activity of the protease(s) involved in the cleavage. One of the factors regulating these proteases is calcium-dependent protein kinase C. Ligand degradation may also contribute to determining ligand availability. Ligand binding to some receptors triggers endocytosis of the ligand–receptor complex into the cell (Sorkin and Waters, 1993), which can lead to three possible fates: (1) the complex may dissociate and the ligand transfer to lysosomes for degradation, while the receptor recycles to the cell surface; (2) the complex may transfer to the lysosome for degradation of both components; (3) both components may recycle to the cell surface. Either mechanism for degrading the ligand reduces its availability. Finally, ligand availability may be determined by interactions with extracellular components, both positively and negatively (Schlessinger *et al*., 1995). Many ligands contain positively charged amino acid sequences which interact with glycosaminoglycans (GAGs), such as heparin, at the cell surface or in the extracellular matrix. This interaction may affect ligand availability in two ways. In a positive sense cell-surface GAGs or other cell-surface components may recruit ligands to the cell surface and vicinity of the receptors. Cell-surface components, such as GAGs, may also act as coreceptors to form multimeric complexes with ligands and receptors which facilitate the activation of the receptor kinases. In a negative sense GAGs may sequester ligands away from their receptors. By combining the two mechanisms, ligands in a tissue may be held in a GAG 'reservoir' until they are released by an acute event that

frees the ligand and consequently initiates receptor activation. Soluble ligand-binding proteins are produced by some tissues and sequester extracellular ligands. For example, a whole family of binding proteins regulate the availability of insulin-like growth factors.

Receptor activation can also be determined by receptor availability, often dictated by receptor turnover, as described above for ligand–receptor complex endocytosis and degradation. These mechanisms appear to be rather receptor specific. For example, ligand binding induces turnover of ErbB1 (EGF receptor) more readily than other ErbB family members. A second aspect of receptor availability is localization. This may occur at either the cellular or subcellular level. At the cellular level only certain cell types or cells in specific locations may contain a given receptor in a tissue. At the subcellular level the receptor location may be restricted to a specific region of the cell surface, e.g. to cell–cell contacts. Localization is particularly important to receptors which act by juxtacrine mechanisms (**Table 8**), since they require appropriate cell–cell contacts for their activation. One example is the Eph family of receptors (**Table 6**), which guide movements of cells and neuronal growth cones in establishing neuronal connections. The mechanism for guidance control involves a repulsive reaction to juxtacrine association of an Eph receptor with its complementary Ephrin ligand on an adjacent cell. This behaviour not only emphasizes the importance of location, but also suggests the possibility of reciprocal signalling, in which both the receptor and ligand initiate pathways and responses to the contact. Once again, this unusual mechanism underlines the diversity which has evolved in receptor tyrosine kinase functions to regulate cellular behaviours in multicellular organisms.

Although ligand-dependent activation is the most common means for stimulation of receptor tyrosine kinase pathways, other mechanisms have been proposed. Overexpression of ErbB2 is sufficient for activation of receptor phosphorylation and cell transformation in some types of cells and has been implicated in neoplasia. One explanation for this effect is that the receptors associate when their concentration reaches a certain level in the membrane, triggering kinase activation and cross-phosphorylation, although the participation of other factors cannot be ruled out. One possible example of indirect activation of ErbB2 through multimerization involves the cell surface molecule CD44, a receptor for the extracellular GAG hyaluronic acid. CD44 has been observed to associate with ErbB2. Hyaluronic acid, an extracellular matrix component, can bind and aggregate CD44 in the cell membrane, thus also aggregating ErbB2 and inducing its activation. Other indirect mechanisms for initiating receptor tyrosine kinase signalling pathways are less easily rationalized. These include stimulation of G protein-coupled receptors, activation by cytokines, cellular stress responses, cell adhesion and membrane depolarization. One possible explanation for these other ligand-independent signalling mechanisms is that the receptors are not acting as enzymes in these instances, but merely serving as docking proteins such as the IRSs (see above) which are phosphorylated by cytoplasmic tyrosine kinases, such as Src, activated by these other signalling mechanisms. An alternative possibility is that these additional mechanisms trigger proteolytic activities which release ligands from their precursors or sequestration sites to activate receptor tyrosine kinases.

Not surprisingly, ligand mimics have evolved which can modulate receptor signalling. In the fruit fly the secreted protein Argos contains an EGF-like domain and blocks signalling through the ErbB family tyrosine kinase receptor, probably acting as a competitor for the activating ligands. In contrast, mammalian epithelia contain a mucin Muc4 with an EGF-like domain which binds to ErbB2 and potentiates ligand-activated phosphorylation through heterodimer ErbB2/ErbB3. The sequence of each of these mimics differs slightly from the sequences of activating ligands of the receptors. Other receptor-binding proteins associate with the receptor extracellular domains to influence ligand binding and signalling or with the cytoplasmic domain to modulate downstream signalling, although the exact mechanisms have not been well studied. Receptor ‘desensitization’ is a common phenomenon for many types of receptors, including tyrosine kinases. A frequent mechanism involves phosphorylation of the receptor at specific sites to repress its activity. For example, ErbB1 (EGF receptor) can be desensitized by phosphorylation of serine residues in its cytoplasmic juxtamembrane domain by protein kinase C.

BINARY RECEPTORS

Activation of the JAK/STAT Pathway

Binary receptors consist of a nonmembrane tyrosine kinase subunit noncovalently associated at the cytoplasmic surface of the plasma membrane with a transmembrane receptor (**Figure 2b**). One advantage of the dual subunit system is the ability to ‘mix and match’ different gene products for the two subunits to create a broader array of different signalling units from the same amount of genetic information, increasing the diversity of signalling potential. Moreover, binary receptors often contain multiple receptor (nonkinase) subunits to provide further specificity and diversity. However, the mechanism of signal transduction across the membrane is remarkably similar to that of the single subunit receptor tyrosine kinases. Binding of ligand to the receptor induces dimerization or multimerization of the receptor with concomitant activation of the kinase. Cross-phosphorylation can then occur on both the receptor and kinase subunits, creating binding sites for cytoplasmic proteins with SH2 or PTB domains. In the case of the binary cytokine receptors (reviewed in detail in the previous chapter, *Signalling by Cytokines*),

the associated tyrosine kinases are the Janus kinases (JAKs). The JAKs participate with signal transducing and transcription (STAT) factors in a direct mechanism for regulating transcription (**Figure 3b**). STATs are recruited to phosphorylated receptors via their SH2 domains. Tyrosine phosphorylation of the STATs by receptor-associated JAKs creates SH2 binding sites which can interact reciprocally with the STAT SH2 domains to induce STAT dimerization. The STAT dimers translocate to the nucleus, where they are able to regulate gene transcription (see the previous chapter, *Signalling by TGF-β* for details). The JAK/STAT pathway is thus more direct than the Ras–MAP kinase pathway, which also regulates cell transcriptional activity. Moreover, the JAK/STAT pathway can also be initiated by some receptor tyrosine kinases, such as ErbB1. Conversely, adaptors such as Grb2 are able to bind to some cytokine receptor phosphotyrosines to couple them to Ras and the MAP kinase signalling pathway, and IRS docking proteins can link them to multiple signalling elements. Thus, there appears to be substantial redundancy in signalling pathways. A major concern in signalling research is the molecular mechanisms for coordinating different signalling pathways, whether they derive from the same receptor (pleiotropy) or from different receptors (cross-talk).

Lymphocyte T cell Complex

Even more complex binary receptor tyrosine kinase complexes participate in lymphocyte regulation (Peterson and Koretzky, 1998), as exemplified by the T cell receptor complex. T lymphocytes are the immune cells responsible for defence against invading organisms, such as bacteria and viruses. The T cell has evolved a very sophisticated mechanism for recognizing invader cells, involving recognition of specific proteins unique to the invaders. However, rather than recognizing the intact protein, specialized cells of the immune system degrade the foreign proteins to peptides, which can then be ‘presented’ for recognition at their cell surfaces. This antigen presentation is performed by a cell-surface protein complex called the major histocompatibility complex (MHC). It is the recognition of the foreign peptide on the MHC molecule by a receptor on the T cell (T cell receptor) that triggers T cell activation. The recognition and activation processes are finely tuned to provide a graded response to the foreign material.

The interaction of the peptide on the MHC of the ‘antigen-presenting cell’ with the T cell receptor on the T cell is a juxtacrine response (**Table 8**). Moreover, the T cell receptor itself is a highly complex moiety, a multicomponent binary receptor. The full receptor contains at least nine polypeptides which can interact with four tyrosine kinases and one or more tyrosine phosphatases. Presumably, this complexity facilitates the regulation of the receptor signalling to provide a graded response instead of a switch (on/off) type of response. The signal development can be best presented as a timed series of events. The T cell receptor initially contains two types of moieties. The peptide recognition moiety is a heterodimer of two transmembrane proteins (α/β or γ/δ, depending on the T cell type) which bind the peptide on the juxtaposed antigen-presenting cell (**Figure 7**). This receptor

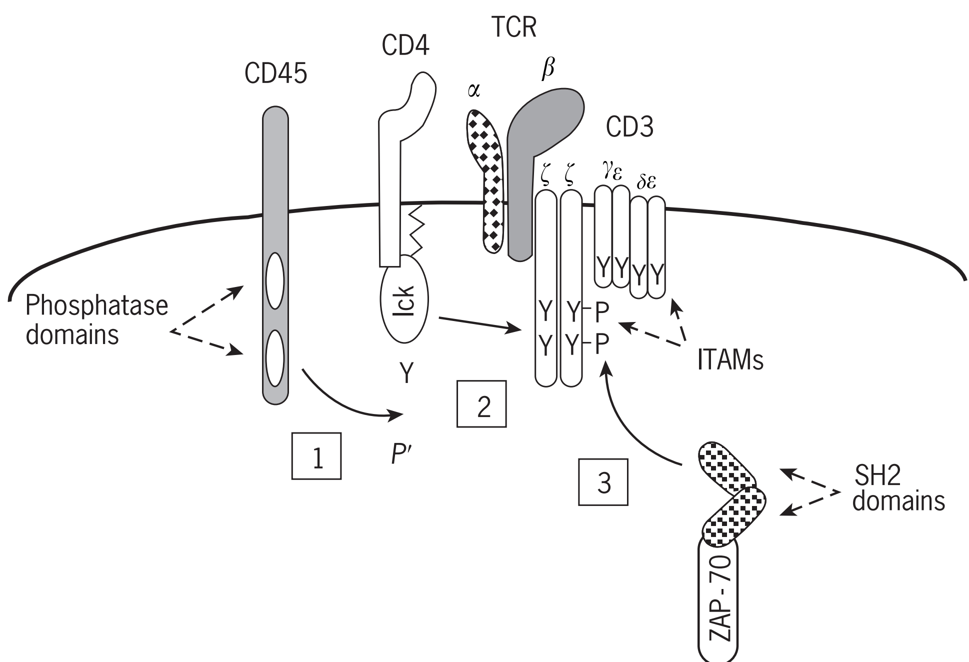


Figure 7 Schematic structure of T cell receptor and some key events in T cell receptor signalling. ITAM, immunoreceptor tyrosine-based activation motif. (From Peterson and Koretsky, 1998, *Clinical and Experimental Rheumatology*, **17**, 107–114.)

heterodimer is associated with two additional transmembrane complexes, a heterodimer and a heterotetramer (CD3). These complexes appear to serve a dual function. First, they act as the tail of a binary receptor to bind specific tyrosine kinases of the Src family for signal initiation. Second, they contain specific sequences called immunoreceptor tyrosine-based activation motifs (ITAMs) which can be phosphorylated to provide binding sites for proteins with SH2 domains, thus serving the function of docking proteins. The kinetic model for T cell activation proceeds when the T cell receptor is engaged by the peptide on the MHC complex of the adjacent cell. This engagement leads to dimerization of the T cell receptor and activation of an associated Src family kinase Fyn. The activated kinase can then phosphorylate the ITAM motifs of the docking proteins to provide recruitment sites for additional molecules. One of the recruited molecules is CD4 (or CD8, depending on the cell type), a binary coreceptor containing a transmembrane component and a second member of the Src family Ick. The CD4 association stabilizes the T cell receptor dimer and provides additional sites for recruitment of SH2-containing components. Among the components recruited to the phosphotyrosine sites created by the Src family members are two additional tyrosine kinases of another family, Zap and Syk.

The result of this complex series of manoeuvers is to recruit a collection of initiators of signalling pathways (**Figure 5** and **Table 5**) to activate the multiple functions of the T cell. The initiators include PLCγ, PI3 kinase, a Ras activator called Vav and the adaptor proteins Shc and Grb2. The activities of the tyrosine kinases are tightly regulated by ancillary molecules. Since Src family kinases can be both positively and negatively regulated by tyrosine phosphorylation (see below), these regulators include the membrane tyrosine phosphatases CD45 and SHP-1 and the nonreceptor tyrosine kinase Csk. CD45 is an activator, while SHP-1 and Csk are both inhibitors of Src family members. The activation mechanisms for Src family members are described in more detail below. Thus, the balance of these activating and inhibiting activities provides mechanisms for generating graded responses. Many of the specifics of the temporal associations and the regulation remain to be discovered, but the T cell receptor provides a clear example of how the binary receptor tyrosine kinase model has evolved to provide a very specific, highly regulated function for higher organisms.

Focal Adhesions

The focal adhesion complex provides another example of the application of the binary receptor principle to a primary cell function. The key tyrosine kinase in this case is the focal adhesion kinase (FAK, see **Figure 1b**). Focal adhesions are sites of cell attachment to the extracellular matrix (see chapter on *the Extracellular Matrix: The Networking Solution*) and of microfilaments to the plasma membrane. Thus, they play a critical role in the determination of cell shape and cell movement (Carraway *et al*., 1997). Furthermore, cell attachment through focal adhesions has been shown to be necessary for normal progression through the cell cycle for cell proliferation. One of the hallmarks of neoplastic transformation is the ability of the transformed cells to escape this adhesion requirement. Focal adhesions are extremely complex structures, containing dozens of components (Carraway *et al*., 1997). Precise assembly and disassembly of these structures is necessary for their roles in cell movements and growth regulation. The objective here is to show how the binary model and tyrosine phosphorylation can contribute to the dynamics of these structures.

The receptors involved in the formation of focal adhesions are integrins, a family containing two types of subunits (α and β), which combine to form a large number of different heterodimeric receptors with different specificities for extracellular matrix components. Ligand binding and aggregation of integrin heterodimers induces their oligomerization in the membrane and recruitment of the nonreceptor tyrosine kinase FAK (**Figure 1B**) from the cytoplasm. This event also stimulates tyrosine phosphorylation of FAK, providing SH2 binding sites for the recruitment of Src. Additional phosphorylation of FAK creates sites for binding docking proteins paxillin and Cas. FAK appears to serve as both kinase and docking protein in these complexes, similar to the cytoplasmic domain of a receptor tyrosine kinase. The combined binding sites in these proteins can initiate further recruitment through SH2 domains and other mechanisms of both structural components of the focal adhesion, such as the actin-binding proteins tensin, talin, vinculin and α-actinin, and signalling pathway components. Included among the latter are PLCγ, PI3 kinase, members of the MAP kinase cascade, members of the Src family and the small G proteins Rac and Rho. These small G proteins are relatives of Ras and act as signalling switches for the processes involved in the organisation of microfilaments in cell movements. One of the enzymes recruited is Csk, which can then serve as a regulator of the Src family members. Thus, the tyrosine kinase FAK, acting through a binary receptor mechanism, is able to initiate the assembly of an extremely complex multimeric structure which is critical to multiple cell functions required for cell movements, such as those involved in tumour invasion.

Src Regulation

Since Src family members play such important roles in many cellular processes involving binary receptors, one of the critical questions in understanding these signalling processes is how the Src activity is regulated. This regulation can most easily be considered as an autorepression mechanism (Hubbard *et al*., 1998) which limits access to the kinase site for catalysis by folding the Src molecule

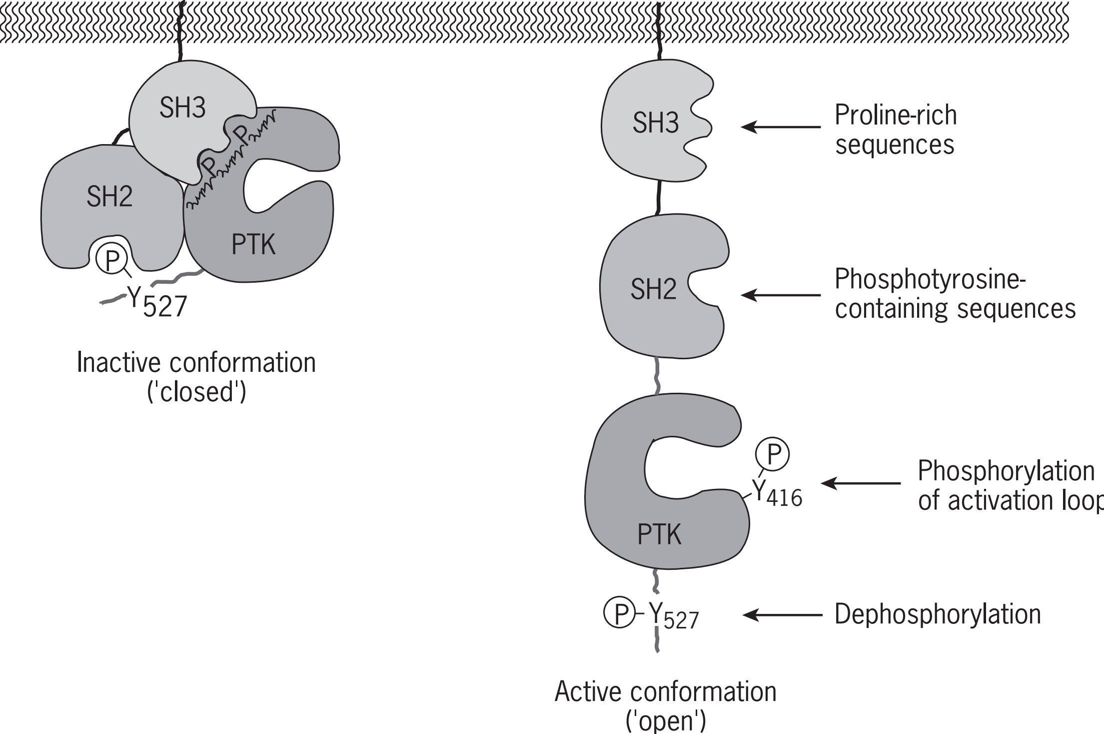


Figure 8 Autorepression/autoactivation model for Src family kinases. (From Hubbard *et al.*, 1998, *Journal of Biological Chemistry*, **273**, 11987-11990.)

into a more compact structure (**Figure 8**). Two interactions are particularly important: (1) the binding of the Src SH3 domain to the peptide linking the SH2 and kinase domains and (2) the binding of the SH2 domain to a phosphorylated tyrosine in the C-terminal tail. Both of these intramolecular interactions must be weak, because they can be displaced by external agents with activation of the enzyme. Thus, Src can be activated by binding appropriate SH3-containing proteins or proteins with SH3-binding motifs. Src can also be activated by protein tyrosine phosphatases, such as CD45 (see **Figure 7**) which cleave the C-terminal phosphotyrosine. Conversely, active Src can be inactivated by phosphorylation of the C-terminal tyrosine by Csk. Activation of Src is also promoted by phosphorylation of a tyrosine near the kinase site to prevent folding of the protein into the compact, inactive form.

NONRECEPTOR TYROSINE KINASE MECHANISMS

Abl

The large number of signalling pathways and cellular functions regulated by the receptor tyrosine kinase and binary receptor mechanisms raises the question whether tyrosine kinases act by any other mode. Unfortunately, it is often difficult to distinguish nonreceptor from receptor mechanisms at the cell level. One highly studied nonreceptor kinase is Abl. Abl is a multidomain molecule whose N-terminal half is similar to the Src family (kinase, SH3 and SH2 domains), but whose C-terminal half has DNA-binding, microfilament-binding and nuclear localization motifs (**Figure 1b**) (Zou and Calame, 1999). Abl is most familiar as the oncogene associated with chronic myelogenous leukaemia. In this cancer Abl is activated as a consequence of a chromosome translocation encoding a fusion protein of Abl and Bcr (breakpoint cluster region). This chimaeric protein appears to be primarily localized to microfilaments in cells. In contrast, normal unactivated Abl is primarily a nuclear protein, although it is able to shuttle between the nucleus and cytoplasmic microfilaments in response to cell adhesion. In the nucleus Abl has been implicated in responses to DNA damage and in cell cycle progression by virtue of its ability to bind p53 and pRb, respectively, both of which are transcription regulators and tumour suppressors. The specific mechanisms for these responses are yet unclear. Abl can also bind RNA polymerase II, the Crk adaptor (**Table 5**) and the protein tyrosine phosphatase SHP1 (**Table 5**). Abl has been implicated in numerous signalling pathways, including those involving PI3K, PKC, Ras, JAK/STAT and Rac. Thus, it has the capacity for participating in multiple cellular functions, arising from both the binding of SH2-containing proteins to its phosphotyrosines and other proteins to other domains or motifs. Other tyrosine kinases which function by nonreceptor mechanisms probably act similarly, although probably not as globally.

Membrane transport and tyrosine kinases

A more direct action of nonreceptor tyrosine kinases is suggested by studies of nonreceptor tyrosine kinase effects on channel activities (Thomas and Brugge, 1997). For example, Src binds K^+ ion channels and induces their phosphorylation, which decreases channel activity. In other cases the relationship between channel function and tyrosine kinase activity is less clear. Aggregation of acetylcholine receptor ion channels can lead to tyrosine phosphorylation of the receptor and association of nonreceptor kinase with the receptor. This and other kinase-channel interaction mechanisms resemble the binary receptor tyrosine kinases, although the consequences for downstream signalling in the specific cases are often unclear. In a number of cases Src family kinases can be activated via ion fluxes into cells through regulated channels. The mechanisms of these activations are varied, involving both receptor and nonreceptor processes. These examples again underline the versatility of the tyrosine kinase functions in many aspects of cell function.

SIGNALLING PATHWAYS AND CELLULAR FUNCTIONS OF TYROSINE KINASES

It should be obvious from the previous sections that tyrosine kinases can initiate multiple signalling pathways. In trying to understand the cellular functions of these kinases, it is useful to describe the downstream effectors and biological effects of these pathways. Such analyses are complicated by the fact that information about some of the pathways is incomplete. Moreover, many of the pathways are branched and have multiple effects. Finally, there are intersections of some pathways which cannot easily be represented by a linear or even a two-dimensional diagram. The major pathways described in this chapter include PLCγ, PI3 kinase and Ras–MAP kinase. However, there are variants of both the PI3 kinases and MAP kinases which involve different cellular functions. For example, at least three different classes of mammalian MAP kinases contribute to cell behaviour by regulating transcription: Erk, p38 and JNK. Each of these responds to different stimuli and results in a different cellular response. Moreover, it is important to remember that these individual pathways are not independent, but form a dynamic network.

The key to understanding the roles of tyrosine kinases is to be able to link these signalling pathways to the cellular functions elaborated by the tyrosine kinases. A preliminary and simplistic effort is shown in **Table 2**. However, all of these complex functions are usually the consequence of multiple pathways. As a result there is not necessarily a linear relationship between any kinase and a function. Moreover, the same pathway may be involved in different, almost contradictory, functions in different cellular contexts. EGF activation of the Ras–Erk pathway in PC12 cells induces proliferation, but NGF stimulation of the same pathway in these cells leads to differentiation. The difference appears to reside in the temporal aspects (kinetics) of the pathway. Another problem in analysing kinase functions is redundancy. This issue is illustrated from 'gene knockout' studies, in which the gene for a particular protein has been eliminated from the mouse genome for analysis of the phenotype of the mutant. For example, the gene for Src can be eliminated from the mouse without severe consequences for the reproduction of the animal. Only when three genes of the Src family are eliminated do the genetic defects become lethal. In contrast, gene deletions are lethal for ErbB2, ErbB3 or ErbB4 and the ErbB ligand heregulin because of a failure of heart development at about embryonic day 10. These studies indicate that the function of a particular kinase in an organism depends on its time and place of expression as well as the consequences of its downstream signalling in the cell and tissue of origin.

TYROSINE KINASES AND CANCER

The functions listed in **Table 2** are important because they include many of the cellular behaviours which are modified in neoplastic transformation of cells (Nicolson, 1976). Thus, the phenotypic changes in tumour cells correspond closely to the functions regulated by tyrosine kinases in cells. However, the relationship of tyrosine kinases and human cancer is not so simple. Formation of diagnosable human tumours appears to require about five genetic changes in a single cell lineage. Contrary to the original oncogene hypothesis, many of these genetic lesions are not in tyrosine kinase-related pathways for cell proliferation. This situation results because tissues have evolved 'tumour suppressors' to act as brakes for cell proliferation and tumour progression. Thus, removal of these suppression mechanisms is as important as enhancement of the progression mechanisms. Surprisingly, reversal of tyrosine phosphorylation by phosphatase action does not appear to be an important tumour-suppressor mechanism. The only tumour-suppressor phosphatase (PTEN) to be identified to date acts more robustly on inositol (lipid) phosphates than on protein phosphates. It may therefore be more important in countering the effects of PI3 kinase than tyrosine kinases.

To understand the role of tyrosine kinases in human cancer, it is instructive to analyse those cancers in which tyrosine kinases have been implicated as contributors (**Table 10**). The list is necessarily limited because providing evidence that a kinase contributes to tumour progression (cause versus effect) is difficult. Furthermore, downstream effectors in tyrosine kinase pathways, such as

Table 10 Tyrosine kinases implicated in neoplasia

Tyrosine kinase	Mechanism of activation	Cancer
ErbB1	Truncation	Glioma
ErbB2	Amplification	Breast
PDGFRβ	Chromosome translocation	Leukaemia
Kit	Mutation	Leukaemia
Met	Mutation, overexpression	Multiple cancers
FGFR	Mutation	Multiple myeloma
Ret	Chromosome translocation	Multiple endocrine
Alk	Chromosome translocation	Lymphoma
Src	Overexpression, activation	Multiple
Yes	Overexpression, activation	Multiple
Abl	Chromosome translocation	CML
JAK2	Chromosome translocation	Leukaemia

R, receptor.

Ras and PI3 kinase, are more proximal to the phenotypic changes in cells and tissues and may thus be more potent oncogenes. However, some tyrosine kinases are very potent oncogenes. A good example is ErbB2, which is an inducer of mammary cancer when its gene is expressed in the mammary gland of mice. Although the reason for this potency is not entirely clear, two observations are likely important. One is that ErbB2 is susceptible to mutations and other events which induce the activation of its tyrosine kinase. The second is that ErbB2 can activate multiple downstream pathways to change the mammary cell phenotype, including both Ras–MAP kinase and PI3 kinase pathways.

The list in **Table 10** suggests that these two factors are important in many instances in which tyrosine kinases contribute to cancer. Particularly noteworthy are activations of tyrosine kinase oncogenes by chromosome translocations. Since tyrosine kinases are frequently autoregulated, displacement of the regulatory regions by truncation (ErbB1) or by chromosome translocation (Abl) can remove the autorepression and activate the kinase. This mechanism is familiar in viral oncogenes. In the instance of v-erbB (ErbB1), the extracellular domain has been truncated, facilitating dimerization and activation of the intracellular kinase. In the case of v-Src, the C-terminal tail is missing. Since this sequence contains the phosphorylated tyrosine which interacts with the SH2 domain (**Figure 8**), phosphorylation of the enzyme at this site is no longer possible as a repression mechanism.

An alternative mechanism for activation of tyrosine kinases in cancer is the autocrine growth loop (Kolibaba and Druker, 1997), in which overexpression or mislocation of a ligand for the receptor kinase aberrantly induces its activation. The cause versus effect relationship of this mechanism in human cancer is also difficult to establish, although it has been well characterized in cell culture studies. Most of the evidence is based on statistical analyses of clinical outcomes. Such studies have implicated TGF-α in oesophageal cancer and insulin-like growth factor in prostate cancer.

OVERVIEW

Tyrosine kinases regulate a large number of cellular functions via an array of signalling pathways. Most, but not all, of these pathways originate at the cell plasma membrane via transmembrane receptor tyrosine kinases or binary receptor–tyrosine kinase couples. Both the tyrosine kinases and their effector pathways are highly regulated and must be integrated both spatially and temporally into the organisation of cellular functions. Mutational activation or overexpression of tyrosine kinases can lead to aberrant cellular behaviours, including neoplastic transformation and tumour progression.

REFERENCES

Carraway, K. L. and Carraway, C. A. C. (1995) Signaling, mitogenesis and the cytoskeleton: where the action is. *BioEssays*, **17**, 171–175.

Carraway, K. L., *et al*. (1997) Signaling and the Cytoskeleton, (Springer, Berlin).

Claesson-Welsh, L. (1994) Platelet-derived growth factor receptor signals. *Journal of Biological Chemistry*, **269**, 32023–32026.

Cohen, P. T. W. (1997) Novel protein serine/threonine phosphatases: variety is the spice of life. *Trends in Biochemical Science*, **22**, 245–251.

Hubbard, S. R., *et al*. (1998) Autoregulatory mechanisms in protein-tyrosine kinases. *Journal of Biological Chemistry*, **273**, 11987–11990.

Hunter, T. (1998) The phosphorylation of proteins on tyrosine: its role in cell growth and disease. *Philosophical Transactions of the Royal Society of London, Series B*, **353**, 583–605.

Kolibaba, K. S. and Druker, B. J. (1997) Protein tyrosine kinases and cancer. *Biochimica Biophysica Acta*, **1333**, F217–F248.

Nicolson, G. L. (1976) Transmembrane control of the receptors on normal and tumor cells. II. Surface changes associated with transformation and malignancy. *Biochimica Biophysica Acta*, **457**, 57–108.

Panayotou, G. and Waterfield, M. D. (1993) The assembly of signalling complexes by receptor tyrosine kinases. *BioEssays*, **15**, 171–177.

Peterson, E. J. and Koretzky, G. A. (1998) Signal transduction in T lymphocytes. *Clinical and Experimental Rheumatology*, **17**, 107–114.

Riese, D. J., II and Stern, D. F. (1998) Specificity within the EGF family/ErbB receptor family signalling network. *BioEssays*, **20**, 41–48.

Schlessinger, J. Lax, I. and Lemmon, M. (1995) Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell*, **83**, 357–360.

Sorkin, A. and Waters, C. M. (1993) Endocytosis of growth factor receptors. *BioEssays*, **15**, 375–382.

Streuli, M. (1996) Protein tyrosine phosphatases in signalling. *Current Opinions in Cell Biology*, **8**, 182–188.

Taylor, S. S., Radzio-Andzelm, E. and Hunter, T. (1995) How do protein kinases discriminate between serine/threonine and tyrosine? Structural insights from the insulin receptor protein–tyrosine kinase. *FASEB Journal*, **9**, 1255–1266.

Thomas, S. M. and Brugge, J. S. (1997) Cellular functions regulated by Src family kinases. *Annual Reviews of Cellular Development*, **13**, 513–609.

Wells, J. A. (1996) Binding in the growth hormone receptor complex. *Proceedings of the National Academy of Sciences of the USA*, **93**, 1–6.

Zou, X. and Calame, K. (1999) Signaling pathways activated by oncogenic forms of Abl tyrosine kinase. *Journal of Biological Chemistry*, **274**, 18141–18144.

FURTHER READING

Alberts, B., *et al*. (1994). *Molecular Biology of the Cell*, 3rd edn. (Garland Publishing, New York).

Hesketh, R. (1997). *The Oncogene and Tumor Suppressor Gene Factsbook* (Academic Press, San Diego).

Kellie, S. (1994). *Tyrosine Kinases and Neoplastic Transformation* (R. G. Landes, Austin, TX).

Lodish, H., *et al*. (1995). *Molecular Cell Biology*, 3rd edn. (Scientific American Books, New York).

Mendelsohn, J., *et al*. (eds) (1995). *The Molecular Basis of Cancer* (W. B. Saunders, Philadelphia).

Peters, G. and Vousden, K. H. (eds) (1997). *Oncogenes and Tumor Suppressors* (IRL Press, Oxford).

Ruddon, R. W. (1995). *Cancer Biology*, 3rd edn. (Oxford University Press, Oxford).

Websites

http://www.kinase.com/.

http://www.sdsc.edu/kinases/.

http://www.stke.org/.

http://www.mshri.on.ca/pawson/research2.html.

Chapter 13

Signalling by TGF-β

Jing Qing and Rik Derynck

University of California at San Francisco, San Francisco, CA, USA

CONTENTS

- Introduction
- TGF-β Expression and Activation
- Cellular Responses to TGF-β
- TGF-β Receptors and Interacting Proteins
- Smad and Non-smad Signalling
- Inactivation of TGF-β's Tumour-suppressor Functions in Carcinomas
- Cell-autonomous, Stimulatory Effects of TGF-β on Tumour Development
- Tumour Progression Stimulated by Effects of TGF-β on the Tumour Environment

INTRODUCTION

While normal, differentiating cells closely coordinate their proliferation and differentiation programmes, this balance is deregulated during malignant transformation into tumour cells. Compared with normal cells, transformed cells acquire a higher proliferative index and decreased differentiation, concomitant with a lower degree of dependence on extracellular cues. These changes in cell behaviour and responsiveness often result in part from genetic alterations, whereby defined mutations activate oncogenes or inactivate tumour-suppressor genes. Additionally, changes in the production of stimulatory or inhibitory growth and differentiation factors and/or in the cellular responsiveness to these factors greatly contribute to the behavioural and phenotypic changes in tumour cells. In fact, tumour cells often display autocrine responsiveness to increased endogenous expression of growth factors, and this autocrine stimulation contributes to tumour formation and cancer progression. Similarly, the invasive and metastatic phenotype of the tumour cells also results from mutations, changes in gene expression and/or altered production of cell surface-associated or extracellular mediators, and altered responsiveness to these factors.

Various growth factors and cytokines have been implicated in the progression and behaviour of tumour cells and cancers. Among these, transforming growth factor-β (TGF-β) and its downstream effectors are key determinants of the tumour cell behaviour of carcinomas. The TGF-β production by tumour cells and the responsiveness of tumour cells to autocrine TGF-β and TGF-β in the tumour microenvironment exert both positive and negative effects on cancer development. Accordingly, TGF-β and the TGF-β signalling pathway have been considered as both a tumour suppressor and a promoter of tumour progression and invasion. This chapter introduces the role of TGF-β in tumour development, specifically of carcinomas, and summarizes our knowledge of the TGF-β signalling mechanisms, i.e. the cell surface receptors and downstream effector proteins.

TGF-β EXPRESSION AND ACTIVATION

TGF-β is a secreted polypeptide, which in its receptor-binding, fully active form is a disulfide-bonded and stable dimer. The TGF-β precursor polypeptide comprises an N-terminal signal peptide followed by a large prosegment, which is about twice as long as the C-terminal, active TGF-β sequence. Protein purification and cDNA cloning have revealed the existence of three TGF-βs, TGF-β1, -β2 and -β3, each encoded by a different gene and made as a precursor with similar structural organisation. TGF-β1 was the first identified and best characterized TGF-β, and is therefore considered as the prototype of these three TGF-βs, which act through the same receptor system and have similar biological activities in cell culture (Derynck and Choy, 1998). TGF-β1 is also the prototype of the TGF-β superfamily, which comprises about 60–70 structurally related proteins, characterized by seven characteristically spaced cysteines in the C-terminal segment of the precursor polypeptide (Derynck and Feng, 1997; Piek *et al*., 1999; Massagué *et al*., 2000). All TGF-β superfamily members are thought to act as dimers that bind to structurally related receptors. TGF-β-related proteins are found in multicellular eukaryotes from *C. elegans* and *Drosophila* to all vertebrates and regulate a variety of developmental and differentiation processes.

TGF-β1 expression is often upregulated in tumour cells and is therefore the focus of most studies on the role of TGF-β in tumorigenesis. With few exceptions, all cells in culture express TGF-β1, but this may be due to the absence of an extracellular matrix substrate. Thus, TGF-β1 expression is repressed when cells are grown on extracellular matrix or once they have deposited their extracellular matrix, a process strongly activated by TGF-β (Derynck and Choy, 1998). TGF-β1 expression in cell culture is therefore not necessarily a reflection of TGF-β expression *in vivo*, but may be a response to injury, consistent with increased TGF-β1 expression at sites of tissue injury. TGF-β itself activates TGF-β1 expression, a positive feedback and amplification response of relevance in tumour development.

Following an intracellular proteolytic cleavage between the prosequence and the active TGF-β sequence, TGF-β is released as a 'latent' complex, which is incapable of binding to the TGF-β receptors and is consequently biologically inactive. This latent complex consists of an active TGF-β dimer in a noncovalent complex with two prosegments, to which one of several 'latent TGF-β binding proteins' is often disulfide linked (Munger *et al*., 1997). This latent complex represents an important safeguard against 'unneeded' or 'inadvertent' activation and these binding proteins may stabilize and target latent TGF-β to the extracellular matrix, where it is stably sequestered (Munger *et al*., 1997). The extracellular matrix thus acts as a reservoir, from which TGF-β can readily be made available to cells without the need to induce TGF-β synthesis.

The secretion of TGF-β as a latent complex necessitates the existence of a regulated activation process. While latent TGF-β is efficiently activated by acidic conditions, its physiological activation occurs most likely through proteases, which degrade the TGF-β prosegments and thereby release the highly stable, active TGF-β dimer. Since plasmin activates latent TGF-β and plasminogen is converted into plasmin at sites of cell migration and invasion, we assume that plasmin-mediated activation of TGF-β occurs at sites of angiogenesis and tumour development, thus exposing endothelial and tumour cells to active TGF-β. Matrix metalloproteases may also play a key role in activation of latent TGF-β. For example, the matrix metalloproteases MMP-9 and MMP-2 have the ability to activate latent TGF-β (Yu and Stamenkovic, 2000). Since metalloproteases are frequently expressed by malignant cells, this mechanism may locally activate TGF-β at sites of tumour cell invasion. Other mechanisms of activation may not depend on proteases. For example, the extracellular matrix protein thrombospondin (Ribeiro *et al*., 1999) and the $\alpha v\beta 6$ integrin, which is expressed at the surface of epithelial cells in response to inflammation (Munger *et al*., 1999), may activate TGF-β through a conformational change in the TGF-β complex. Thus different mechanisms may regulate TGF-β activation in different physiological contexts, and tumour cells are well equipped to activate TGF-β locally. Finally, high-energy irradiation, which is often used in cancer therapy, also activates TGF-β (Barcellos-Hoff *et al*., 1994).

CELLULAR RESPONSES TO TGF-β

TGF-β exerts a large variety of cellular responses, which often depend on the cell type and physiological conditions, thus making the responses highly context dependent. Since the role of TGF-β in tumour development is best studied in carcinoma development, we will briefly describe the cellular responses to TGF-β that play a role in the development of carcinomas from normal epithelial cells.

While TGF-β stimulates proliferation of various mesenchymal cell types, including fibroblasts, it is a very potent inhibitor of proliferation of epithelial cells and cells of haematopoietic origin, including various immune cells (Derynck and Choy, 1998; Massagué *et al*., 2000). The TGF-β signalling process that leads to growth inhibition is considered as a tumour suppressor pathway, which is often inactivated to allow tumour development and cancer progression (de Caestecker *et al*., 2000; Massagué *et al*., 2000). Cell cycle progression and proliferation are driven by a complex and interdependent regulation by cyclins, cyclin-dependent kinases (cdks) and cdk inhibitors, which form complexes with each other (Sherr and Roberts, 1999). Accordingly, TGF-β-induced growth arrest, which occurs in the late G_1 phase of the cell cycle, induces a variety of alterations in the levels and activities of cyclins, cdks and cdk inhibitors, although many of these effects could be considered as indirect. Several direct TGF-β effects may play key roles in TGF-β-mediated growth arrest.

The key event that leads to TGF-β-induced growth arrest is the induction of expression of the cdk inhibitors $p15^{Ink4B}$ and/or $p21^{Cip1}$, depending on the cell type and context (Massagué *et al*., 2000) (**Figure 1**). $p21^{Cip1}$, similarly to the related $p27^{Kip1}$, interacts with complexes of cdk2 with cyclin E or A and thereby inhibits the cdk2 activity, thus preventing cell cycle progression. $p21^{Cip1}$ and $p27^{Kip1}$ also interact with cdk4 and cdk6 and stabilize their interaction with cyclin D, thus playing a role in the activation of these complexes (Sherr and Roberts, 1999). Since the complexes of cdk2 with cyclin E or A act downstream from the complexes of cdk4 or 6 with cyclin D, TGF-β-induced expression of $p21^{Cip1}$ leads to cell cycle arrest in late G_1. In contrast, $p15^{Ink4B}$ interacts with and inactivates cdk4 and 6 and prevents their complex formation with cyclin D. In addition, $p15^{Ink4B}$ binds to the complexes of cdk4 or 6 with cyclin D and thereby not only inactivates the catalytic activity of these cdks, but also displaces $p21^{Cip1}$ and $p27^{Kip1}$ from these complexes, thus allowing them to bind to and inactivate the cdk2 complexes with cyclin A and E. Consequently,

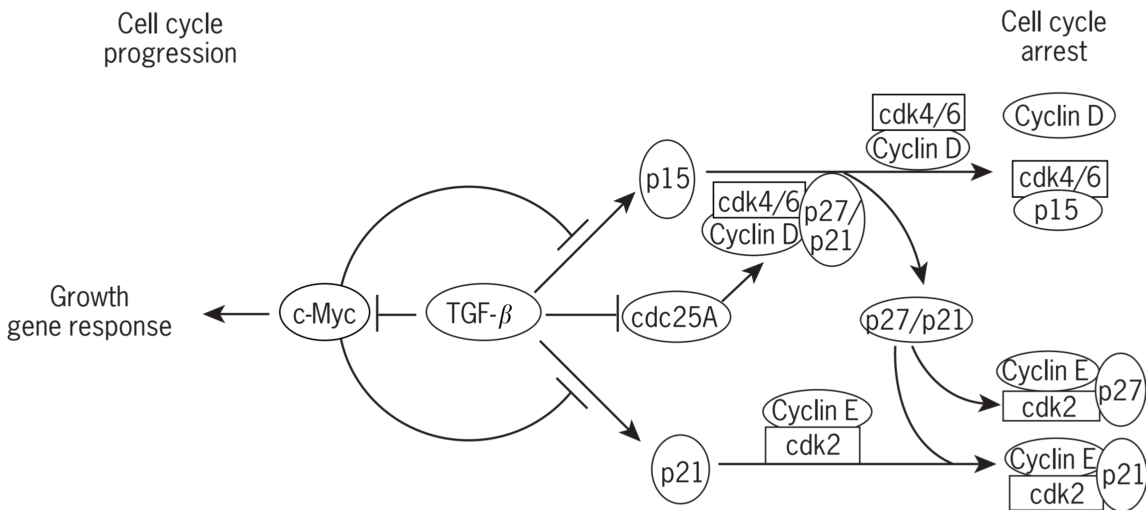


Figure 1 Schematic representation of TGF-β-induced mechanism of growth arrest in late G_1. TGF-β induces the expression of p15Ink4B and p21Cip1 and down-regulates the expression of c-myc and cdc25A. The down-regulation of c-myc expression relieves the inhibition of p15Ink4B and p21Cip1 expression, thus allowing induction of these cdk inhibitors by Smads.

TGF-β-induced expression of $p15^{Ink4B}$ inactivates cell cycle progression at two levels, $p15^{Ink4B}$-mediated inactivation of cdk4 and cdk6 and, as a result of $p21^{Cip1}$ or $p27^{Kip1}$ displacement, inactivation of cdk2 through $p21^{Cip1}$ or $p27^{Kip1}$ binding (Massagué *et al.*, 2000).

Additional mechanisms may also contribute to TGF-β-mediated growth arrest, again depending on the cell type. For example, TGF-β inhibits cdk4 expression, presumably at the translational level (Ewen *et al.*, 1993). TGF-β also down-regulates the levels of cdc25A (Iavarone and Massagué, 1997), a tyrosine phosphatase, and this decreased activity leads to increased tyrosine phosphorylation of cdk4 and cdk6 and consequent inhibition of their kinase activity. TGF-β also inhibits c-Myc expression in normal epithelial cells, and c-Myc plays a role in growth arrest in response to TGF-β (Massagué *et al.*, 2000; Chen *et al.*, 2001). High levels of c-Myc repress $p15^{Ink4B}$ and $p21^{Cip1}$ expression, and a decrease in c-Myc levels results in derepression, thus allowing TGF-β-induced transcription of the $p15^{Ink4B}$ and $p21^{Cip1}$ genes. The TGF-β-mediated decrease in c-Myc levels may also play a role in the down-regulation of cdc25A expression in response to TGF-β, since c-Myc may positively regulate cdc25A expression. Finally, the TGF-β-induced interaction of protein phosphatase 2A with S6 kinase and consequently decreased S6 kinase activity may contribute to TGF-β-mediated growth arrest (Petritsch *et al.*, 2000).

TGF-β regulates the expression of a large variety of genes through activation or repression of transcription. Among the many genes, TGF-β regulates the expression of transcription factors, secreted cytokines and growth factors, extracellular matrix proteins, proteases and integrins (Derynck and Choy, 1998; Massagué *et al.*, 2000). Which genes are regulated by TGF-β and the extent of this regulation are highly cell type and context dependent. The induction of transcription factors by TGF-β results in a variety of indirect responses, thus enhancing the complexity of the TGF-β response. Similarly, the induction of cytokine and growth factor expression, e.g. interleukin 1 and PDGF, by TGF-β results in indirect, yet physiologically very important, cellular responses. The potent ability of TGF-β to induce the expression of extracellular matrix proteins stands in contrast to other growth factors. TGF-β induces the expression of some, but not all, common extracellular matrix proteins, including collagens and fibronectin. Consequently, TGF-β expression and activation are major determinants of extracellular matrix synthesis and deposition. This activity is often complemented by a TGF-β-induced decrease in protease activity and increased expression of protease inhibitors, which together enhance the increased extracellular matrix deposition. TGF-β also enhances the expression of a variety of integrins, depending on the cell type, often resulting in increased cell adhesion to the extracellular matrix and presumably also increased integrin signalling. Finally, TGF-β is a potent chemoattractant of monocytes, macrophages and fibroblasts (Derynck and Choy, 1998). Consequently, TGF-β activation often results in localized inflammation and chemoattraction of fibroblasts, which, together with the mitogenic effect of TGF-β on fibroblasts and stimulation of extracellular matrix deposition, results in fibrosis (Roberts *et al.*, 1986). These responses may also explain radiation-induced fibrosis, a consequence of TGF-β activation (Barcellos-Hoff *et al.*, 1994).

TGF-β RECEPTORS AND INTERACTING PROTEINS

TGF-β signals through a cell surface receptor complex of two types of transmembrane serine/threonine kinases, named type I and type II receptors (Derynck and Feng, 1997; Massagué, 2000) (**Figure 2**). The type II and type I receptors are structurally related with a high level of sequence conservation in their cytoplasmic kinase domains, besides similarities in their extracellular domains. They autophosphorylate and phosphorylate target proteins on serines and threonines, yet also have tyrosine kinase activity. Both receptor types form cell surface dimers in the absence of TGF-β binding. These type II and type I receptor dimers also have an intrinsic affinity for each other, thus resulting in a heterotetrameric complex, which is stabilized by TGF-β binding. Among the receptors for TGF-β superfamily members, several heteromeric combinations of type II and type I receptors form functional signalling complexes, thus providing a variety of combinatorial type II/type I receptor interactions. In the case of TGF-β, only one type II TGF-β receptor, i.e. TβRII, has been identified. Three different type I receptors have been proposed as type I TGF-β receptors. Most gene expression responses in most cell types are mediated by the TβRI receptor, while the potential roles of the ALK-1/TSR-1 and ALK-2/ActRI/Tsk7L type I receptors in TGF-β signalling remain to be better explored. TGF-β receptors are expressed on most, if not all, cell types (Derynck and Feng, 1997).

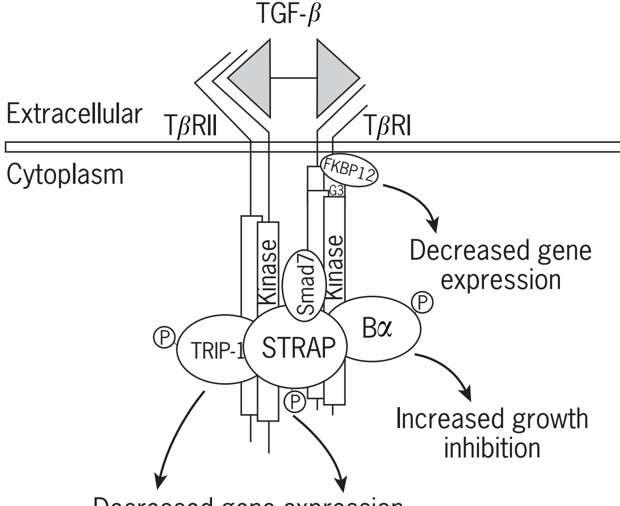


Figure 2 The TGF-β receptor complex with interacting regulatory proteins. The receptor complex consists of two TβRII and two TβRI polypeptides. FKBP12 interacts constitutively with TβRI, while the WD-repeat proteins TRIP-1, STRAP and the Bα subunit of protein phosphatase 2A interact following ligand activation of kinases and are phosphorylated. TRIP-1 interacts with TβRII, while the Bα subunit of protein phosphatase 2A interacts with TβRI. STRAP can interact with either receptor and also interacts and synergizes with Smad7.

All three TGF-βs interact with the same TGF-β receptor complex (Derynck and Feng, 1997). TGF-β1 and -β3 interact primarily with the TβRII receptor, yet also contact the type I receptor in the complex. Thus, TβRII binds TGF-β1 and -β3 without a requirement for a type I receptor. In contrast, TGF-β2 does not bind to either TβRII or TβRI alone, but binds efficiently to the heteromeric receptor complex. TGF-β binding to the receptor complexes is enhanced in the presence of the type III receptors β-glycan and endoglin. β-Glycan is an abundant cell surface proteoglycan with a short cytoplasmic domain and no known signalling function. β-Glycan binds all three TGF-βs with high efficiency and may enhance the efficiency of receptor binding of TGF-β, most notably TGF-β2. Endoglin, a structurally related glycoprotein, is primarily expressed by vascular endothelial cells. Endoglin binds TGF-β1 and -β3 with high affinity, but not TGF-β2, and may enhance TGF-β's presentation to the TGF-β signalling receptor complex, similarly to β-glycan (Derynck and Feng, 1997).

Following TGF-β binding to the receptor complex, signalling is initiated through the activities of the cytoplasmic kinase domains of both receptors. The TβRII kinase, when overexpressed, is constitutively active and phosphorylated. Whether this is also the case at endogenous expression levels or whether TGF-β induces TβRII autophosphorylation remains to be clarified. A key event in receptor activation is the phosphorylation, and consequent activation, of the TβRI kinase by the TβRII kinase. TβRII phosphorylates the cytoplasmic domain of TβRI on serine and threonine residues in the 'GS sequence', a sequence which is conserved among the type I receptors and is located immediately upstream from the kinase domain. This phosphorylation then presumably induces a conformational change in the TβRI cytoplasmic domain that activates the kinase and consequently allows TβRI autophosphorylation and phosphorylation of downstream target proteins (Derynck and Feng, 1997; Massagué, 2000).

Several proteins have been shown to interact with the cell surface TGF-β receptor complexes (**Figure 2**). Among these, FKBP12 interacts with the juxtamembrane domain of type I receptors and may regulate its conformation, and dampens the TGF-β receptor activation. FKBP12 interacts constitutively with, yet is not phosphorylated by, the TGF-β receptor (Chen *et al.*, 1997). In contrast, three WD-repeat containing proteins associate with and are phosphorylated by the receptor complex following ligand-induced activation. TRIP-1 interacts with TβRII (Choy and Derynck, 1998), while the Bα subunit of the protein phosphatase 2A interacts with type I receptors (Griswold-Prenner *et al.*, 1998). STRAP, on the other

hand, interacts with both TβRII and TβRI and in turn interacts with Smad7 (see below) to decrease TGF-β signalling (Datta and Moses, 2000). Since WD repeats mediate protein–protein interactions, it is likely that these proteins allow interactions of multiprotein complexes with the receptors. These proteins regulate the TGF-β receptor response, but there is currently no solid evidence that they act as effectors of TGF-β responses.

SMAD AND NON-SMAD SIGNALLING

The only characterized signalling effector pathway, initiated by activated TGF-β receptors, is provided by the Smads (Piek *et al*., 1999; Itoh *et al*., 2000; Massagué, 2000; Massagué *et al*., 2000). The Smads, a family of structurally related proteins, are directly activated by the receptors and then translocate into the nucleus to act as ligand-dependent transcriptional regulators of target genes. The Smads are structurally related to each other in two domains, an N- or MH1 domain which corresponds to the N-terminal third of the protein, and a C- or MH2 domain which corresponds to the C-terminal third of these proteins. Based on structural and functional characteristics, the Smads are divided into three subfamilies. Smad1, -5 and -8 and Smad2 and -3 are 'receptor-activated' Smads that are phosphorylated on C-terminal serines by the activated type I receptor. The activated TβRI phosphorylates and thereby activates Smad2 and -3, whereas Smad1, -5 and -8 are activated by BMP receptors. Following their release from the receptors, the activated Smads form a heterotrimeric complex with Smad4, which serves as a common mediator for all receptor-activated Smads. This complex consists of two receptor-activated Smads and one Smad4, raising the possibility of combinatorial interactions among receptor-activated Smads. This Smad complex translocates into the nucleus and cooperates with other transcription factors to activate or repress transcription of defined genes in response to TGF-β (Piek *et al*., 1999; Itoh *et al*., 2000; Massagué, 2000; Massagué *et al*., 2000) (**Figure 3**).

Besides the receptor-activated Smads and Smad4, which act as ligand-induced effectors, two other 'inhibitory' Smads regulate Smad signalling. Smad6 and -7 have much less sequence conservation in their MH1 domain than the other Smads, and interfere with the activation of effector Smads (Piek *et al*., 1999; Itoh *et al*., 2000; Massagué, 2000). Smad6 and -7 are able to interact with type I receptors, thus competitively preventing the 'receptor-activated' Smads from being phosphorylated, whereas Smad6 additionally interferes with the heteromeric complex formation of Smad1 with Smad4. Although Smad6 and -7 seem to interact nonspecifically with type I receptors, Smad7 may primarily inhibit TGF-β signalling, whereas Smad6 primarily inhibits BMP signalling.

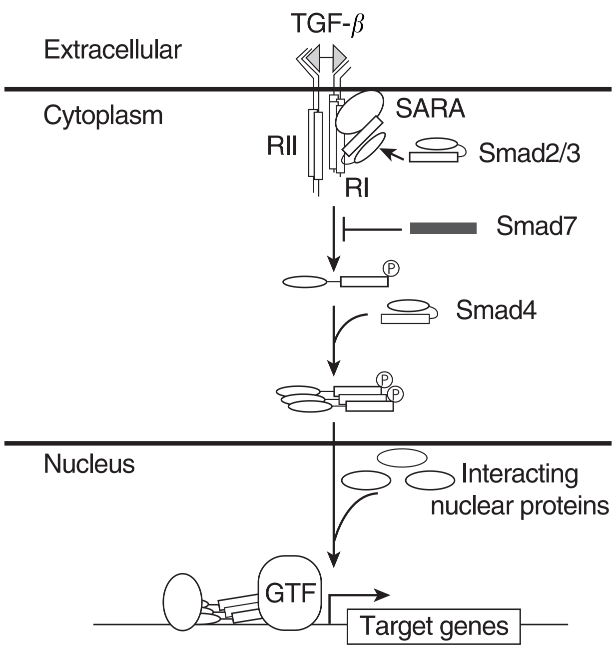


Figure 3 TGF-β-induced signalling through Smads. Following ligand-induced receptor activation, Smad2 and/or -3 interact transiently with the TβRI receptor and this interaction is stabilized by the FYVE-protein SARA. Smad2 and -3 are C-terminally phosphorylated by TβRI and then dissociate from the receptor to form a heterotrimeric complex consisting of two receptor-activated Smads and Smad4. This complex is then translocated into the nucleus where it interacts at the promoter with other transcription factors to regulate gene expression. Smad7 inhibits activation of Smad2 and/or -3 by the receptors.

Accordingly, TGF-β signalling induces Smad7 expression, thus providing a TGF-β induced negative feedback loop, whereas BMP signalling induces Smad6 expression. While the inhibitory functions of Smad6 and -7 can be easily conceptualized as a mechanism of competitive inhibition, the functions of Smad6 and -7 may be more complex (Piek *et al*., 1999; Itoh *et al*., 2000). This is suggested by the observations that Smad6 and -7 are primarily localized in the nucleus and that both Smads can cooperate with TGF-β signalling in inhibiting adipocyte differentiation. Smad7 may also act as an effector of TGF-β-induced cell death, whereas Smad6 can function as a corepressor (Piek *et al*., 1999; Itoh *et al*., 2000). Further studies are needed to define the functions of Smad6 and -7.

Once inside the nucleus, the Smad complexes function as ligand-dependent transcriptional regulators of target genes (Derynck *et al*., 1998; Piek *et al*., 1999; Massagué, 2000; Massagué and Wotton, 2000). TGF-β activates or represses transcription of defined target genes, and many of these responses are direct, i.e. immediate early, responses to TGF-β receptor activation. Smads have

mainly been characterized as TGF-β-induced transcription activators, although they may also mediate TGF-β-induced repression of gene expression. Smads activate transcription from a variety of promoters through physical interaction with the transcriptional coactivator CBP/p300, a variety of DNA-binding transcription factors, and defined sequences in the promoter DNA (Derynck *et al*., 1998; Massagué, 2000; Massagué and Wotton, 2000) (**Figure 4**). Most TGF-β-inducible genes are transcriptionally activated through Smad3, even though TGF-β activates both Smad2 and -3, and some responses, e.g. the induction of $p15^{Ink4B}$ and $p21^{Cip1}$, require the participation and cooperation of both Smad2 and -3 (Feng *et al*., 2000; Pardali *et al*., 2000).

The receptor-activated Smads link to the general transcriptional machinery through a direct association with the transcriptional coactivator CBP/p300. This interaction, which is stabilized by Smad4, occurs through the C-terminal sequence of the Smads, and requires receptor-mediated phosphorylation of the C-terminal serines. CBP/p300 thus serves as transcriptional coactivator for Smads, and this interaction is essential for transcriptional activation (Derynck *et al*., 1998; Massagué, 2000; Massagué and Wotton, 2000). Smads also interact with a variety of DNA binding transcription factors and it is this interaction, together with the DNA binding of Smads, which specifies the promoter binding and transcriptional activation (Derynck *et al*., 1998; Itoh *et al*., 2000; Massagué, 2000; Massagué and Wotton, 2000). The versatility of physical interactions with a variety of transcription factors is best illustrated for Smad3. Among others, Smad3 has been shown to interact with bZIP transcription factors (e.g. c-Jun, ATF-2 or CREB), bHLH transcription factors (e.g. TFE3), runt domain transcription factors (AML-1 and -2 and CBFA1/AML3), nuclear receptors (e.g. the vitamin D_3 and glucocorticoid receptors) and Sp1, either through the MH1- or MH2-domains of the Smad, depending on the interacting transcription factor. While this interaction is usually direct, stable interaction with the transcription factor may depend on the concomitant interaction of the Smad and the transcription factor with CBP/p300. This physical interaction not only localizes the interacting Smad to a defined promoter sequence, but also allows the Smad to enhance the activity of the interacting transcription factor, presumably as a result of the increased interaction with CBP/p300.

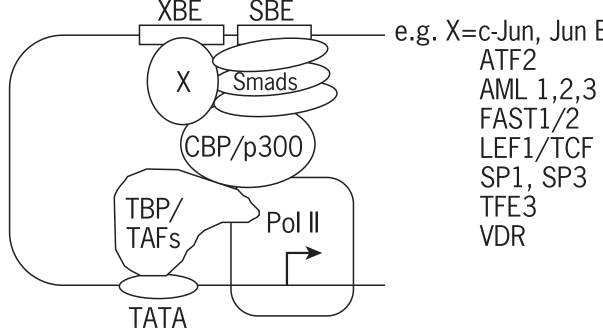


Figure 4 Physical and transcriptional cooperation of TGF-β-activated Smad2 or -3 with other transcription factors at the promoter. The heterotrimeric Smad complex interacts with the CBP/p300 coactivator, which connects to the general transcription machinery. While also interacting with DNA at a Smad binding element (SBE), the Smad complex interacts with one of several possible transcription factors with sequence-specific DNA binding (XBE).

Thus, Smads often serve as coactivators of other transcription factors. This, however, may not always be the case. For example, Smad2 interacts with the DNA binding protein FAST1/2 at an activin-responsive promoter sequence. FAST1/2 does not have transcription activity by itself and therefore serves as a DNA sequence-specific scaffold to allow transcriptional activation by Smad2. Also, the TGF-β-induced transcription of the Smad7 gene is mediated by Smad3, presumably without involvement of other DNA binding, interacting transcription factors. This opportunity may be provided by tandem Smad-binding DNA sequences, which enable permit Smad3 binding to the Smad7 promoter. Finally, while the interacting transcription factor provides high-affinity, sequence-specific binding to the promoter, Smad3 and -4, but not Smad2, are also able to bind DNA. However, Smads have a much lower DNA binding affinity and sequence specificity than most DNA binding transcription factors. Thus, a Smad binding sequence may primarily provide a sequence context conducive to Smad binding in close proximity to the sequence for the high-affinity DNA binding, interacting transcription factor. In this way, the DNA context-dependent binding of a Smad to both the interacting transcription factor and the promoter DNA may explain why TGF-β activates only a select set of the promoters, which show productive DNA binding of the Smad-interacting transcription factor (Derynck *et al*., 1998; Itoh *et al*., 2000; Massagué, 2000; Massagué and Wotton, 2000). This cooperation of Smads with other DNA-binding transcription factors also explains why, prior to the characterization of Smads, no consensus TGF-β response element could be defined in TGF-β-responsive genes and why AP-1 and Sp1 binding sequences were shown to be required for TGF-β responsiveness of various promoters.

The TGF-β-dependent recruitment of Smad complexes to the transcription machinery also allows for interactions with additional coactivators or corepressors, which regulate the amplitude of TGF-β-dependent transcriptional activation through Smads (Itoh *et al*., 2000; Massagué, 2000; Massagué and Wotton, 2000). Besides the interaction of the Smad complex with the CBP/p300 coactivator, Smad4 is able to engage the MSG1 coactivator into the transcription complex to enhance the Smad response. In contrast, recruitment of a corepressor into the complex decreases or inhibits the Smad and TGF-β response.

Several corepressors can interact directly with TGF-β Smads to decrease TGF-β responses. For example, the proto-oncogene product Evi-1 interacts with Smad3 and represses the gene expression and growth-inhibitory responses of TGF-β. Similarly, c-Ski also interacts with receptor-activated Smads and recruits a histone deacetylase to repress Smad-mediated transcription. A similar mechanism may also explain the inhibition by other Smad-interacting corepressors, such as the c-Ski-related SnoN, TGIF, SNIP and SIP1, although, in some cases, their interaction with CBP/p300, or interference with the Smad interaction with CBP/p300, may also play a role (Itoh *et al*., 2000; Massagué, 2000; Massagué and Wotton, 2000). Enhanced expression of corepressors in tumour cells could selectively block the TGF-β response and its growth-inhibitory effect in tumour cells and thus contribute to cancer progression.

The physical interactions and functional cooperativity of Smads with other transcription factors allows for cross-talk with other signalling pathways (Zhang and Derynck, 1999; Massagué, 2000; ten Dijke *et al*., 2000) (**Figure 5**). For example, UV radiation, stress and mitogens activate MAP kinase pathways that activate Jun N-terminal kinase, which phosphorylates and activates c-Jun. Activation of the c-Jun–c-Fos complex by MAP kinase–JNK cascades is therefore likely to regulate TGF-β-induced transcription from promoters with TGF-β-responsive AP-1- and Smad-binding sites. These promoters thus represent targets of convergence for these two types of signalling pathways. TGF-β signalling can also converge with Wnt signalling. This pathway is mediated by β-catenin, which shuttles between the cytoplasm and the nucleus, and serves as coactivator for the nuclear LEF/TCF transcription factors. Smad4 forms a complex with β-catenin and LEF, while Smad3 associates with LEF/TCF, and both Smads can cooperate with these transcription factors at promoters with LEF/TCF and Smad binding sites. Finally, Smad signalling can also converge with STAT signalling. STATs are activated by receptors and translocate into the nucleus to act as DNA binding transcription factors. BMP-activated Smad1 has been shown to form a complex with STAT3, which was activated in response to LIF. This interaction required the presence of the coactivator p300, with which both transcription factors interact. This complex formation is presumably at the basis of the functional cooperativity between Smad and STAT signalling.

The cross-talk of TGF-β/Smad signalling with other pathways can also occur prior to nuclear translocation of

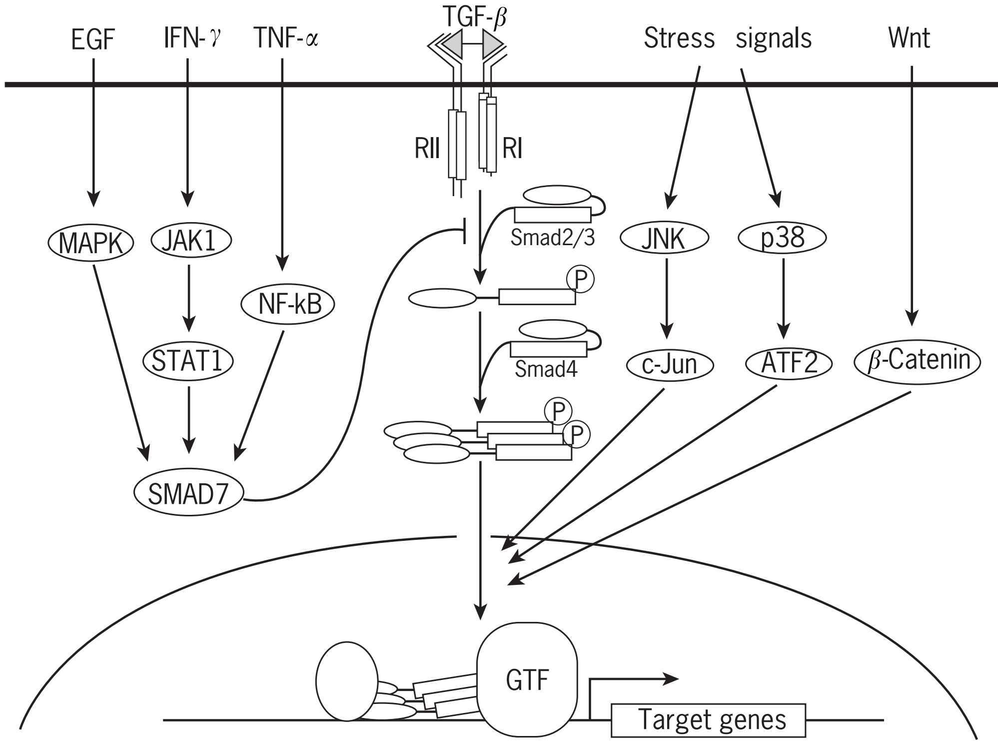


Figure 5 Cross-talk with TGF-β-induced signalling through Smads. A variety of extracellular signals activate JNK and/or p38 MAP kinase signalling, which leads to phosphorylation of c-Jun or ATF-2, two transcription factors, with which Smads can cooperate at defined promoters. Wnt signalling also cross-talks with Smad signalling, through the ability of activated Smads to associate with β-catenin and/or LEF/TCF at some promoters. Activation of MAP kinase signalling by growth factors, such as EGF, and activation of STAT1 by interferon-γ and of NF-κB by TNF-α all induce Smad7 expression, thus leading to inhibition of TGF-β-induced Smad2/3 activation by Smad7.

Smads (Piek *et al*., 1999; Zhang *et al*., 1999; Massagué, 2000; ten Dijke *et al*., 2000). For example, the linker regions of Smad1, -2 and -3 can be phosphorylated by Erk MAP kinase, which is activated in response to receptor tyrosine kinase and Ras signalling, and this phosphorylation then regulates ligand-induced nuclear translocation (Piek *et al*., 1999). Some signalling pathways activate expression of an inhibitory Smad, thus leading to decreased Smad activation. While BMPs and TGF-β induce the expression of the inhibitory Smad6 and -7, thus activating autoregulatory negative feedback loops, EGF can also induce Smad6 and -7 expression. This observation suggests that receptor tyrosine kinases regulate Smad signalling through induction of an inhibitory Smad, although other observations suggest Smad activation in response to receptor tyrosine kinase activation. Finally, interferon-γ-induced signalling through STATs and TNF-α-induced activation of NF-κB also activate Smad7 expression, which in turn inhibits TGF-β/Smad signalling. Thus, upregulation of Smad7 may represent a convenient mechanism used by several signalling pathways to inhibit TGF-β responsiveness (Piek *et al*., 1999; Zhang and Derynck, 1999; Itoh *et al*., 2000; Massagué, 2000; ten Dijke *et al*., 2000).

The mechanisms of Smad signalling and regulation, outlined above, now also explain the versatility and context dependence of the TGF-β responses. Indeed, the levels of interacting transcription factors and their activation state, as regulated by other signalling pathways, are important determinants of the TGF-β response. For example, TGF-β-induced transcription resulting from interactions of Smads with AP-1 transcription factors are likely to be regulated by MAP kinase signalling (Piek *et al*., 1999; Zhang and Derynck, 1999; Itoh *et al*., 2000; Massagué, 2000; ten Dijke *et al*., 2000). In contrast, the interaction of the heteromeric Smad2–Smad3–Smad4 complex with Sp1 induces expression of the cdk-inhibitors $p15^{Ink4B}$ or $p21^{Cip1}$ in a manner that is unlikely to require MAP kinase signalling (Feng *et al*., 2000; Pardali *et al*., 2000). Finally, the presence and identity of corepressors may determine the amplitude of the Smad response and even whether TGF-β receptor activation results in transcriptional activation or repression.

Finally, recent evidence strongly suggests that TGF-β signals through other pathways, distinct from the Smad-mediated regulation of gene expression (Piek *et al*., 1999; Itoh *et al*., 2000; Massagué, 2000) (**Figure 6**). TGF-β is able to activate MAP kinase signalling, although the extent and kinetics of activation differ substantially among different cell lines and types. These activation events have been shown to result in activation of Erk MAP kinase, p38 MAP kinase and Jun N-terminal kinase (JNK). p38 MAP kinase and JNK activation occurs through phosphorylation by MKK3, MKK4 and/or MKK6. Activation of p38 MAP kinase and JNK by TGF-β may enhance Smad signalling through Smad phosphorylation or phosphorylation of c-Jun and ATF-2, transcription factors that cooperate with TGF-β Smads, thus resulting in cross-talk with Smad-mediated transcription. TAK1, a MAP kinase, which is rapidly activated by TGF-β, yet is also involved in other unrelated signalling pathways, may serve as a TGF-β-dependent initiator of these signalling cascades. TGF-β may also activate or stabilize the small GTPases RhoA and RhoB and these may in turn play roles in several TGF-β responses, e.g. a requirement of RhoB for JNK activation (Itoh *et al*., 2000; Massagué, 2000). Finally, TGF-β also induces an interaction of protein phosphatase 2A with p70/S6 kinase, a kinase known to regulate protein translation and growth control, thus decreasing its activity (Petritsch *et al*., 2000). While the mechanisms of activation by TGF-β and roles of these non-Smad signalling cascades remain to be characterized, these observations strongly suggest that inactivation of Smad pathways may not leave the cell unresponsive to TGF-β.

INACTIVATION OF TGF-β's TUMOUR-SUPPRESSOR FUNCTIONS IN CARCINOMAS

The growth-inhibitory response of epithelial cells to TGF-β strongly suggests that TGF-β signalling may exert tumour suppression. On the other hand, TGF-β1 expression is often upregulated in carcinomas, which then would suggest that tumour cells benefit from TGF-β expression. Accordingly, there is substantial evidence for both tumour promoting and tumour-suppressor roles of TGF-β in carcinoma development. The tumour-suppressor role of TGF-β signalling is best supported by the presence of inactivating mutations in TGF-β receptors and Smads in human carcinomas and by tumour development studies in mouse models.

Somatic mutations in TβRII are common in tumours from patients with hereditary nonpolyposis colorectal cancer (HNPCC), who have germ-line defects in their capacity for DNA mismatch repair (de Caestecker *et al*., 2000; Massagué *et al*., 2000) (**Figure 7**). Nucleotide additions to or deletions from the stretch of adenines within the *TβRII* coding sequence give rise to a truncated TβRII, which is incapable of signalling. Consequently, the cells acquire a selective growth advantage, allowing them to progress for tumour development. While TβRII mutations occur frequently in colon cancers, gastric cancers and gliomas with microsatellite instability, they are less common in tumours from the endometrium, pancreas, liver and breast with microsatellite instability. Missense and inactivating mutations in the kinase domain of TβRII have also been reported in colon cancers, which do not display microsatellite instability. Together, inactivating TβRII mutations may be present in 20–25% of all colon cancers.

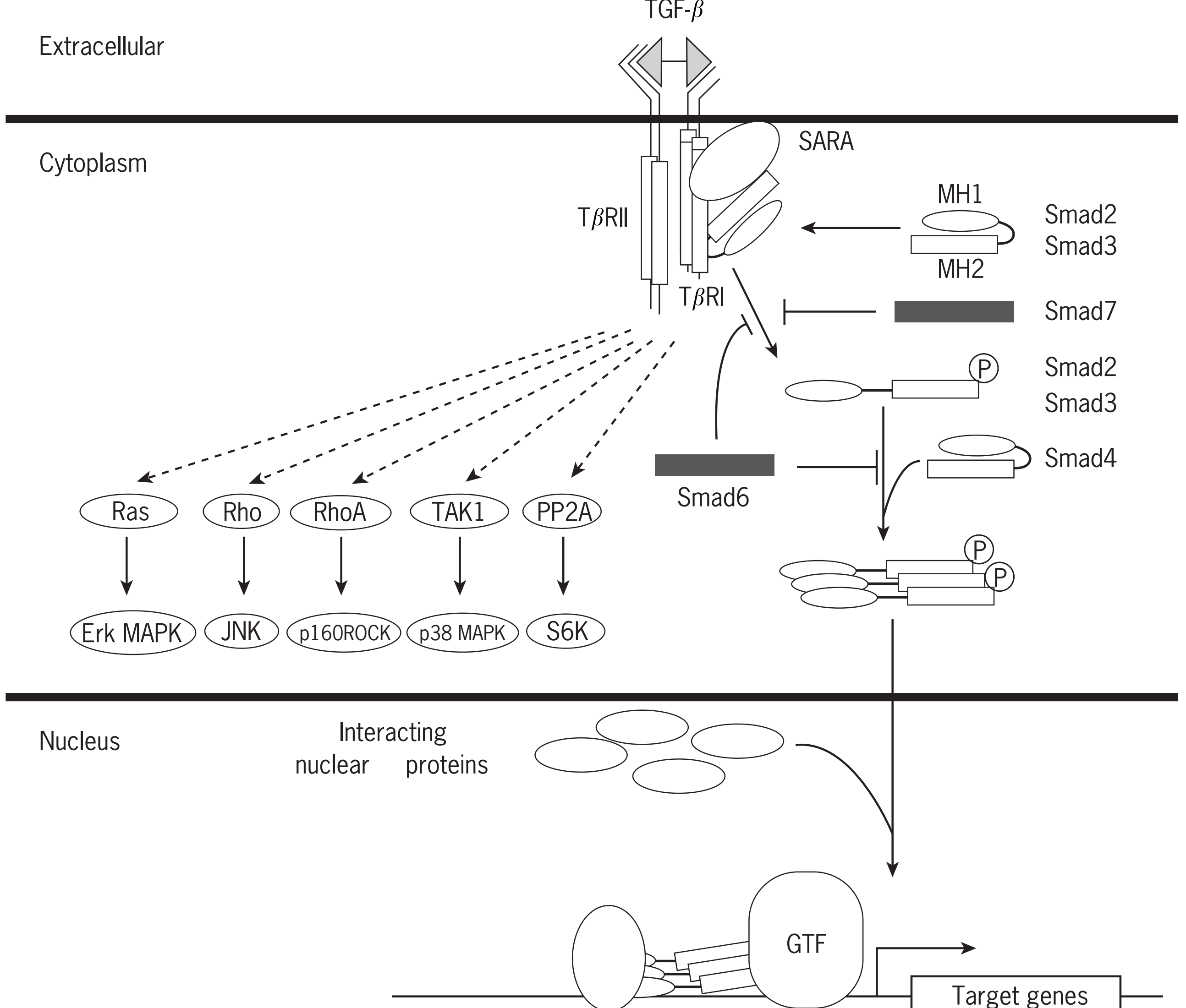


Figure 6 TGF-β-induced signalling through Smads is complemented by several non-Smad signalling mechanisms activated by the TGF-β receptors. Although the exact mechanism of activation remains to be better characterized, TGF-β induces activation of Ras, RhoB and RhoA, as well as TAK1 and the protein phosphatase 2A, thus leading to activation of several MAP kinase pathways, and down-regulation of S6 kinase activity.

Although less common, inactivating TβRI mutations also occur in ovarian cancers, metastatic breast cancers, pancreatic carcinomas and T cell lymphomas. Together, these mutations suggest a function of TβRII and TβRI as tumour suppressors in carcinoma development (de Caestecker *et al*., 2000; Massagué *et al*., 2000).

TGF-β receptor expression is often decreased in carcinomas and this down-regulation may allow cells to escape growth inhibition by TGF-β. Several Ets transcription factors regulate the expression of TβRII and reduced expression of these factors correlates with reduced receptor expression in gastric cancers. Transcriptional silencing may also result from hypermethylation of CpG islands in the *TβRI* or *TβRII* promoters, or mutations in the *TβRII* promoter which interfere with transcription factor binding. Decreased TβRII function has been shown to confer resistance against the growth-inhibitory effect of TGF-β, whereas other TGF-β responses, e.g. extracellular matrix protein expression, may not be similarly affected, and increased TβRII expression correlates with sensitivity to the growth inhibitory response of TGF-β. Thus, different signalling threshold requirements may be the basis for the observation that a decrease in TβRII function may primarily affect the growth responsiveness (de Caestecker *et al*., 2000; Massagué *et al*., 2000).

Expression of wild-type TβRII in colon or breast carcinoma cell lines, which lack a functional *TβRII* allele, provided evidence that the TβRII acts as a tumour

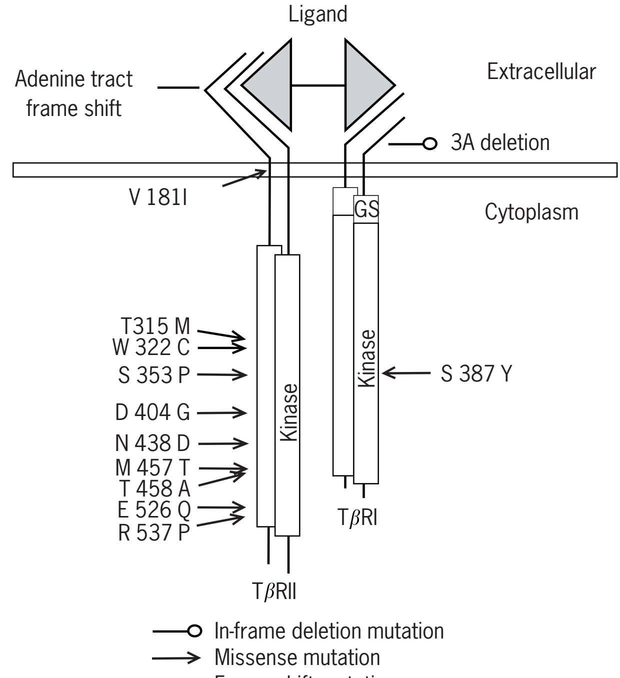


Figure 7 Summary of mutations in TβRII and TβRI found in various cancers.

suppressor (de Caestecker *et al*., 2000; Massagué *et al*., 2000). The TβRII-expressing cells were growth inhibited and had suppressed anchorage independence and strongly reduced tumour formation in nude mice, when compared with the parental cells. Transgenic expression of dominant negative forms of TβRII in the skin or mammary gland increased tumour formation, further supporting a tumour suppressor role of TβRII. In addition, mice with an inactivated *TGF-β1* gene also show increased carcinoma development, either spontaneously (Engle *et al*., 1999) or after carcinogen treatment (Tang *et al*., 1998). In the latter case, haploinsufficiency of the *TGF-β1* gene increased tumour susceptibility, since the tumours retained one wild-type allele. Consistent with these results, decreased TβRII expression correlated with high tumour grade of human breast cancers. While these observations suggest that loss of TGF-β responsiveness provides a distinct advantage for tumour development, most tumours do not have inactivated TGF-β receptors, and HNPCC patients, who frequently have TβRII mutations in their tumours, have a better prognosis than patients with sporadic colon cancer, who do not have TβRII mutations. Therefore, abrogation of TGF-β signalling, while leading to loss of the growth inhibition by TGF-β and early tumour onset, paradoxically protects against tumour progression, since the tumours do not adopt an invasive phenotype in response to autocrine or paracrine TGF-β (see below). This possibility is supported by mouse studies, in which a functional TβRII was expressed in colon cancer cells that lack TGF-β receptor expression. Although the transfected cells showed reduced growth rate, they had a strongly increased invasive and metastatic capacity (Oft *et al*., 1998). Together these findings illustrate the tumour-suppressor role of TGF-β receptors in carcinoma development, and the distinct advantages, provided by TGF-β responsiveness, for cancer progression at later stages.

Mutations in Smads have been found in a variety of carcinomas (**Figure 8**), and, even though generally uncommon, they suggest that some Smads act as tumour suppressors (de Caestecker *et al*., 2000; Massagué *et al*., 2000). While no mutations in Smad3 or the inhibitory Smad 6 or -7 have been reported, inactivation of Smad2 or -4 genes occurs by loss of chromosome segments, small deletions, frameshift, nonsense or missense mutations. Smad4 mutations occur primarily in pancreatic carcinomas, in which Smad4 was originally identified as *DPC4* (deleted in pancreatic carcinoma), and colon carcinomas, and less frequently in other carcinomas. While biallelic inactivation of Smad4 is most commonly observed, haploinsufficiency of the Smad4 locus may also contribute to cancer progression, although this interpretation may be confounded by the presence of the Smad2 and DCC loci in close proximity to the Smad4 gene. In contrast to Smad4, Smad2 mutations are rare and occur primarily in colorectal and lung carcinomas. Finally, enhanced Smad7 levels are observed in pancreatic carcinomas and may also decrease Smad responsiveness.

Tumour-associated mutations in Smad4 and Smad2 occur most frequently in the MH2 domain, which mediates heteromeric complex formation and transcriptional activation. C-terminal deletions or mutations often inactivate the Smad, and provide dominant negative interference with wild-type Smad function and TGF-β-induced nuclear translocation. Many mutations also map at the interfaces of Smad heteromerization, suggesting interference with heteromerization. Other mutations decrease the stability, e.g. through increased ubiquitin-mediated degradation of the Smad proteins. Finally, mutations in the MH1 domain of Smad4 interfere with its DNA binding. While most if not all mutations impair Smad functions, some mutations may alter TGF-β signalling to the tumour's advantage. Indeed, the Smad2 mutations *D450E* and *P445H*, found in colorectal carcinomas, enhance the invasive behaviour of the tumour cells, when overexpressed (de Caestecker *et al*., 2000; Massagué *et al*., 2000).

While homologous inactivation of the Smad2 or -4 genes results in embryonic lethality and therefore prevents an assessment in tumour progression, several mouse studies suggest a role of Smad4 as tumour suppressors. Mice heterozygous for an inactivated Smad4 gene develop intestinal polyps, which can progress into carcinomas. When combined with an inactivated allele of the adenomatous polyposis coli (*APC*) gene, this double heterozygosity allows the development of highly invasive colon carcinomas with both the *APC* and Smad4 genes

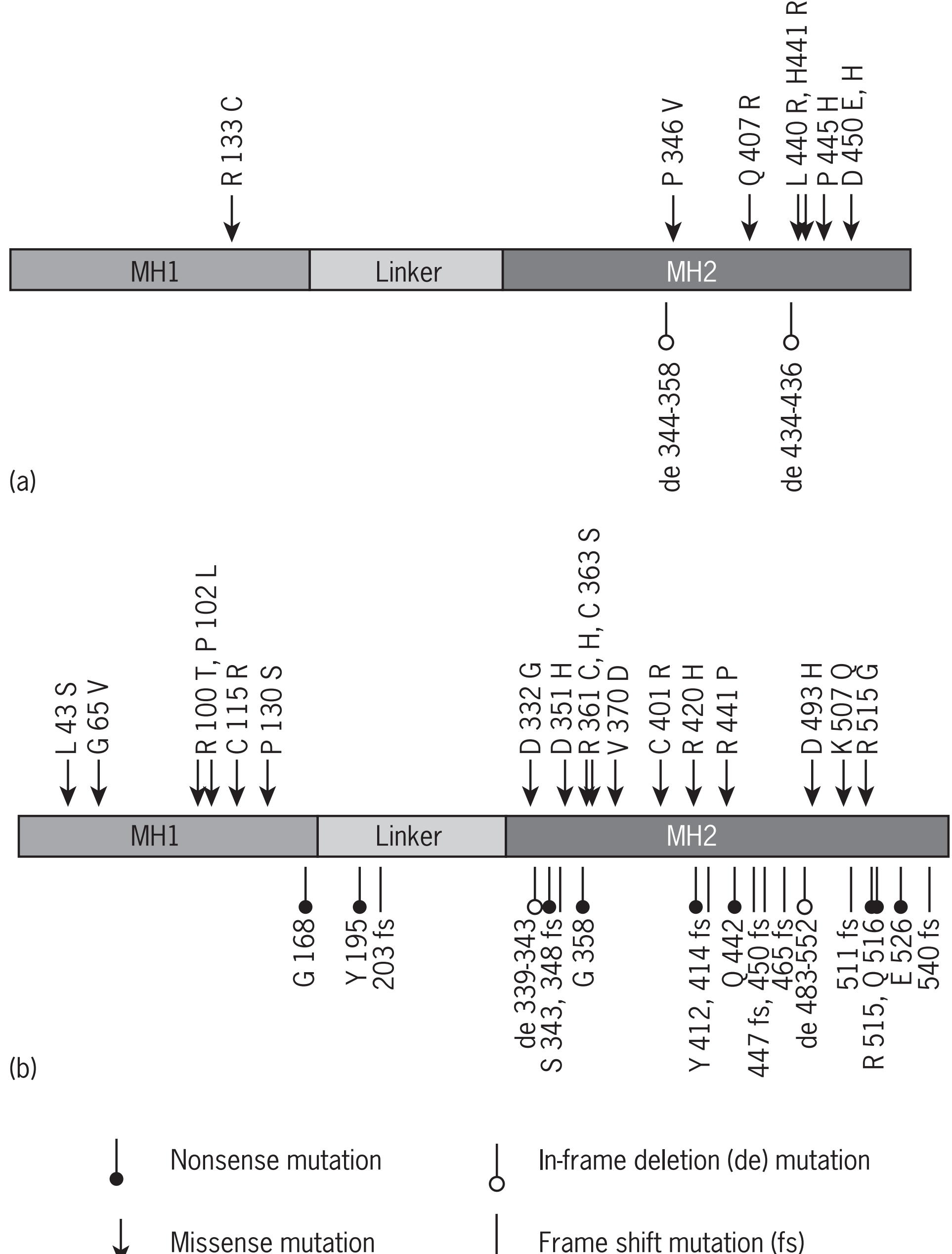


Figure 8 Summary of mutations in (a) Smad2 and (b) Smad4 found in various carcinomas.

inactivated (Takaku *et al*., 1998). Mice with homozygous inactivation of the Smad3 gene have also been shown to develop colon carcinomas (Zhu *et al*., 1998), although this phenotype was not seen in two other similar studies. While a role of Smad3 as tumour-suppressor gene is conceivable, no inactivating Smad3 mutations have been observed in human tumours. This could be rationalized by the observation that Smad3 activates most gene expression responses to TGF-β, whereas Smad2 cooperates with Smad3 for TGF-β-induced expression of the p21^{Cip1} or p15^{Ink4B} cdk inhibitors (Feng *et al*., 2000; Pardali *et al*., 2000). Smad2 inactivation may therefore inactivate the growth-arrest response, without affecting Smad3 responses, e.g. induction of extracellular matrix proteins, which provide an advantage for tumour development. Importantly, while Smad4 is required for many if not most TGF-β gene expression responses, Smad4-deficient cells may have retained a variety of responses (Sirard *et al*., 2000). Thus, inactivation of Smad4 or Smad2 function may not abolish, but rather perturb, the complex TGF-β response, and thereby primarily targets the growth inhibitory response to TGF-β. Moreover, different TGF-β responses may have differential sensitivity to Smad signalling. Accordingly, Smad7 overexpression in a colon

carcinoma cell line suppressed TGF-β-induced growth arrest, without affecting TGF-β-induced expression of plasminogen activator inhibitor 1, and enhanced anchorage-independent growth and tumorigenicity. Finally, inactivating Smad4 mutations have been found in conjunction with mutations in TβRII or TβRI, strongly suggesting that Smad4 has tumour-suppressing activities which are unrelated to TGF-β signalling.

Oncogene expression or increased proto-oncogene expression may also down-regulate the TGF-β responsiveness and allow for escape of TGF-β's tumour-suppressor effects. For example, *Evi-1*, an oncogene involved in leukaemic transformation of haematopoietic cells, inhibits Smad3 function and thus decreases or abolishes TGF-β responsiveness. The oncogenic effect of *Evi-1* may therefore result in part from interference with TGF-β signalling. Conceptually similarly, c-Ski represses Smad-mediated transcription, thus raising the possibility that its interference with TGF-β responsiveness may contribute to its oncogenic activity. Accordingly, c-Ski expression is often elevated in melanomas and carcinomas, and this has been correlated with decreased TGF-β responsiveness. Finally, c-Myc also inhibits the antiproliferative response to TGF-β. TGF-β downregulates c-Myc expression in epithelial cells, but this repression is lost in various cancer cell lines, concomitant with the loss of the growth inhibitory response to TGF-β (Massagué *et al.*, 2000; Chen *et al.*, 2001). Thus, downregulation of c-Myc by TGF-β may be required for the growth-inhibitory response to TGF-β, and the absence of this downregulation may confer resistance to the tumour-suppressor activity of TGF-β. Clearly, carcinoma cells have developed several strategies to escape the growth-inhibitory response of TGF-β.

CELL-AUTONOMOUS, STIMULATORY EFFECTS OF TGF-β ON TUMOUR DEVELOPMENT

Even though TGF-β signalling exerts tumour-suppressor activity, TGF-β1 expression is often increased in tumour cells and stimulates tumour development and cancer progression. Indeed, autocrine TGF-β signalling can induce morphological changes and invasive behaviour of the tumour cells, while increased TGF-β1 production and activation make the microenvironment more permissive for tumour progression (Akhurst and Balmain, 1999; Dumont and Arteaga, 2000).

Tumour metastasis depends on various factors, including the ability of tumour cells to migrate and invade the stroma and to migrate in and out of the blood vessels. Epithelial to mesenchymal differentiation of tumour cells plays an important role in this invasive phenotype. Fibroblastoid or 'spindle cell' tumours of epithelial origin have been characterized as highly malignant and invasive (Cui *et al.*, 1996). TGF-β and TGF-β-related proteins have been implicated in epithelial to mesenchymal transdifferentiation both in normal development and in tumour progression. This phenotypic transition is characterized by extensive changes in expression of cell adhesion molecules and a switch from a cytoskeleton of mainly cytokeratin intermediate filaments to one predominantly composed of vimentin. Epithelial to mesenchymal differentiation in culture is thought to correlate with these cell changes that facilitate invasion and metastasis *in vivo*.

TGF-β has been shown to induce epithelial to mesenchymal transition in culture of normal and transformed breast epithelial cells, squamous carcinoma, ovarian adenosarcoma and melanoma cells. The phenotypic changes have been best characterized in NMuMG cells, in which TGF-β induces cell shape changes, down-regulates expression of E-cadherin, ZO-1, vinculin and keratin and induces expression of vimentin and N-cadherin, which has been shown to increase cell motility and scattering (Miettinen *et al.*, 1994; Bhowmick *et al.*, 2001). Although the mechanism of TGF-β-induced epithelial to mesenchymal differentiation is presumably complex, PI-3-kinase signalling and TGF-β-induced Smad and RhoA activation appear to play a role (Bhowmick *et al.*, 2001). In addition, the morphological changes occur rapidly in response to TGF-β, and are reversible. *In vivo*, the effect of TGF-β on the spindle cell phenotype and invasive behaviour of carcinomas has been well documented for skin carcinomas in mice (Cui *et al.*, 1996). Consistent with a tumour-suppressor role, increased expression of activated TGF-β1 reduced the hyperplastic response to the tumour promoter PMA and the number of papillomas after treatment with DMBA. On the other hand, increased TGF-β1 expression enhanced the malignant conversion of skin carcinomas in a carcinogenesis mouse model, and increased the incidence of spindle cell carcinomas (Cui *et al.*, 1996). In another study, expression of a dominant negative TβRII prevented squamous carcinoma cells from undergoing mesenchymal differentiation in response to TGF-β both *in vitro* and *in vivo*. Consequently, the tumours had a differentiated epithelial phenotype and were less malignant and invasive than the parental cells, which developed fibroblastoid spindle cell carcinomas (Portella *et al.*, 1998). A similar role of TGF-β signalling was also apparent in Ras-transformed mammary epithelial cells and a fibroblastoid colon carcinoma cell line (Oft *et al.*, 1998). Expression of a dominant negative TβRII prevented TGF-β-induced epithelial to mesenchymal changes and reverted the colon carcinoma cells to an epithelial phenotype. In addition, blocking TβRII function suppressed invasion *in vitro* and the metastatic phenotype of this colon carcinoma cell line. Finally, restoration of TβRII signalling in HNPCC cells with a mutated TβRII rendered the cells invasive, in contrast to their normally, noninvasive phenotype.

TUMOUR PROGRESSION STIMULATED BY EFFECTS OF TGF-β ON THE TUMOUR ENVIRONMENT

The increased expression and activation of TGF-β1 by tumour cells also makes the microenvironment more conducive to tumour development. Increased TGF-β1 expression by tumour cells, presumably a result of activated Ras/MAP kinase signalling and signal amplification in response to TGF-β1 itself, enhances the TGF-β levels in the tumour microenvironment. The increased protease expression and plasmin generation by tumour cells and the TGF-β1-induced expression of collagenases and other proteases, such as the matrix metalloproteases MMP-2 and MMP-9, result most likely in TGF-β activation and degradation of the extracellular matrix with consequent release of stored TGF-β. Increased TGF-β1 production stimulates synthesis of extracellular matrix proteins and chemoattraction of fibroblasts. All these changes together result in a microenvironment that is conducive for tumour growth and invasion, and for angiogenesis (Akhurst and Balmain, 1999; Dumont and Arteaga, 2000).

Tumour angiogenesis is critical for tumour growth and invasion, since blood vessels are required to deliver nutrients and oxygen to the tumour cells and allow tumour cells to intravasate the blood system, leading to metastatic spread. TGF-β1 acts as a potent inducer of angiogenesis in several assays (Roberts *et al*., 1986), while mouse models defective in TGF-β signalling illustrate the important role of TGF-β1 in normal vascular development. For example, targetted inactivation of the *TGF-β1* or *TβRII* genes results in embryonic lethality due to defective vasculogenesis and angiogenesis (Dickson *et al*., 1995; Oshima *et al*., 1996), while angiogenesis-defective phenotypes are also apparent in mice with null mutations of the genes for *Alk-1* (Oh *et al*., 2000), a TGF-β type I receptor that is expressed in endothelial cells, or endoglin, a TGF-β type III receptor expressed by endothelial cells (Arthur *et al*., 2000). (See chapter on *Angiogenesis*.)

Several tumour models illustrate the importance of tumour cell-secreted TGF-β1 in tumour angiogenesis. Increased TGF-β1 secretion by transfected pancreas carcinoma (Stearns *et al*., 1999) or CHO (Ueki *et al*., 1992) cells enhanced tumour angiogenesis in immunodeficient mice, whereas local administration of neutralizing antibodies to TGF-β1 strongly reduced tumour angiogenesis (Ueki *et al*., 1992). TGF-β antibodies also reduced angiogenesis and tumorigenicity of a renal carcinoma cell line in T, NK and B cell-deficient mice. These cells did not have cell-autonomous responses to TGF-β since they lacked TβRII (Ananth *et al*., 1999). In humans, histological studies of breast tumours correlate high levels of TGF-β1 mRNA with high microvessel density, and each of these factors correlated with poor patient prognosis. Diagnostic studies on other carcinoma types correlate high tumour burden and circulating plasma levels of TGF-β1, and enhanced tumour angiogenesis and poor patient prognosis. In one study, TGF-β1 levels were also correlated with expression levels of the angiogenic growth factor VEGF (Saito *et al*., 1999).

The mechanisms whereby TGF-β1 stimulates angiogenesis remain to be further characterized, but presumably combine direct and indirect effects. TGF-β induces expression of VEGF, a potent angiogenic growth factor, which directly stimulates endothelial cell proliferation and migration. TGF-β also induces capillary formation of endothelial cells, cultured on collagen matrix (Choi and Ballermann, 1995). Indirect stimulation of angiogenesis may also occur, since TGF-β is a potent chemoattractant for monocytes, which release angiogenic cytokines, and the TGF-β1-induced changes in the microenvironment stimulate endothelial cell migration and capillary formation. Moreover, the TGF-β-induced expression of MMP-2 and MMP-9, and down-regulation of the protease inhibitor TIMP in both tumour cells and endothelial cells, are expected to enhance the migratory and invasive properties of endothelial cells required for angiogenesis. Thus, both direct effects of TGF-β and effects on the microenvironment stimulate tumour angiogenesis.

Local immunosuppression in response to increased TGF-β1 levels allows tumour cells to escape from immunosurveillance and thus stimulates tumour development and progression (Dumont and Arteaga, 2000). TGF-β1 is the most potent immunosuppressive cytokine known to date, and inhibits proliferation and functional differentiation of T lymphocytes, lymphokine-activated killer cells, natural killer cells, neutrophils, macrophages and B cells (Derynck and Choy, 1998; Letterio and Roberts, 1998; de Visser and Kast, 1999). Several findings illustrate the role of TGF-β-induced, local immunosuppression in tumorigenicity. For example, increased TGF-β1 expression in a tumour cell line prevented cytotoxic T lymphocyte activation and enhanced tumorigenicity, in contrast to parental cells. Such repression can be elaborated by the ability of TGF-β to inhibit the expression and function of interleukin 2 and interleukin 2 receptors (Letterio and Roberts, 1998; de Visser and Kast, 1999). Accordingly, mammary tumour cells, which produce high TGF-β1 levels, inhibited the cytotoxic T lymphocyte response, but this inhibition was overcome by expressing interleukin 2 in these cells. These and other results suggest that TGF-β-mediated suppression of the cytotoxic T cell response promotes tumour development.

TGF-β1 also inhibits other immune functions of relevance to tumour development. Increased TGF-β1 expression by tumour cells decreased natural killer cell activity and promoted tumour formation in nude mice that lack T cells. In addition, anti-TGF-β antibodies suppressed tumour formation and metastasis of a breast carcinoma cell line in nude mice, while enhancing natural killer cell function. This suppression was not seen in beige mice,

which lack natural killer cells, thus implicating TGF-β1-induced suppression of natural killer cells in cancer progression (Arteaga *et al*., 1993). TGF-β-mediated suppression of neutrophil function may also be involved in tumour progression. Indeed, Fas-ligand expressing carcinoma cells underwent neutrophil-mediated rejection, but this rejection did not occur at a site with high TGF-β levels or when TGF-β1 was injected at the tumour site (Chen *et al*., 1998). Finally, TGF-β down-regulates the expression of the major histocompatibility complex (MHC) class II antigens (Letterio *et al*., 1996), suggesting that TGF-β1 expression renders the tumour cells less immunogenic. This regulation contributes to the local immunosuppression and to the escape from immune surveillance.

In summary, the mechanisms through which TGF-β signals and exerts its multiple responses are rapidly being characterized. Smads exert multiple gene expression responses, but cross-talk with other signalling pathways and non-Smad TGF-β signalling provide further complexity to the TGF-β response. TGF-β signalling plays an important role in tumour cell behaviour and cancer progression. The growth-inhibitory activity of TGF-β acts as a tumour suppressor in carcinoma development, whose function is often eliminated through mutations in receptors or Smads. In contrast, the increased TGF-β1 expression by tumour cells stimulates tumour development, both in a cell-autonomous manner and through effects on its environment, e.g. stimulation of angiogenesis and localized immunosuppression.

REFERENCES

Akhurst, R. J. and Balmain, A. (1999). Genetic events and the role of TGF-β in epithelial tumour progression. *Journal of Pathology*, **187**, 82–90.

Ananth, S., *et al*. (1999). Transforming growth factor β1 is a target for the von Hippel–Lindau tumour suppressor and a critical growth factor for clear cell renal carcinoma. *Cancer Research*, **59**, 2210–2216.

Arteaga, C. L., *et al*. (1993). Anti-transforming growth factor (TGF)-β antibodies inhibit breast cancer cell tumorigenicity and increase mouse spleen natural killer cell activity. Implications for a possible role of tumor cell/host TGF-β interactions in human breast cancer progression. *Journal of Clinical Investigation*, **92**, 2569–2576.

Arthur, H. M., *et al*. (2000). Endoglin, an ancillary TGFβ receptor, is required for extraembryonic angiogenesis and plays a key role in heart development. *Developmental Biology*, **217**, 42–53.

Barcellos-Hoff, M.-H., *et al*. (1994). Transforming growth factor-β activation in irradiated murine mammary gland. *Journal of Clinical Investigation*, **93**, 892–899.

Bhowmick, N. A., *et al*. (2001). Transforming growth factor-β1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. *Molecular Biology of the Cell*, **12**, 27–36.

Chen, Y. G., *et al*. (1997). Mechanism of TGF-β receptor inhibition by FKBP12. *EMBO Journal*, **16**, 3866–3876.

Chen, J. J., *et al*. (1998). Regulation of the proinflammatory effects of Fas ligand (CD95L). *Science*, **282**, 1714–1717.

Chen, C. R., *et al*. (2001). Defective repression of c-myc in breast cancer cells: A loss at the core of the transforming growth factor β growth arrest program. *Proceedings of the National Academy of Sciences of the USA*, **98**, 992–999.

Choi, M. E. and Ballermann, B. J. (1995). Inhibition of capillary morphogenesis and associated apoptosis by dominant negative mutant transforming growth factor-β receptors. *Journal of Biological Chemistry*, **270**, 21144–21150.

Choy, L. and Derynck, R. (1998). The type II transforming growth factor-β (TGF-β) receptor-associated protein TRIP-1 acts as a modulator of the TGF-β response. *Journal of Biological Chemistry*, **273**, 31455–31462.

Cui, W., *et al*. (1996). TGFβ1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. *Cell*, **86**, 531–542.

Datta, P. K. and Moses, H. L. (2000). STRAP and Smad7 synergize in the inhibition of transforming growth factor β signaling. *Molecular and Cell Biology*, **20**, 3157–3167.

de Caestecker, M. P., *et al*. (2000). Role of transforming growth factor-β signaling in cancer. *Journal of the National Cancer Institute*, **92**, 1388–1402.

Derynck, R. and Choy, L. (1998). Transforming growth factor-β and its receptors. In Thompson, A. (ed.), *The Cytokine Handbook*, 3rd edn 593–636 (Academic Press, Boston).

Derynck, R. and Feng, X.-H. (1997) TGF-β receptor signaling. *BBA Reviews in Cancer*, **1333**, F105–F150.

Derynck, R., *et al*. (1998). Smads: transcriptional activators of TGF-β responses. *Cell*, **95**, 737–740.

de Visser, K. E. and Kast, W. M. (1999). Effects of TGF-β on the immune system: implications for cancer immunotherapy. *Leukemia*, **13**, 1188–1199.

Dickson, M. C., *et al*. (1995). Defective haematopoiesis and vasculogenesis in transforming growth factor-β1 knock out. *Development*, **121**, 1845–1854.

Dumont, N. and Arteaga, C. L. (2000). Transforming growth factor-β and breast cancer: tumor promoting effects of transforming growth factor-β. *Breast Cancer Research*, **2**, 125–132.

Engle, S. J., *et al*. (1999). Transforming growth factor β1 suppresses nonmetastatic colon cancer at an early stage of tumorigenesis. *Cancer Research*, **59**, 3379–3386.

Ewen, M. E., *et al*. (1993). TGF β inhibition of Cdk4 synthesis is linked to cell cycle arrest. *Cell*, **74**, 1009–1020.

Feng, X.-H., *et al*. (2000). Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15Ink4B expression in response to TGF-β. *EMBO Journal*, **19**, 5178–5193.

Griswold-Prenner, I., *et al*. (1998). Physical and functional interactions between type I transforming growth factor-β receptors and Bα, a WD-40 repeat subunit of phosphatase 2A. *Molecular and Cell Biology*, **18**, 6595–6604.

Iavarone, A. and Massagué, J. (1997). Repression of the CDK activator Cdc25A and cell-cycle arrest by cytokine TGF-β in cells lacking the CDK inhibitor p15. *Nature*, **387**, 417–422.

Itoh, S., *et al*. (2000). Signaling of transforming growth factor-β family members through Smad proteins. *European Journal of Biochemistry*, **267**, 6954–6967.

Letterio, J. J. and Roberts, A. B. (1998). Regulation of immune responses by TGF-β. *Annual Review of Biochemistry*, **16**, 137–161.

Letterio, J. J., *et al*. (1996). Autoimmunity associated with TGF-β1-deficiency in mice is dependent on MHC class II antigen expression. *Journal of Clinical Investigation*, **98**, 2109–2119.

Massagué, J. (2000). How cells read TGF-β signals. *Nature Reviews of Molecular and Cellular Biology*, **1**, 169–178.

Massagué, J. and Wotton, D. (2000). Transcriptional control by the TGF-β/Smad signaling system. *EMBO Journal*, **19**, 1745–1754.

Massagué, J., *et al*. (2000). TGF-β signaling in growth control, cancer, and heritable disorders. *Cell*, **103**, 295–309.

Miettinen, P. J., *et al*. (1994). TGF-β induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *Journal of Cell Biology*, **127**, 2021–2036.

Munger, J. S., *et al*. (1997). Latent transforming growth factor-β: structural features and mechanisms of activation. *Kidney International*, **51**, 1376–1382.

Munger, J. S., *et al*. (1999). The integrin αv β6 binds and activates latent TGF β1: a mechanism for regulating pulmonary inflammation and fibrosis. *Cell*, **96**, 319–328.

Oft, M., *et al*. (1998). TGF-β signaling is necessary for carcinoma cell invasiveness and metastasis. *Current Biology*, **19**, 1243–1252.

Oh, S. P., *et al*. (2000). Activin receptor-like kinase 1 modulates transforming growth factor-β1 signaling in the regulation of angiogenesis. *Proceedings of the National Academy of Sciences of the USA*, **97**, 2626–2631.

Oshima, M., *et al*. (1996). TGF-β receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. *Developmental Biology*, **179**, 297–302.

Pardali, K., *et al*. (2000). Role of Smad proteins and transcription factor Sp1 in p21(Waf1/Cip1) regulation by transforming growth factor-β. *Journal of Biological Chemistry*, **275**, 29244–29256.

Petritsch, C., *et al*. (2000). TGF-β inhibits p70 S6 kinase via protein phosphatase 2A to induce G1 arrest. *Genes and Development*, **14**, 3093–3101.

Piek, E., *et al*. (1999). Specificity, diversity, and regulation in TGF-β superfamily signaling. *FASEB Journal*, **13**, 2105–2124.

Portella, G., *et al*. (1998). Transforming growth factor β is essential for spindle cell conversion of mouse skin carcinoma *in vivo*: implications for tumor invasion. *Cell Growth and Differentiation*, **9**, 393–404.

Ribeiro, S. M., *et al*. (1999). The activation sequence of thrombospondin-1 interacts with the latency-associated peptide to regulate activation of latent transforming growth factor-β. *Journal of Biological Chemistry*, **274**, 13586–13593.

Roberts, A. B., *et al*. (1986). Transforming growth factor type β: rapid induction of fibrosis and angiogenesis *in vivo* and stimulation of collagen formation *in vitro*. *Proceedings of the National Academy of Sciences of the USA*, **83**, 4167–4171.

Saito, H., *et al*. (1999). The expression of transforming growth factor-β1 is significantly correlated with the expression of vascular endothelial growth factor and poor prognosis of patients with advanced gastric carcinoma. *Cancer*, **86**, 1455–1462.

Sherr, C. J. and Roberts, A. B. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes and Development*, **13**, 1501–1512.

Sirard, C., *et al*. (2000). Targeted disruption in murine cells reveals variable requirement for Smad4 in transforming growth factor β-related signaling. *Journal of Biological Chemistry*, **275**, 2063–2070.

Stearns, M. E., *et al*. (1999). Role of interleukin 10 and transforming growth factor β1 in the angiogenesis and metastasis of human prostate primary tumor lines from orthotopic implants in severe combined immunodeficiency mice. *Clinical Cancer Research*, **5**, 711–720.

Takaku, K., *et al*. (1998). Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. *Cell*, **92**, 645–656.

Tang, B., *et al*. (1998). Transforming growth factor-β1 is a new form of tumor suppressor with true haploid insufficiency. *Nature Medicine*, **4**, 802–807.

ten Dijke, P., *et al*. (2000). Signaling inputs converge on nuclear effectors in TGF-β signaling. *Trends in Biochemical Science*, **25**, 64–70.

Ueki, N., *et al*. (1992). Excessive production of transforming growth-factor β1 can play an important role in the development of tumorigenesis by its action for angiogenesis: validity of neutralizing antibodies to block tumor growth. *Biochimica Biophysica Acta*, **1137**, 189–196.

Yu, Q. and Stamenkovic, I. (2000). Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-β and promotes tumor invasion and angiogenesis. *Genes and Development*, **14**, 163–176.

Zhang, Y. and Derynck, R. (1999). Regulation of Smad signaling by protein associations and signaling crosstalk. *Trends in Cell Biology*, **9**, 274–279.

Zhu, Y., *et al*. (1998). Smad3 mutant mice develop metastatic colorectal cancer. *Cell*, **94**, 703–714.

FURTHER READING

Chang, H., *et al*. (1999). Smad5 knockout mice die at mid-gestation due to multiple embryonic and extraembryonic defects. *Development*, **126**, 1631–1642.

Daopin, S., *et al*. (1992). Crystal structure of transforming growth factor-β2: an unusual fold for the superfamily. *Science*, **257**, 369–373.

Derynck, R., *et al*. (1988). A new type of transforming growth factor-β, TGF-β3. *EMBO Journal*, **7**, 3737–3743.

Hata, A., *et al*. (1998). TGF-β signaling and cancer: structural and functional consequences of mutations in Smads. *Molecular Medicine Today*, **4**, 257–262.

Kim, S. J., *et al*. (2000). Molecular mechanisms of inactivation of TGF-β receptors during carcinogenesis. *Cytokine Growth Factor Review*, **11**, 159–168.

Markowitz, S. D. and Roberts, A. B. (1996). Tumor suppressor activity of the TGF-β pathway in human cancers. *Cytokine Growth Factor Review*, **7**, 93–102.

Massagué, J. (1998) TGF-β signal transduction. *Annual Review of Biochemistry*, **67**, 753–791.

McAdam, A. J., *et al*. (1994). Transfection of transforming growth factor-β producing tumor EMT6 with interleukin-2 elicits tumor rejection and tumor reactive cytotoxic T-lymphocytes. *Journal of Immunotherapy with Emphasis on Tumor Immunology*, **15**, 155–164.

Miyazono, K., *et al*. (2000). TGF-β signaling by Smad proteins. *Advances in Immunology*, **75**, 115–157.

Torre-Amione, G., *et al*. (1990). A highly immunogenic tumor transfected with a murine transforming growth factor type β1cDNA escapes immune surveillance. *Proceedings of the National Academy of Sciences of the USA*, **87**, 1486–1490.

Wotton, D. and Massagué, J. (2001). Smad transcriptional corepressors in TGF β family signaling. *Current Topics in Microbiology and Immunology*, **254**, 145–164.

Chapter 14

Wnt Signal Transduction

Jeffrey R. Miller
University of Minnesota, Minneapolis, MN, USA

CONTENTS

- Introduction
- The Wnt Signal Transduction Pathway
- Additional Regulators of Wnt Signal Transduction
- Oncogenic Activation of the Wnt Pathway
- Conclusion

INTRODUCTION

Whether it is a meeting with a colleague, a conversation over coffee with a friend or a telephone call to a loved one, interpersonal communication is an essential part of our daily lives. In our bodies, cells also communicate with one another to coordinate their behaviour and determine their specialized role in the body. Cells utilize elaborate systems of proteins to transmit, receive and respond to signals from neighbouring cells. These systems, or signal transduction pathways, utilize secreted signalling proteins, cell surface and intracellular receptor proteins, protein kinases, transcription factors and other intracellular proteins. Intercellular signals control a variety of processes in the body and ensure, for example, that each cell divides only when its neighbours dictate that it should do so. The importance of such signals becomes apparent when this communication breaks down and cells begin to divide uncontrollably, resulting in cancer. For example, colon cells normally divide at a rate that balances the loss of colon cells due to attrition. Colon cells receive signals that tell them to divide when more cells are needed and are told to stop dividing when the appropriate number of cells is reached. In colon cancer, like many other cancers, defects in this regulatory network cause colon cells to divide continuously leading to tumour formation.

The past 25 years of cancer research have revealed that cancer is a complex disease involving dynamic changes in the genome. This was realized through the discovery that mutations in specific genes, called oncogenes and tumour-suppressor genes, played critical roles in tumour formation. Mutations in these genes result in defects in regulatory pathways that control normal cell proliferation and homeostasis. Dominant gain-of-function mutations produce oncogenes, a gene that is locked in the 'ON' position and leads to hyperactivation of a regulatory pathway. Conversely, recessive mutations inactivate tumour-suppressor genes that normally keep a regulatory pathway 'OFF'. Coming back to the example of colon cancer, both types of mutations have been uncovered with the end result being the uncontrolled proliferation of colon cells. Thus, understanding the function of oncogenes and tumour-suppressors is vital to gain insights into the molecular causes of cancer and for designing therapeutic agents to treat cancer.

Through modern cloning techniques, a number of oncogenes and tumour suppressors have been identified. Characterization of these genes led to the realization that many oncogenes and tumour suppressors encode components of evolutionarily conserved signal transduction pathways important for controlling embryonic development. The Wnt signal transduction pathway provides one of the most striking examples of this connection. Inappropriate activation of the Wnt signal transduction pathway is implicated in a variety of human cancers, most notably colon cancer. Many cases of colon cancers are associated with either oncogenic mutations in β-catenin, a positive regulator of Wnt signal transduction, or inactivating mutations in APC (the tumour-suppressor protein encoded by the adenomatous polyposis coli gene), a negative regulator of Wnt signalling. In both cases, these mutations lead to the aberrant activation of the Wnt pathway and tumorigenesis. During development, Wnt signalling plays critical roles in controlling a variety of processes including cell fate determination and cell proliferation. For example, Wnt signalling is required for the patterning of the central nervous system and the establishment of the dorsal–ventral axis in frogs. The involvement of Wnt signalling in both embryonic development and cancer has fuelled an extraordinary explosion of interest in understanding the underlying molecular mechanism of Wnt signal transduction. This chapter summarizes the current model for Wnt signal transduction and then discusses how inappropriate activation of Wnt signalling causes cancer. Owing to constraints on space and an effort to present a simplified and coherent picture

of Wnt signalling, certain aspects of Wnt signalling will not be covered in this review. For more information, readers should refer to the list of further reading located at the end of this chapter. Additional information on Wnt signalling can also be found on the Wnt gene homepage (http://www.stanford.edu/~rnusse/wntwindow.html) and at the connections map at *Science's* STKE Web site (http://www.stke.org).

THE Wnt SIGNAL TRANSDUCTION PATHWAY

The Wnt signal transduction pathway is one of the major developmentally important signalling pathways with well-characterized roles in a variety of organisms including mice, frogs, fish, flies and worms. Like other intercellular signal transduction pathways, the Wnt pathway utilizes a secreted signalling protein a transmembrane receptor protein and complex intracellular machinery to relay signals from the cell surface to the nucleus. The major components of the Wnt signalling pathway include the Wnt family of secreted proteins, the Frizzled family of transmembrane receptor proteins and the intracellular proteins Casein Kinase Iε (CKIε), Dishevelled, GBP/Frat, Glycogen Synthase Kinase 3 (GSK3), APC, Axin, β-catenin (Armadillo) and the TCF/LEF family of transcriptional regulators. **Figure 1** presents a schematic representation of each of these proteins showing their important structural and functional domains.

The first insights into the mechanism of Wnt signal transduction came from pioneering studies in the fruit fly, *Drosophila melanogaster*. Researchers used the awesome power of genetics to characterize several fly genes with mutant phenotypes similar to that seen in embryos lacking *wingless*, the fly ortholog of vertebrate *Wnt-1*. These genes were then ordered into a genetic pathway, which has served as a paradigm for understanding Wnt signal transduction in other model systems. In recent years, additional components of the pathway have been identified through a variety of methods, including protein–protein interaction screens. Now, researchers are working to put the pieces of the Wnt signalling puzzle together, a task that is proving to be challenging. However, recent studies examining the complex biochemical relationships between components of the Wnt pathway have provided exciting new insights into the mechanism of Wnt signalling and provides the working model for Wnt signal transduction shown in **Figure 2** and described below.

Wnt signalling is dependent on the presence or absence of the intracellular protein β-catenin. In the absence of Wnt signal, the destruction complex, a multiprotein machine composed of Axin, APC and GSK3, tags β-catenin for destruction by the addition of phosphate groups to serine and threonine residues near the N-terminus of β-catenin. These phosphoamino acids then act as a binding site for a second complex of proteins, the ubiquitin ligase complex, which covalently adds a small protein called ubiquitin to β-catenin. Proteins tagged with ubiquitin are targetted to the proteasome, the cell's protein incinerator, where they are rapidly destroyed. This process serves to keep the levels of β-catenin in the cell low and the Wnt pathway is OFF. When cells perceive a Wnt signal, a group of proteins including Dishevelled, CKIε and GBP/Frat are activated and together they inactivate the destruction complex. As a result, β-catenin is no longer ubiquitinated and is protected from degradation by the proteasome. These events lead to the accumulation of β-catenin in the cell. As the level of β-catenin rises, it enters the nucleus and interacts with a DNA-binding protein of the TCF/LEF family. Together, β-catenin and TCF/LEF activate expression of specific cassettes of target genes. During development these targets include genes that direct cells to adopt specific cell fates, whereas in human colon cancer cells these targets include genes that control cell growth and proliferation. With this brief introduction in hand, the role that each of these proteins plays in regulating Wnt signalling will now be examined in greater detail.

The Messengers – *Wnt* Genes

W*nt* genes were first identified independently by researchers in two different fields. *Wnt-1* (first called *int-1*) was identified as a preferred integration site for mouse mammary tumour virus (Nusse and Varmus, 1982). Insertion of the mouse mammary tumour virus in regions surrounding the *Wnt-1* gene led to its inappropriate activation and breast cancer. *Wingless (wg)*, the fly counterpart of *Wnt-1*, was identified in the Nobel Prize-winning screen of Nüsslein-Volhard and Weischaus (1980) as a mutation that resulted in segment polarity defects. Cloning of *Wnt-1* and *wg* showed that these genes shared a high degree of sequence identity. This finding brought together researchers in the fields of cancer biology and developmental biology and greatly accelerated our understanding of Wnt signalling. Many additional members of the *Wnt* gene family have since been cloned in many organisms from nematode worms to humans, with each organism possessing multiple related *Wnt* genes. For example, the fruit fly possesses seven and the mouse at least 18 *Wnt* genes. A list of known *Wnt* genes and sequence comparisons can be found on the *Wnt* gene homepage (http://www.stanford.edu/~rnusse/wntwindow.html).

Wnt genes encode secreted glycoproteins typically 350–400 amino acids in length and are characterized by a conserved pattern of 23–24 cysteine residues (**Figure 1**). Wnt proteins appear to act as short-range messengers, acting within several cell diameters of the producing cell. Although Wnt proteins are secreted and can be found associated with the cell surface and extracellular matrix,

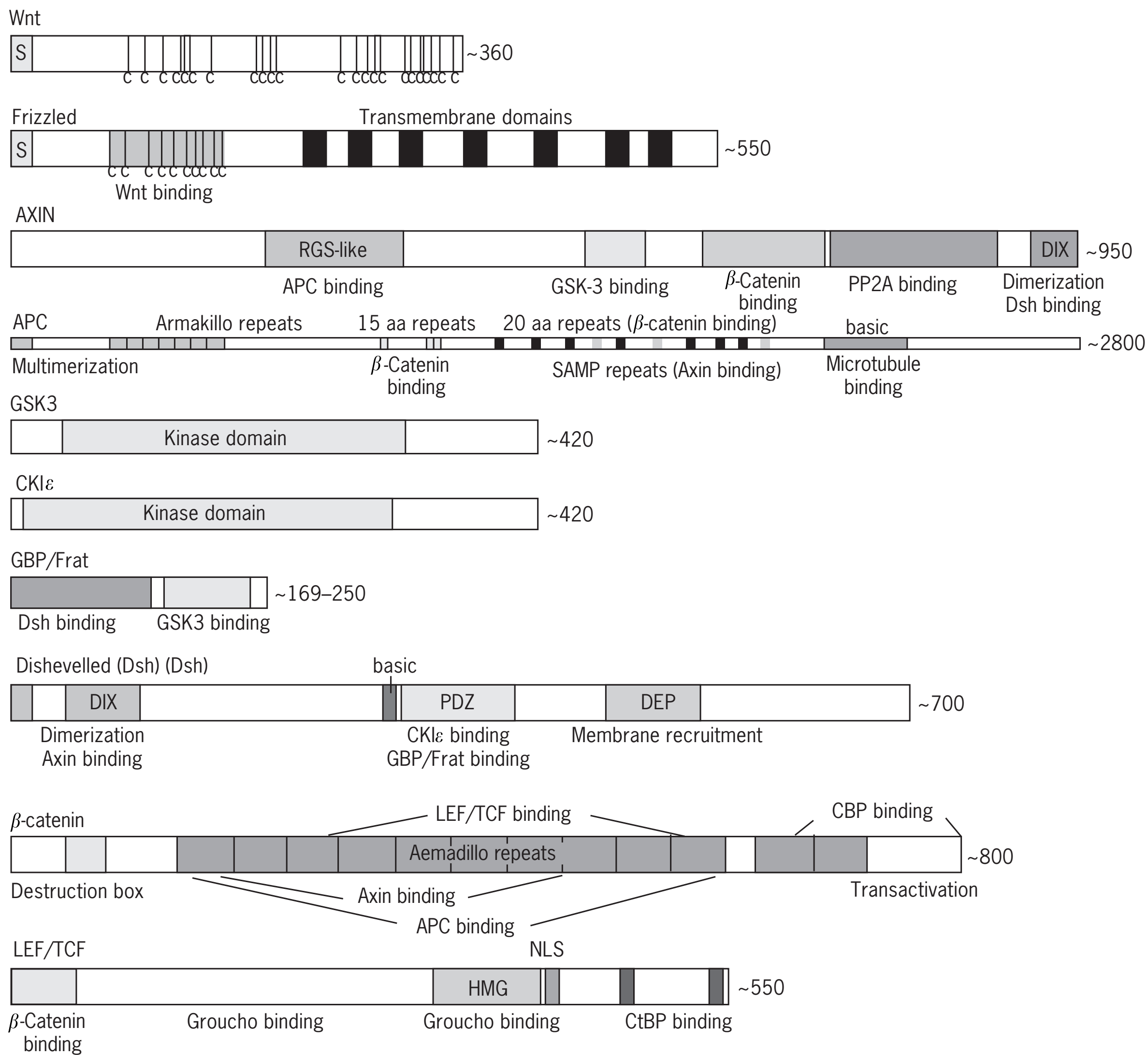


Figure 1 Schematic representation of the major components of the Wnt pathway showing their important structural and functional characteristics. For Wnt and Frizzled: (s) indicates signal sequence and (c) indicates conserved cysteine residues. Axin possesses an RGS domain, found in a family of proteins that regulate signalling via heterotrimeric G-proteins, and a DIX domain, found in both Axin and Dishevelled. Regions implicated in the binding of APC, GSK3, β-catenin, Protein Phosphatase 2A (PP2A) and Dishevelled (Dsh) are shown. APC is very large protein comprised of the following domains: a multimerization domain; seven Armadillo repeats important for localization; a series of β-catenin binding sites (15 amino acid repeats and 20 amino acid repeats), a series of SAMP repeats important for Axin binding; and a basic microtubule binding domain. Both GSK3 and CKIε are serine–threonine protein kinases and the kinase domain is indicated. GBP/Frat is a small protein with an N-terminal Dishevelled binding domain and a C-terminal GSK3 binding domain. Dishevelled possesses the following conserved domains: an N-terminal DIX domain important for dimerization and Axin binding, a centrally located PDZ domain that binds GBP/Frat and CKIε and a C-terminal DEP domain important for membrane recruitment and planar cell polarity signalling. β-Catenin possesses the N-terminal destruction box containing four GSK3 phosphorylation sites and a series of Armadillo repeats. Binding sites for Axin, APC and LEF/TCF have been mapped to the Armadillo repeat region. The N-terminal domain is important for trans-activation. TCF/LEF contains a nuclear localization sequence (NLS) and an HMG box that mediates DNA binding. Binding sites for β-catenin, Groucho, CtBP and CBP/p300 are indicated.

they are notoriously insoluble and troublesome to work with biochemically. The difficulty to acquire soluble forms of Wnt proteins has hindered progress in understanding how cells send and receive Wnt signals. However, several forms of Wnt protein have been recovered from the medium of cultured cells and, using *in vitro* assays for activity, these soluble forms have been shown to be biologically active. This work will, it is hoped, provide the tools for

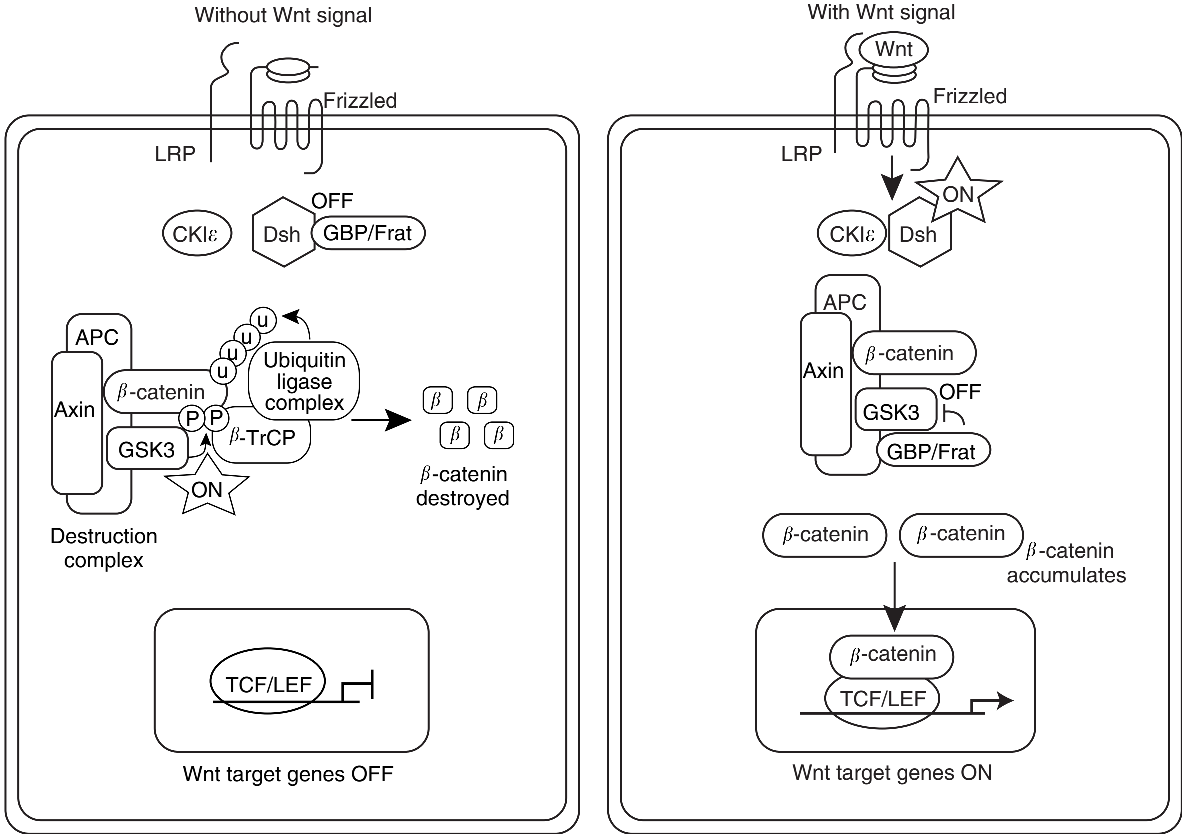


Figure 2 The Wnt signalling pathway. Left: in the absence of Wnt signal, Axin and APC facilitate the addition of phosphate groups to β-catenin by GSK3. Phosphorylated β-catenin binds to β-TrCP and is modified by the addition of a polyubiquitin tag. Proteins tagged with ubiquitin are degraded by the proteasome. The pathway is OFF because β-catenin is rapidly destroyed and its levels in the cell are low. Right: binding of Wnt to cell surface Frizzled and LRP receptors 'activates' Dishevelled (Dsh), CKIε and GBP/Frat by an unknown mechanism. Activation leads to inhibition of the destruction complex and a decrease in the phosphorylation of β-catenin. Thus, β-catenin evades the proteasome and accumulates in the nucleus where it interacts with a DNA-binding protein of the LEF/TCF family. Together, they activate expression of Wnt target genes and the pathway is ON.

understanding how Wnt proteins interact with components of the extracellular matrix and cell surface receptors.

At the Cell Surface – The Reception of Wnt Signals

For many years the identity of Wnt receptors remained elusive, leaving a large gap in our understanding of Wnt signalling. This hole has now been filled with the finding that members of the *Frizzled* gene family can function as Wnt receptors (**Figures 1** and **3**; Bhanot *et al*., 1996; Yang-Snyder *et al*., 1996). *Frizzled* genes encode seven-transmembrane proteins with an N-terminal cysteine-rich domain that binds Wnts with high affinity. Like the *Wnt* gene family, members of the *Frizzled* gene family have been identified in a number of organisms from worms to humans with each organism possessing multiple *Frizzled* genes. A list of known *Frizzled* genes and sequence alignments can be found on the *Wnt* gene homepage (http://www.leland.stanford.edu/~rnusse/wntwindow.html). Work in vertebrate systems has demonstrated that different Wnt proteins preferentially interact with and signal through specific subsets of Frizzled receptors. Given the large number of Wnt and Frizzled genes and, as a consequence, the large number of possible Wnt–Frizzled combinations, deciphering the importance of specific Wnt–Frizzled interactions *in vivo* will be challenging.

Very little is known about how Frizzled proteins function. Structurally, Frizzled receptors resemble other seven-transmembrane receptor proteins that signal through heterotrimeric GTP-binding proteins (G-proteins). Recent evidence suggests that in vertebrates a subset of Frizzled receptors may signal through G-proteins to affect levels of intracellular Ca^{2+}. However, there is no biochemical evidence that Frizzled proteins directly bind G-proteins or that G-proteins are directly involved in promoting the stabilization of β-catenin. Thus, the mechanism of Frizzled

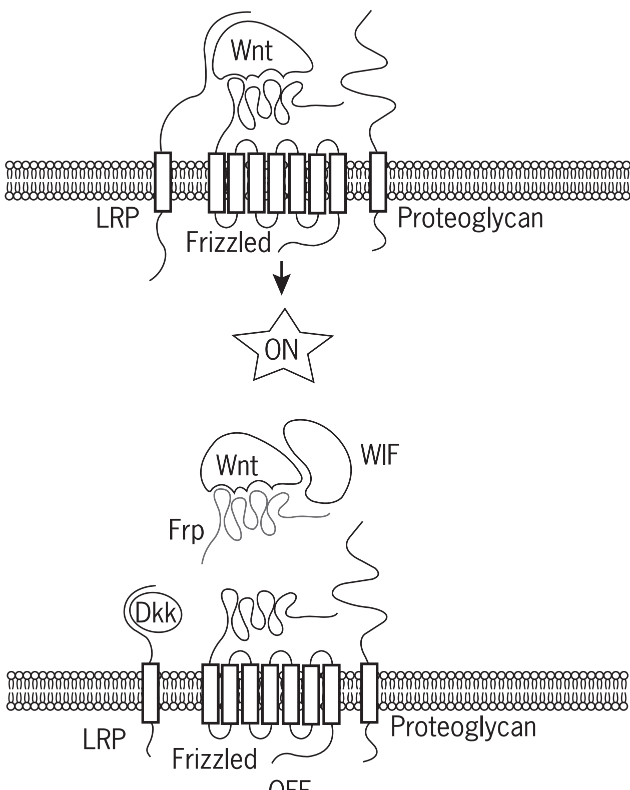


Figure 3 Wnt signalling at the cell surface. Top: Wnt proteins are secreted and bind Frizzled receptors with high affinity at a conserved domain in the N-terminal region. The mechanism by which Frizzled proteins transduce the Wnt signal across the membrane is unclear but may involve signalling through heterotrimeric G-proteins. Wnt proteins also interact with LRP co-receptors. Proteoglycans are also involved in the reception of Wnt signals through an unknown mechanism. Bottom: Wnt signalling is modulated extracellularly by a variety of secreted Wnt inhibitors including members of the secreted Frizzled-related proteins (FRPs), WIF-1 and Dickkopf (Dkk). Binding of these proteins to Wnt is thought to prevent the interaction of Wnt with Frizzled, thereby preventing activation of the pathway.

action and the potential importance of G-proteins in Wnt signalling await further experimentation.

Although compelling evidence exists that Frizzled proteins act as Wnt receptors, it is clear that other cell surface and extracellular molecules also play roles in the reception of Wnt signals (**Figure 3**). For example, members of the LDL-receptor-related family of transmembrane proteins (LRP 5/6) have recently been shown to act as Wnt co-receptors. Proteoglycans, extracellular or cell surface proteins that consist of a protein core and at least one glycosaminoglycan sugar side chain, have been shown to be important regulators of Wnt signalling. In addition, reception of Wnt signals is also modulated extracellularly by a diverse group of secreted Wnt inhibitors. At present, these Wnt inhibitors include members of the secreted Frizzled-Related Protein family (called FrzB or FRP), Wnt-inhibitory factor-1 (WIF-1), Cerberus and Dickkopf (Dkk). FRPs, WIF-1 and Cerberus can directly bind Wnt proteins and are thought to antagonize Wnt function by preventing their interaction with Frizzled receptors. FRP can also interact with Frizzled, suggesting that FRPs may also antagonize Wnt signalling through the formation of a nonfunctional complex with Frizzled receptors. Dkk functions by binding to LRP and preventing the interaction between LRP, Wnt, and Frizzled. Together, these data demonstrate that the reception of Wnt signals is complex, involving the interplay between Wnt protein, Frizzled and LRP cell surface receptors, proteoglycans and secreted Wnt inhibitors. Understanding the nature and importance of these interactions will require further biochemical analyses of the Wnt reception complex and how this complex transduces the Wnt signal from the outside to the inside of the cell.

From the Membrane to the Nucleus – Transduction of Wnt Signals inside the Cell

Wnt signalling is dependent on the levels of β-catenin in the cell. The pathway is OFF when levels of β-catenin are low and the pathway is ON when levels of β-catenin are high. This begs the question of how the levels of β-catenin are regulated in the cell. The answer to this question requires understanding the function of two competing groups of proteins. On one side, the destruction complex, GSK3, Axin and APC work to destroy β-catenin keeping levels low and the pathway OFF. On the other side, CKIε, Dishevelled and GBP/Frat are activated in response to Wnt and work to antagonize the destruction complex, increasing the levels of β-catenin and turning the pathway ON. The following sections examine the role of each of these groups of proteins in Wnt signalling in more detail.

The Destruction Complex: GSK3, Axin and APC

GSK3, a serine–threonine protein kinase, is the central player in the destruction complex (**Figures 1** and **4**). In the absence of Wnt signal, GSK3 is active and adds phosphate groups to four N-terminal sites of β-catenin (S33, S37, T41 and S45 in human β-catenin). These phosphoamino acids act as a tag on β-catenin, marking it for destruction by the proteasome. The amino acid sequence in β-catenin that is phosphorylated by GSK3 is called the 'destruction box' to denote its involvement in regulating the stability of β-catenin. Mutation of the GSK3 phosphorylation sites within the destruction box significantly diminishes the phosphorylation of β-catenin and results in highly stable forms of β-catenin with increased activity (Yost *et al.*, 1996; Pai *et al.*, 1997). The importance of the destruction box sequence is highlighted by

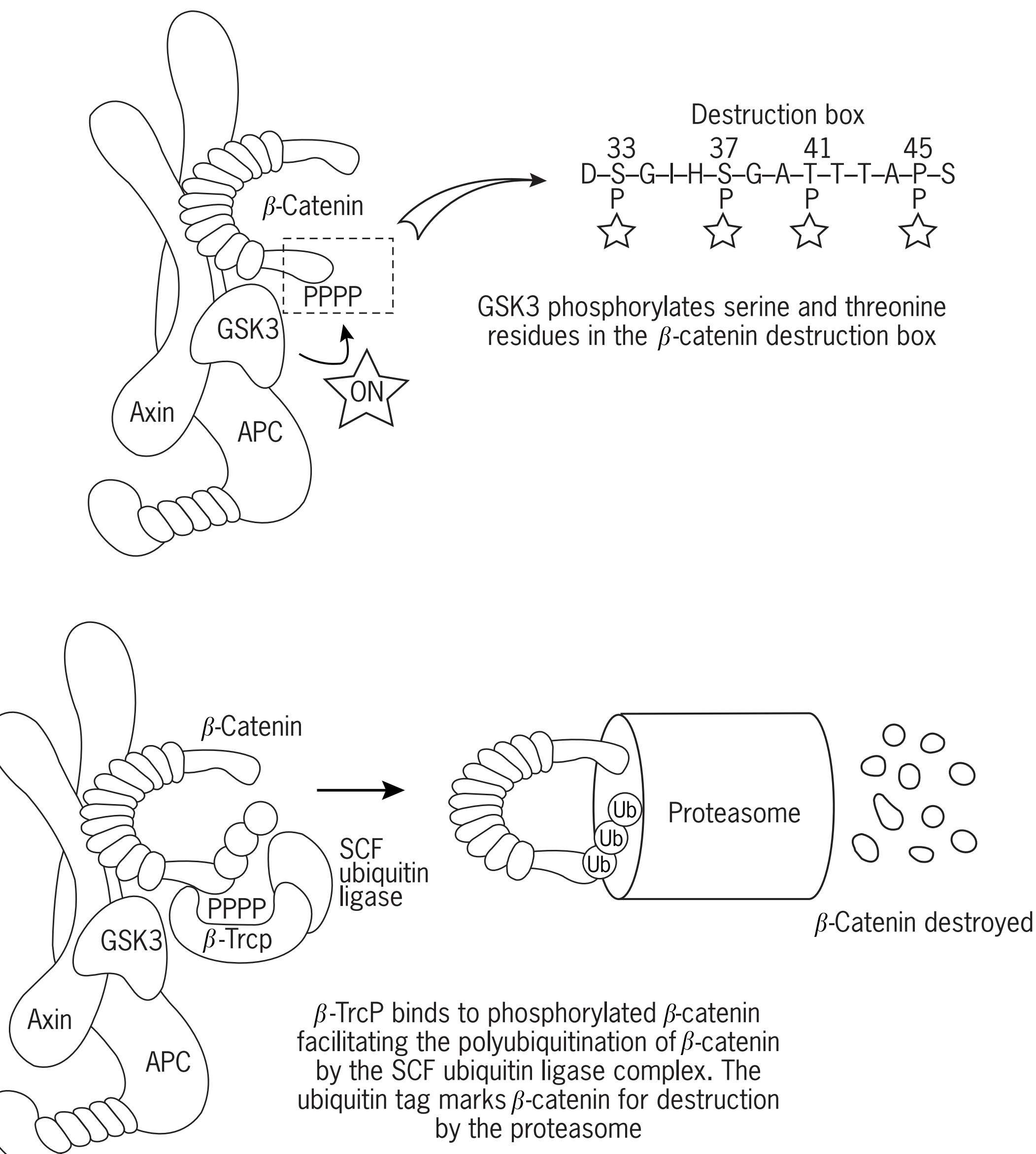


Figure 4 The destruction complex. In the absence of Wnt signal a large multiprotein machine, called the destruction complex, facilitates the rapid destruction of β-catenin. Top: two proteins, Axin and APC, act as scaffold proteins forming the underlying structure of the complex. APC and Axin bind themselves, each other and β-catenin. Additionally, Axin binds GSK3, Dishevelled and PP2A. The binding of GSK3 and β-catenin to Axin is critical for bringing GSK3 in close proximity to β-catenin and stimulating the GSK3-mediated phosphorylation of β-catenin at serine and threonine residues in the destruction box. Bottom: phosphorylated β-catenin is then bound by β-TrCP, which promotes the addition of a polyubiquitin tag to β-catenin. This ubiquitin tag marks β-catenin for rapid destruction by the proteosome, the cell's protein incinerator.

the recent finding that residues within the destruction box are often mutated in human cancers (**Figures 3** and **7**, discussed in detail below).

Axin is a second key component of the destruction complex (**Figures 1** and **4**). Mice lacking functional Axin develop with defects in the patterning of the dorsal–ventral axis, a phenotype similar to that seen following ectopic activation of the Wnt pathway in frogs (Zeng *et al*., 1997). In addition, over-expression of Axin can inhibit Wnt signalling and promote the degradation of β-catenin. Axin possesses multiple protein–protein interaction domains and appears to act as scaffold proteins, i.e. it serves as a building block for the construction of multiprotein complexes. Axin binds several components of the Wnt pathway including Dishevelled, APC, GSK3 and β-catenin. The binding of both GSK3 and β-catenin appears to be critical for the function of Axin as this interaction greatly enhances the phosphorylation of β-catenin by GSK3. Thus, Axin

appears to promote β-catenin degradation by bringing GSK3 and β-catenin into close proximity, thereby facilitating the phosphorylation of β-catenin by GSK3. Recently, mutations in Axin have been found in hepatocellular carcinomas and colon cancer, underscoring the importance of Axin in regulating the activity of the Wnt signalling pathway (Satoh *et al.*, 2000; Webster *et al.*, 2000).

APC is the third critical component of the destruction complex (**Figures 1** and **4**). APC was originally identified as a tumour-suppressor protein and mutations in APC are found in >80% of all colorectal tumours. The *APC* gene encodes a large, multidomain protein that, like Axin, appears to function as a scaffold protein. APC binds to several components of the Wnt pathway, including Axin, GSK3 and β-catenin. The first clue into the function of APC came from the finding that colorectal adenocarcinoma cell lines harbouring mutations in the *APC* gene possess high levels of β-catenin. Expression of wild-type APC in these cells resulted in a dramatic reduction in β-catenin levels, suggesting that APC is negative regulator of β-catenin stability (Munemitsu *et al.*, 1995). How does APC promote the degradation of β-catenin? Current models predict APC functions in a similar fashion to Axin, stimulating GSK3-mediated phosphorylation of β-catenin. An alternative idea is that APC may function to localize the destruction complex to a specific location in the cell. In support of this hypothesis, mutations in APC that perturb its normal cortical location in the cell also perturb its ability to promote β-catenin degradation (McCartney *et al.*, 1999). These two models are not mutually exclusive and further analysis of APC in a variety of systems should clarify the role APC plays in regulating β-catenin stability.

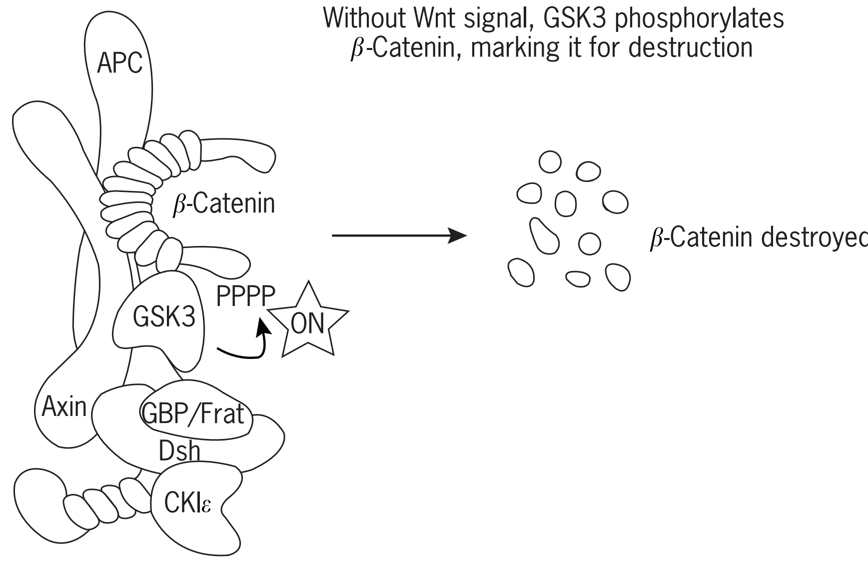
How does phosphorylation of β-catenin stimulate ubiquitination? The answer to this question came recently with the finding that β-TrCP/Slimb, a component of the SCF ubiquitin ligase complex, plays a critical role in regulating β-catenin degradation (**Figure 4**; Jiang and Struhl, 1998; Kitagawa *et al.*, 1999). β-TrCP/Slimb specifically binds the phosphorylated destruction box of β-catenin, resulting in ubiquitination of β-catenin and subsequent proteolysis by the proteasome. What remains unclear, however, is how β-catenin is delivered to its final destination, the proteosome. Together, these data shed light on how mutations in APC, Axin and β-catenin lead to hyperactivation of the Wnt pathway and cancer. Recessive mutations in the tumour suppressors APC and Axin would lead to the inability of the destruction complex to target β-catenin for degradation in the absence of Wnt signals. Oncogenic mutations in the destruction box of β-catenin would prevent phosphorylation by GSK3 and/or the interaction of β-catenin with β-TrCP/Slimb. In each of these cases, β-catenin would evade proteasomal degradation and accumulate in the cell, leading to inappropriate activation of the pathway. This idea is supported by a number of studies showing that primary human tumour cells harbouring mutations in APC, Axin or β-catenin display elevated levels of β-catenin. This knowledge provides potential targets for clinical intervention and will be invaluable for the designing and testing new therapeutic agents for the treatment of cancer in humans.

Antagonizing the Destruction Complex: CKIε, Dishevelled and GBP/Frat

How does activation of the Wnt pathway promote the stabilization of β-catenin in the cell? The answer to this question hinges on understanding how Wnt signals, once transduced across the plasma membrane, act to antagonize the destruction complex. Until recently, our knowledge of the molecular events that occur upon Wnt stimulation was limited, but major gaps have been filled with the characterization of the biochemical relationships between intracellular components of the Wnt pathway.

Dishevelled is the most upstream intracellular component of the Wnt pathway and is 'activated' in response to Wnt signals (**Figures 1** and **5**). Dishevelled appears to function as a scaffold protein, acting through its association with other signalling proteins. Dishevelled possesses three highly conserved domains important for its function. The first is an N-terminal DIX domain that shares sequence identity to a C-terminal domain of Axin. The DIX domain has been shown to be important for homodimerization of Dishevelled and the binding of Axin. Second, is a centrally located PDZ domain. PDZ domains are present in a variety of proteins and serve as sites for protein–protein interactions. A number of proteins have been shown to bind Dishevelled through the PDZ domain including CKIε, GBP/Frat and Casein Kinase II. Both the DIX and PDZ domains are required for Dishevelled function in Wnt signalling. Third is the DEP domain found in the C-terminal third of Dishevelled. The DEP domain shares sequence similarity to the *Caenorhabditis elegans* gene *egl-10* and pleckstrin. Although the DEP domain is not required for the ability of Dishevelled to stabilize β-catenin, it is required for the function of Dishevelled in regulating cell polarity in flies and cell movements during gastrulation in vertebrates. It remains unresolved how Dishevelled transduces Wnt signals, but recent studies showing that Dishevelled interacts with Axin suggest that it may play a direct role in antagonizing the destruction complex.

Dishevelled function may also be dependent on its localization within the cell. In *Xenopus*, activation of the Wnt pathway is required for the establishment of dorsal cell fates. Examination of the localization of Dishevelled in early embryos revealed that it associates with small vesicle-like organelles (0.5–1.0 μm in diameter) that are enriched on the dorsal side of the embryo (Miller *et al.*, 1999). This localization appears to be important since treatments that prevent dorsal development also prevent the dorsal enrichment of Dishevelled. Interestingly, time-lapse confocal microscopy analysis of Green Fluorescent Protein tagged Dishevelled (Dishevelled-GFP)



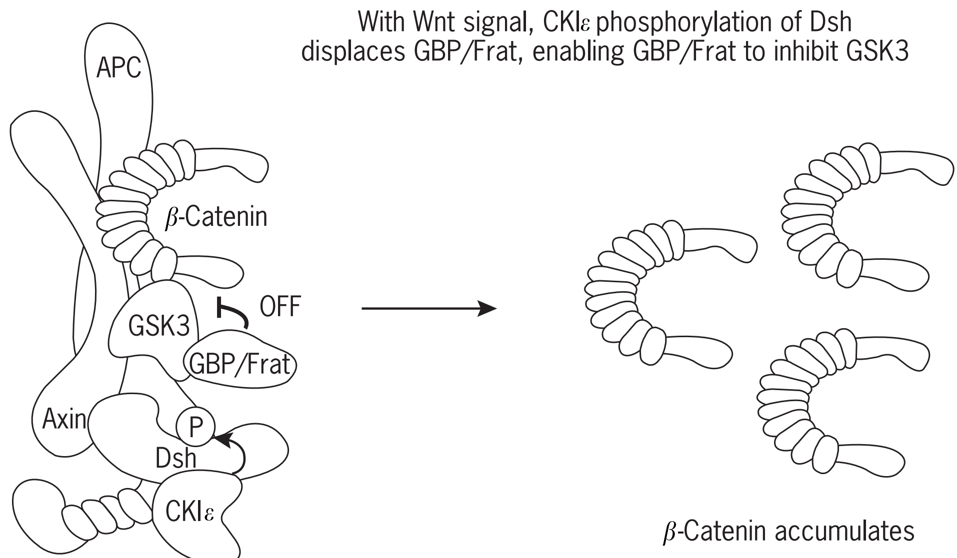


Figure 5 A current model of Wnt signalling and stabilization of β-catenin. Top: without Wnt signal, it is thought that Dishevelled might bind Axin and GBP/Frat, sequestering GBP/Frat from GSK3. Bottom: Wnt signal leads to antagonism of the destruction complex. One current model predicts that Dishevelled, through its interaction with Axin, might bring GBP/Frat to the destruction complex. Upon Wnt stimulation, CKIε might phosphorylate Dishevelled, displacing GBP/Frat, allowing it to bind and inhibit GSK3. Through this, or a similar mechanism, Wnt signal may also promote the dissolution of the destruction complex. Inhibition of GSK3 protects β-catenin from degradation and promotes the accumulation of β-catenin in the cell. This model is supported by recent studies but there are other possible mechanisms. Further characterization of the biochemical relationships between components of the destruction complex will help to resolve how Wnt signals stabilize β-catenin.

localization in early embryos demonstrated that Dishevelled-GFP associates with and is transported along the microtubule cytoskeleton towards the prospective dorsal side of the embryo. To view movies of Dishevelled transport in frog embryos, visit the *Journal of Cell Biology* web page at http://www.jcb.org/cgi/content/full/146/2/427/DC1. In addition, Dishevelled has also been shown to be associated with the actin cytoskeleton in embryonic kidney cells (Torres and Nelson, 2000). Together these data suggest that Dishevelled localization, perhaps through its association with the cytoskeleton, plays an important role in modulating the activity of the Wnt pathway.

GBP (GSK3 Binding Protein) and its mammalian orthologue Frat function as positive regulators of the Wnt signalling pathway (**Figures 1** and **5**; Yost *et al*., 1998). GBP/Frat can inhibit GSK3 kinase activity, suggesting that it promotes β-catenin stabilization through direct inhibition of GSK3-mediated phosphorylation of β-catenin. However, it is unclear whether this affect is due to a change in GSK3 activity or through steric blockade of GSK3–substrate interactions.

More recently, CKIε, a serine–threonine protein kinase, was identified as a positive regulator of the Wnt pathway (**Figures 1** and **5**; Peters *et al*., 1999; Sakanaka *et al*., 1999). Expression of CKIε stabilizes β-catenin and expression of dominant negative forms of CKIε antagonizes Wnt signalling. Overexpression studies have placed CKIε downstream of Dishevelled and upstream of GSK3. In addition, CKIε can bind to and phosphorylate Dishevelled, suggesting that CKIε could directly affect the activity of Dishevelled.

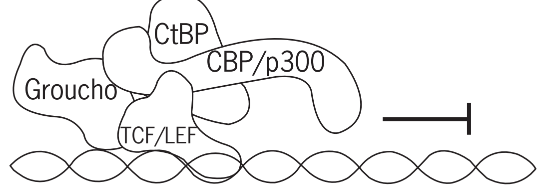
How do Dishevelled, GBP/Frat and CKIε antagonize the destruction complex? In the absence of Wnt signals, Dishevelled might associate with Axin and bind GBP/Frat, sequestering it from GSK3. Upon Wnt stimulation, CKIε might phosphorylate Dishevelled, causing the displacement of GBP/Frat. Released GBP/Frat would then bind GSK3 that is associated with Axin, thereby inhibiting GSK3 activity. This model is consistent with current data, but so are many others and many questions are unresolved. How do Frizzled receptors transduce Wnt signals? How are Dishevelled, GBP/Frat and CKIε activated in response to Wnt signals? Where in the cell is the destruction complex located and is this location important for signal transduction? These and other questions await further experimental analysis.

In the Nucleus – Regulation of Gene Expression by β-Catenin and TCF/LEF

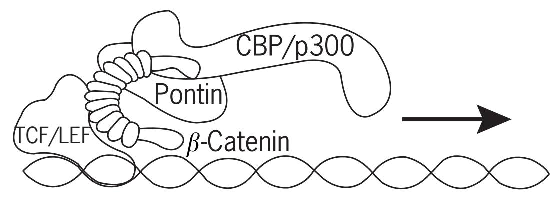
As we have seen, cells have evolved a very elaborate and complex mechanism for regulating intracellular levels of β-catenin. Wnt signalling frees β-catenin from the destruction complex and it now accumulates in the cell. What happens next? It is now dogma that upon Wnt stimulation β-catenin accumulates and enters the nucleus, where it regulates gene expression. However, when first proposed the idea that the nucleus was the primary location of β-catenin function in Wnt signalling seemed hard to believe. At the time, β-catenin was known as a protein that localized to cell–cell junctions and played a crucial role in regulating cell–cell adhesion. This raised the perplexing question of how a cell adhesion protein could also be a signalling protein in Wnt pathway that affected gene expression in the nucleus. The answer came from studies showing that β-catenin's roles in cell adhesion and signalling were separable and involved the interaction of β-catenin with distinct sets of protein partners. Several groups also showed that β-catenin, in addition to its membrane localization, also localizes to the nucleus and that Wnt signalling caused an enrichment of β-catenin in the nucleus. Now this idea is so prevalent that nuclear β-catenin localization is used as a diagnostic tool for Wnt pathway activation in development and oncogenic activation of the Wnt pathway in cancer.

In the nucleus, β-catenin binds to a number of different protein partners to regulate gene expression (**Figure 6**). These partners include members of the LEF/TCF family of transcription factors (**Figures 1** and **6**). The LEF/TCF proteins are sequence-specific DNA binding proteins and serve to localize β-catenin to the promoters of Wnt target genes. A number of these target genes have been identified in the past several years and include developmental regulatory genes such as *siamois*, *twin* and *Xnr-3* in *Xenopus* and *ultrabiothorax* in *Drosophila*. Additional targets include regulators of cell growth and proliferation, c-*myc*

Wnt signalling in the nucleus



Wnt target genes OFF



Wnt target genes ON

Figure 6 Regulation of gene expression by β-catenin and TCF/LEF. In the absence of Wnt signal, TCF/LEF transcription factors specifically bind to sequences in the promoters of Wnt target genes and act as repressors keeping these genes OFF. This repression is mediated through the interaction of TCF/LEF with a number of transcriptional repressor proteins including members of the Groucho family, CtBP and CBP/p300. Upon Wnt stimulation, β-catenin accumulates in the nucleus and interacts with TCF/LEF. This complex specifically binds to sites in the promoters of Wnt target genes and through interactions with additional transcription factors (e.g. CBP/p300 and Pontin 52) activate transcription. The mechanism by which the β-catenin–TCF/LEF complex activates transcription is unclear, but may involve the displacement of repressors bound to TCF/LEF by β-catenin.

and *cyclin D1*. For a complete list of known targets of the Wnt pathway, see the Wnt gene homepage (http://www.stanford.edu/~rnusse/wntwindow.html).

The mechanism by which β-catenin activates transcription remains unclear. Mutational analyses have identified two regions of β-catenin, one near the N-terminus and the other at the C-terminus, that are important for transcriptional activation. It appears that these sites may serve as protein–protein interaction domains enabling β-catenin to regulate transcription through binding of additional transcriptional regulators (**Figure 6**). Several binding partners have been identified including CBP/p300 (Creb Binding Protein) and Pontin 52. In addition, Lef-1 has been shown to form complexes with several members of the SMAD family of signalling proteins that play a role in Transforming Growth Factor β signalling (see below). Lef-1 and SMADs appear to function cooperatively to stimulate expression of specific developmental target genes such as *twin*.

Another possible mechanism by which β-catenin may activate transcription is by displacing co-repressors from TCF/LEF (**Figure 6**). In the absence of Wnt signals, TCF/LEF proteins can act as transcriptional repressors, preventing transcription of Wnt/β-catenin target genes. TCF/LEF proteins do not appear to act alone but instead require interactions with one of several identified co-repressors. These repressors include members of the Groucho family, CtBP (C-terminal Binding Protein), and CBP/p300. It is thought that the ability of CBP/p300 to act as an activator and a repressor may be due to differences in cellular context. In other words, CBP/p300 can function as either an activator or repressor depending on the situation.

ADDITIONAL REGULATORS OF Wnt SIGNAL TRANSDUCTION

Although the mechanism of Wnt signal transduction described thus far may seem complicated enough, additional players continue to be identified. For example, protein phosphatase 2A appears to play a role in regulating β-catenin stability although it is unclear whether it promotes β-catenin degradation or stabilization. The *Drosophila naked cuticle* gene has recently been shown to be a novel cytoplasmic antagonist that may limit the potency, duration or distribution of Wnt signals. Recent evidence also implicates components of the mitogen-activated protein kinase (MAPK) pathway, transforming growth factor-β-activated kinase (TAK-1) and NEMO-like kinase (NLK) as regulators of Wnt signalling. However, it remains to be determined whether these genes are true components of the Wnt pathway or whether they act in parallel to the canonical Wnt pathway.

In addition to these players, it has become clear that cross-talk between the Wnt pathway and the Transforming Growth Factor β (TGF-β) signalling pathway plays an important role in regulating Wnt/β-catenin signalling both during development and in human disease. For example, the secreted Wnt antagonist Cerberus can also interact with members of the BMP and Nodal families of TGF-β-like signalling molecules, suggesting that Cerberus can function as a multivalent modulator of both Wnt and TGF-β/BMP signalling. In addition, Wnt and TGF-β signalling pathways also cross-talk inside the cell. Several recent papers have shown that TCF/LEF transcription factors interact with members of the SMAD family of TGF-β/BMP signal mediators. Specifically, Lef-1 can form complexes with three different SMAD proteins: SMAD-2 and -3, effectors for TGF-β and Activin signals; and SMAD-4, a ubiquitous effector for all TGF-β/BMP signalling pathways. Through these interactions, SMAD proteins were found to stimulate synergistically transcription of specific Wnt target genes. These data are intriguing since mutations in components of both the Wnt (APC and β-catenin) and TGF-β signalling pathways (TGF-β receptor type II, SMAD-2, -3 and -4) are associated with colorectal cancers. Furthermore, mice double heterozygous for both *APC* and *SMAD-4* display intestinal polyps that develop into more malignant tumours than those in mice heterozygous for *APC* alone (Takaku *et al*., 1998). Together these data argue that the Wnt and TGF-β pathways cross-talk to regulate cooperatively gene expression and that this synergistic interaction may be important both during development and in cancer.

β-Catenin levels can also be regulated by Wnt-independent mechanisms. For example, expression of integrin-linked kinase in mammalian cells promotes the stabilization and nuclear accumulation of β-catenin (Novak *et al*., 1998). Presenilin proteins have also been implicated as regulators of β-catenin stability. Mutations in presenilin associated with the rapid onset of Alzheimer disease decrease the stability of β-catenin in neurons. This effect on β-catenin was also correlated with an increase in the susceptibility of neurons to apoptosis resulting from the accumulation of β-amyloid protein. Given the ability of these signalling pathways to modulate β-catenin stability, it seems likely that β-catenin may regulate many cellular processes independent of its role in Wnt signalling.

ONCOGENIC ACTIVATION OF THE Wnt PATHWAY

Tumour formation results from the loss of control over cell proliferation. This occurs through mutations that produce oncogenes with a dominant gain of function or inactivate tumour suppressor genes through recessive loss of function mutations. Both types of mutations lead to defects in regulatory pathways that normally control cell proliferation. Recently, it has become clear that components of the

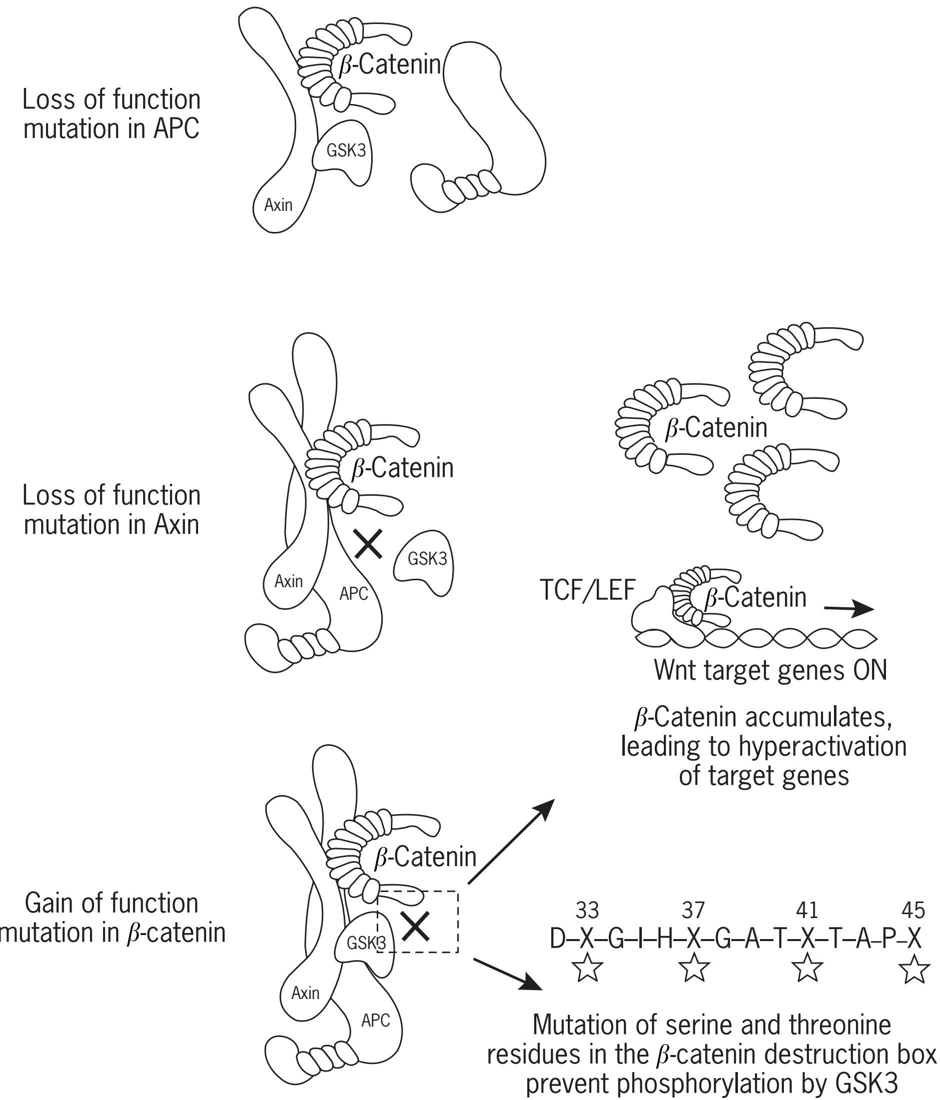


Figure 7 Oncogenic activation of the Wnt signalling pathway. Recessive loss of function mutations in the tumour suppressors APC and Axin disable the destruction complex and lead to the inappropriate accumulation of β-catenin in the cell. Most mutations isolated in *APC* are nonsense mutations that lead to the premature truncation of the protein and loss of critical binding sites for Axin and β-catenin. One mutation found in Axin appears to decrease the binding affinity between Axin and GSK3. Gain of function mutations in β-catenin that allow it to escape regulation by the destruction complex also lead to constitutive activation of the Wnt pathway. The majority of mutations in β-catenin found thus far are missense mutations that alter one of the four potential GSK3 phosphorylation sites in the destruction box. The inappropriate accumulation of β-catenin in the cell then leads to hyperactivation of target genes such as c-*myc* and *cyclin D1*, and uncontrolled cell division.

Wnt signalling pathway are mutated in a variety of human cancers. Thus far, mutations in APC, Axin and β-catenin have been identified in various cancers and these findings have fuelled a great explosion of interest in the relationship between Wnt signalling and cancer. These mutations result in the inability of the cell to regulate appropriately levels of β-catenin (**Figure 7**). Recessive mutations in the tumour suppressor genes *APC* and *Axin* result in defects in the destruction complex allowing β-catenin to escape degradation. Conversely, mutations in β-catenin produce a dominant gain of function protein that evades regulation by the destruction complex. The end result of these mutations is the constitutive activation of Wnt target genes and uncontrolled cell proliferation.

Germline mutations of the *APC* gene lead to familial adenomatous polyposis characterized by the development

Table 1 Current list of human cancers associated with mutations in β-catenin, APC and Axin

β-Catenin	**APC**	**Axin**
Colorectal adenoma and carcinoma	Colorectal adenoma and carcinoma	Colorectal
Endometrial carcinoma	Breast cancer	Hepatocellular carcinoma
Hepatoblastoma	Medulloblastoma	
Hepatocellular carcinoma		
Malignant fibrous histiocytoma		
Medulloblastoma		
Ovarian carcinoma		
Pilomatricoma		
Prostate		
Synovial sarcoma		
Uterine		
Wilms' tumours		

of colorectal polyps in the second to third decade of life. In addition, somatic mutations of the *APC* gene are associated with >80% of sporadic colorectal adenomas and carcinomas. More than 95% of germ-line and somatic mutations of the *APC* gene are nonsense mutations that result in the synthesis of a truncated protein that lacks the region of APC important for its function in the destruction complex. Significantly, these truncations in APC remove binding sites for β-catenin and Axin and also putative phosphorylation sites for GSK3 (**Figures 1** and **7**). Colon cancer cells expressing such a truncated form of APC possess very high levels of β-catenin, suggesting that the inability of APC to promote β-catenin degradation is causally linked to tumour formation. In addition to the well-documented link between APC function and colorectal cancer, mutation of *APC* is also implicated in other human cancers including aggressive fibromatosis and breast cancer (**Table 1**). Together these findings highlight the importance of APC in Wnt signalling and oncogenesis.

Recent evidence demonstrates that *Axin*, like *APC*, functions as a tumour-suppressor gene (**Figure 7**, **Table 1**). The *Axin1* gene is mutated in hepatocellular carcinomas and missense mutations in *Axin* have also been isolated in colon cancer cells (Satoh *et al.*, 2000; Webster *et al.*, 2000). One of these mutations was found to diminish the interaction between Axin and GSK3, providing a potential mechanism for oncogenic activation of the Wnt pathway.

The striking connection between Axin- and APC-mediated regulation of β-catenin and oncogenesis suggested that mutations in β-catenin itself might play a role in tumour formation. Consistent with this idea, targeted mutation or deletion of the destruction box sequence results in highly stable forms of β-catenin that are hyperactive (Yost *et al.*, 1996; Pai *et al.*, 1997). Cell lines harbouring activating mutations of β-catenin often display high levels of both cytoplasmic and nuclear β-catenin and constitutive activation of TCF/LEF reporter genes. In addition, expression of an N-terminal truncated form of β-catenin lacking the destruction box in the epidermis of transgenic mice resulted in the formation of hair follicle-related tumours. These data demonstrate that increasing levels of β-catenin are sufficient to promote tumour formation and implicate the N-terminal destruction box sequence as potential sites for oncogenic activation of β-catenin.

Over the past several years, many studies have shown that mutations in β-catenin exist in a variety of human cancers (**Table 1**). Strikingly, the identified mutations are missense or deletion mutations in the destruction box and most of these alter one of the GSK3 phosphorylation sites (**Figure 7**). In addition to the GSK3 sites, missense mutations at aspartate-32 and glycine-34 have also been reported. These residues, along with serine-37, have been characterized as a ubiquitination target motif based on its similarity with Iκ-B, another protein targeted for degradation by the $SCF^{\beta\text{-TrCP}}$ ubiquitin ligase complex. Together, these data strongly argue that mutations in β-catenin that enable it to evade regulation by the destruction complex play an important role in tumorigenic transformation of many cell types.

What is the consequence of constitutive β-catenin activity? β-Catenin functions as a transcriptional activator in Wnt signalling, suggesting that its role in tumour formation may be through inappropriate activation of genes important for regulating cell division and growth. This idea has gained support with the finding that the c-*myc* and *cyclin D1* genes, both of which are known to promote cell proliferation, are direct targets of β-catenin. c-*myc* is a potent oncogene that regulates cell cycle progression, promoting the G_1/S phase transition. Similarly, *cyclin D1* also plays an important role regulating movement through the cell cycle. Thus, mutations in Axin, APC or β-catenin that result in the stabilization and accumulation of β-catenin may lead to inappropriate expression of target genes, such as c-*myc* and *cyclin D1*. Expression of c-*myc* and *cyclin D1* would then expedite the G_1 to S transition, leading to uncontrolled cell proliferation. Additionally, evidence also suggests that β-catenin may act as a survival factor protecting cells from cell

death. Since the transition from G_1 to S phase requires the presence of survival factors, β-catenin may also stimulate this transition directly by preventing apoptosis and permitting cell cycle progression.

CONCLUSION

With the recent completion of the human genome sequence, we have entered a new era in biology and medicine. This achievement has been compared to putting a man on the moon and will undoubtedly revolutionize basic biological and medical sciences. This revolution, however, is only in its infancy. What remains is put the pieces of the puzzle together by characterizing the function of each of the estimated 80 000 human genes. How do these genes instruct cells to divide, to migrate or to die? How do these genes control embryogenesis? How do mutations in specific genes contribute to human disease?

Although many aspects of Wnt signalling remain unclear, we are beginning to put the pieces of the Wnt puzzle together to understand how the genes involved in Wnt signalling communicate signals between cells. In particular, great strides have been made towards understanding the molecular and biochemical mechanics of Wnt signal transduction. These findings underscore the important predictive value of analysing the fundamental mechanisms by which cells signal to one another: by establishing how genes function in specific signalling pathways (e.g. whether they act as repressors or activators), one can make educated guesses how these genes might contribute to human disease. A repressor could be a tumour-suppressor gene and an activator could be an oncogene. Thus, a clear understanding the basics of how cells communicate will lead us to an understanding of how inappropriate activation of signalling pathways leads to cancer. With this information, researchers will hopefully be able to design new therapeutic reagents for treating and preventing cancer.

REFERENCES

Bhanot, P., *et al*. (1996). A new member of the *frizzled* family from *Drosophila* functions as a Wingless receptor. *Nature*, **382**, 225–230.

Jiang, J. and Struhl, G. (1998). Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. *Nature*, **391**, 493–496.

Kitagawa, M., *et al*. (1999). An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of beta-catenin. *Embo Journal*, **18**, 2401–2410.

McCartney, B. M., *et al*. (1999). *Drosophila* APC2 is a cytoskeletally-associated protein that regulates *wingless* signaling in the embryonic epidermis. *Journal of Cell Biology*, **146**, 1303–1318.

Miller, J. R., *et al*. (1999). Establishment of the dorsal–ventral axis in *Xenopus* embryos coincides with the dorsal enrichment of dishevelled that is dependent on cortical rotation. *Journal of Cell Biology*, **146**, 427–437.

Munemitsu, S., *et al*. (1995). Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proceedings of the National Academy of Sciences of the USA*, **92**, 3046–3050.

Novak, A., *et al*. (1998). Cell adhesion and the integrin-linked kinase regulate the LEF-1 and beta-catenin signaling pathways. *Proceedings of the National Academy of Sciences of the USA*, **95**, 4374–4379.

Nusse, R. and Varmus, H. E. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell*, **31**, 99–109.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature*, **287**, 795–801.

Pai, L.-M., *et al*. (1997). Negative regulation of Armadillo, a Wingless effector in *Drosophila*. *Development*, **124**, 2255–2266.

Peters, J. M., *et al*. (1999). Casein kinase I transduces Wnt signals. *Nature*, **401**, 345–350.

Sakanaka, C., *et al*. (1999). Casein kinase I-epsilon in the Wnt pathway: regulation of beta-catenin function. *Proceedings of the National Academy of Sciences of the USA*, **96**, 12548–12552.

Satoh, S., *et al*. (2000). *AXIN1* mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of *AXIN1*. *Nature Genetics*, **24**, 245–250.

Takaku, K., *et al*. (1998). Intestinal tumorigenesis in compound mutant mice of both *Dpc4* (*Smad4*) and *Apc* genes. *Cell*, **92**, 645–656.

Torres, M. A. and Nelson, W. J. (2000). Colocalization and redistribution of Dishevelled and Actin during Wnt-induced mesenchymal morphogenesis. *Journal of Cell Biology*, **149**, 1433–1442.

Webster, M. T., *et al*. (2000). Sequence variants of the *axin* gene in breast, colon, and other cancers: an analysis of mutations that interfere with GSK3 binding. *Genes, Chromosomes and Cancer*, **28**, 443–453.

Yang-Snyder, J., *et al*. (1996). A *frizzled* homolog functions in a vertebrate Wnt signaling pathway. *Current Biology*, **6**, 1302–1306.

Yost, C., *et al*. (1996). The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in *Xenopus* embryos by glycogen synthase kinase 3. *Genes and Development*, **10**, 1443–1454.

Yost, C., *et al*. (1998). GBP, an inhibitor of GSK-3, is implicated in *Xenopus* development and oncogenesis. *Cell*, **93**, 1031–1041.

Zeng, L., *et al*. (1997). The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell*, **90**, 181–192.

FURTHER READING

Behrens, J. (2000). Cross-regulation of the Wnt signalling pathway: a role of MAP kinases. *Journal of Cell Science*, **113**, 911–919.

Bienz, M. (1999). APC: the plot thickens. *Current Opinions in Genetics and Development*, **9**, 595–603.

Bienz, M. and Clevers, H. (2000). Linking colorectal cancer to Wnt signalling. *Cell*, **103**, 311–320.

Boutros, M. and Mlodzik, M. (1999). Dishevelled: at the crossroads of divergent intracellular signaling pathways. *Mechanics of Development*, **83**, 27–37.

Cadigan, K. M. and Nusse, R. (1997). Wnt signaling: a common theme in animal development. *Genes and Development*, **11**, 3286–3305.

Czech, C., *et al*. (2000). Presenilins and Alzheimer's disease: biological functions and pathogenic mechanisms. *Progress in Neurobiology*, **60**, 363–384.

Dierick, H. and Bejsovec, A. (1999). Cellular mechanisms of wingless/Wnt signal transduction. *Current Topics in Development in Biology*, **43**, 153–190.

Kikuchi, A. (1999). Roles of Axin in the Wnt signalling pathway. *Cell Signalling*, **11**, 777–788.

Kuhl, M., *et al*. (2000). The Wnt/Ca^{2+} pathway: a new vertebrate wnt signaling pathway takes shape. *Trends in Genetics*, **16**, 279–283.

Miller, J. R., *et al*. (1999). Mechanism and function of signal transduction by the Wnt/beta-catenin and Wnt/Ca^{2+} pathways. *Oncogene*, **18**, 7860–7872.

Peifer, M. and Polakis, P. (2000). Wnt signaling in oncogenesis and embryogenesis – a look outside the nucleus. *Science*, **287**, 1606–1609.

Perrimon, N. and Bernfield, M. (2000). Specificities of heparan sulphate proteoglycans in developmental processes. *Nature*, **404**, 725–728.

Polakis, P. (1999). The oncogenic activation of beta-catenin. *Current Opinions in Genetics and Development*, **9**, 15–21.

Polakis, P. (2000). Wnt signaling and cancer. *Genes and Development*, **14**, 1837–1851.

Roose, J. and Clevers, H. (1999). TCF transcription factors: molecular switches in carcinogenesis. *Biochimica Biophysica Acta*, **1424**, M23–37.

Waltzer, L. and Bienz, M. (1999). The control of beta-catenin and TCF during embryonic development and cancer. *Cancer and Metastasis Reviews*, **18**, 231–246.

Wodarz, A. and Nusse, R. (1998). Mechanisms of Wnt signaling in development. *Annual Reviews in Cell and Development Biology*, **14**, 59–88.

Chapter 15

Extracellular Matrix: The Networking Solution

Nancy Boudreau
University of California San Francisco, San Francisco, CA, USA

Mina J. Bissell
Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA, USA

CONTENTS

- Introduction
- ECM Proteins and Integrin Receptors
- Integration of Cell Morphology and Signalling by ECM
- 3D Organisation and Integrated Tissue Responses
- Conclusion
- Acknowledgements

INTRODUCTION

One of the fundamental properties of functional differentiated tissues is their unique three-dimensional (3D) organisation which allows individual cells to act in a coordinated manner to carry out complex tissue functions. Although individual cells may be capable of eliciting certain responses to external stimuli such as proliferation, many differentiated functions can be achieved only via the coordinated actions of cells within a tissue. For example, endothelial cells which line the walls of functional blood vessels may individually produce necessary anti-clotting factors but alone cannot function as a conduit for blood. Similarly, although rounded single mammary epithelial cells are capable of expressing the tissue-specific milk protein β-casein, they cannot secrete it and expression is significantly enhanced when cells form multi-cellular complexes (Streuli *et al*., 1991). Three-dimensional organisation was also found to be essential for the generation of significant populations of functional T cells from CD34+ progenitors. Only by embedding the progenitor cells into a 3D carbon matrix seeded with thymus extract could mature functional T cells be obtained. (Poznansky *et al*., 2000) Thus 3D organisation is essential not only for developing and establishing functional differentiated phenotypes but also for maintaining tissue specific gene expression and function. Moreover, 3D tissue architecture acts to override the genetic information contained within normal and malignant cells (Weaver *et al*., 1997).

In order for multi-cellular organisms to acquire and maintain a 3D organisation of their tissues, they have elaborated a complex network of extracellular proteins referred to as the extracellular matrix (ECM). This network of extracellular proteins allows cells to adhere, migrate, proliferate and undergo morphogenesis or alternatively instructs them to undergo programmed cell death (Boudreau and Bissell, 1996).

The significance of the ECM in influencing cellular behaviour was initially overlooked as the ECM was considered merely to provide scaffolding to cells and tissues. However, it has become clear that cellular interactions with the ECM also provide essential information to the cell by initiating intracellular signalling cascades which culminate in changes in gene expression. Furthermore, the ECM also provides structural cues to adherent cells. In turn, these ECM-induced changes in cell morphology and gene expression allow cells to subsequently alter their interactions with the extracellular environment. This 'dynamic reciprocity' is the basis for integrated tissue function and allows cells monitor to constantly their extracellular environment and adjust their responses to maintain differentiated tissue phenotypes. In contrast, tumorigenesis results from the loss of this dynamic interaction between cells and their ECM, and the subsequent aberrant dialogue between cells and their microenvironment prevents cells from achieving or maintaining their functional differentiated state. How cells acquire and maintain their 3D organisation is a fundamental question which largely remains to be understood. In order to understand how the various physical and biochemical properties of ECM components contribute to cellular responses, it is necessary to review the components contributing to ECM-mediated responses in cells.

ECM PROTEINS AND INTEGRIN RECEPTORS

ECM Proteins

The ECM proteins comprise a large family of glycoproteins. These large proteins are often comprised of several distinct subunits. For example, laminin, which is composed of α-, β- and γ-chains, and has a molecular mass of over 10^6 kDa, has up to three variants of each of these chains. The final composition of laminin is tissue specific. For example, laminin 5 is found in basement membranes associated with epithelia and endothelium, while laminin-2 is preferentially found in muscle (for a review, see Ekblom, 1996).

Another major component of the extracellular matrices are the collagens. This family is comprised of over 18 members, which are often expressed in tissue-specific patterns (for a review, see Prockop and Kivirikko, 1995). In addition to laminins and collagens, other major ECM components include fibronectin found in association with cells as well as in the serum, vitronectin also found in serum, and tenascin. These matrix proteins are often found in connective tissue matrices which are rich in collagens and to a lesser extent in association with basement membranes.

Another major class of ECM proteins are the proteoglycans, which, with the exception of hyaluronic acid, consist of membrane-associated globular proteins with specialized sulfated N- and O-linked carbohydrate chains called glycosaminoglycans (GAGs). The most common GAG chains found are the heparan and chondroitin chains which are found in almost all mammalian tissues (for a review, see Bernfield *et al*., 1999).

The structural and biochemical composition of many ECM proteins has been known for some time. These large proteins are usually multidomain structures capable of interacting with other matrix proteins as well as acting as a potential reservoir for soluble mitogens and morphogens (Woodley *et al*., 1983; Saskela and Rifkin, 1990; Vu *et al*., 1998). Indeed, specific matrix proteins rarely exist in isolation but instead are found as members of larger complex matrices comprised of various ratios of individual components. It must be emphasized that the net effect of these complex matrices on cell behaviour is equally complex and not simply the sum of the individual components. One fairly well-characterized type of complex ECM is the basement membrane (BM), which consists primarily of laminin-1 and -5, type IV collagen and other minor components including nidogen, fibronectin and proteoglycans (for a review, see Schwarzbauer, 1999). As the name implies, BMs are found at the basal lateral surface of most epithelial tissues including mammary and intestinal epithelium, hepatocytes, keratinocytes and endothelial cells which line the blood-vessel walls. Although laminin can comprise up to 80% of BM, cells behave very differently when in contact with laminin alone as compared with a complete BM. For example, when endothelial cells are plated on a complete reconstituted BM, they rapidly form an anastomosing network of hollow tube-like structures reminiscent of capillaries (Kubota *et al*., 1998). In contrast, when plated on laminin alone, they form a continuous 'cobblestone' monolayer of cells (Madri and Williams, 1983). Similarly, mammary epithelial cells will form 3D alveolar structures and express milk proteins when cultured on a complete BM but not when plated on laminin alone (Roskelley *et al*., 1994). Together these findings emphasize that complex ECMs are necessary to direct complex tissue-type organisation and gene expression. A recent comparative analysis of the *Drosophila* and *C. elegans* genomes indicated that these BM proteins have remained highly conserved throughout evolution from *C. elegans* to vertebrates, emphasizing the essential nature of these proteins in multi-cellular organisms (Hynes and Zhao, 2000).

More often than not, contact with a BM-type ECM leads to growth arrest and differentiation and promotes tissue-specific gene expression. In contrast to BMs, stromal matrices or provisional matrices formed at sites of injury often consist of various ratios of the fibrillar collagens, tenascin, fibronectin as well as other matrix components derived from the circulation, including vitronectin and fibrinogen. These matrices often comprise the bulk of connective tissues surrounding cells other than epithelium such as fibroblasts and chondrocytes. Adhesion to these connective-tissue type matrices often promotes cell migration and proliferation. Interestingly, increased amounts of stromal ECM are often found adjacent to many epithelial tumor cells (Weaver *et al*., 1996).

In attempting to understand how different ECM components give rise to these markedly different phenotypes, much attention has been focused on different cell-surface receptors for these various ECM components.

Integrins

Cells recognize and respond to different ECM matrices by interacting with cell-surface receptors called integrins. Integrins are a large, specialized family of transmembrane heterodimeric proteins which consist of an α subunit and an often larger β subunit with a cytoplasmic domain. These cytoplasmic domains may directly interact with cytoskeletal proteins and thereby serve to 'integrate' the extracellular and intracellular environments.

The α and β subunits can form up to 19 different combinations which have some selective but also overlapping affinities for various ECM components (see **Tables 1** and **2**). For example, the $\alpha2\beta1$ heterodimer has been shown to bind both collagen I and laminin, whereas the $\alpha6\beta1$ integrin will bind laminin only and not collagen.

Table 1 Integrin heterodimers

Beta subunits	Alpha partners
$\beta1$	$\alpha1$, $\alpha2$, $\alpha3$, $\alpha4$, $\alpha5$, $\alpha6$, $\alpha7$, $\alpha8$, $\alpha9$, αv
$\beta4$	$\alpha6$
$\beta3$, $\beta5$, $\beta6$, $\beta8$	αv
$\beta2$	αL, αM, αX
$\beta7$	$\alpha4$

Table 2 Integrins for common ECM ligands

ECM ligand	Integrin
Laminin	$\alpha1\beta1$, $\alpha2\beta1$, $\alpha3\beta1$, $\alpha6\beta1$, $\alpha7\beta1$, $\alpha6\beta4$
Collagen I	$\alpha1\beta1$, $\alpha2\beta1$, $\alpha3\beta1$,
Fibronectin	$\alpha4\beta1$, $\alpha5\beta1$, αv$\beta1$, αv$\beta3$, αv$\beta6$, αv$\beta8$, $\alpha4\beta7$
Tenascin	αv$\beta3$, αv$\beta6$, αv$\beta8$
Vitronectin	αv$\beta1$, αv$\beta3$, αv$\beta6$, αv$\beta8$

The ligand recognition sequences of the integrins are determined by the combination of α and β subunits of the heterodimeric molecules, rather than by the α or β chains alone, which may account for some of the substrate overlap. The integrin heterodimers can recognize distinct amino acid sequences present in different ECM ligands. For example, the arginine–glycine–aspartic acid (RGD) recognition domain present in many integrin heterodimers including the αv$\beta3$ and $\alpha5\beta1$ integrins can bind to any ECM component which contains an exposed RGD sequence such as fibronectin and tenascin (Pierschbacher and Ruoslahti, 1984; Joshi *et al*., 1993). Furthermore, because most ECM proteins contain multiple domains, they may also contain multiple integrin recognition sites. For example, in addition to the RGD domain, tenascin also contains fibronectin-type repeats which can bind other integrins such as $\alpha9\beta1$ in an RGD-independent manner (Yokosaki *et al*., 1994).

Like the tissue-specific distribution of some ECM components, many of the integrin subunits are often distributed in cell- and tissue-specific patterns. For example, the $\beta2$ integrin is almost exclusively expressed by leukocytes, whereas others such as the $\beta1$ subunit are widely expressed. Although it is tempting to suggest that tissue-specific responses might be determined by the composition of the adjacent ECM as well the particular species of integrin available to interact with it, this relationship is not straightforward. Many integrins exist in inactive conformations, whereby the ligand recognition domains are masked or unavailable to interact with the ECM. Integrin activation can be induced by many different stimuli and provides an additional level of control by which cells can regulate their interaction with the microenvironment (for a review, see Hughes and Pfaff, 1998). This activation phenomenon, referred to as 'inside-out signalling,' is often induced by binding of the ECM to a particular integrin, which in turn induces signalling cascades to allow other surface integrins to become activated. On the other hand, it has also been shown that activating intracellular signalling pathways can also lead to inactivation of integrins (Hughes *et al*., 1997). 'Inside-out signalling' is yet another example of the dynamic interplay between cells and their matrices and emphasizes the reciprocal nature of these interactions.

In addition, proteolytic fragments of ECM molecules appear capable of binding many of the same integrins as their full-length counterparts, yet often give rise to dramatically different cellular responses. For example, both the full-length laminin-5 and the $\alpha3$ fragment of laminin-5 which is generated via proteolytic cleavage can bind the $\alpha3\beta1$ integrin. However, in contrast to full-length laminin, the proteolysed fragments increase cellular migration (Gianelli *et al*., 1997). Thus the net balance of proteolytic activity and structural integrity of the ECM as well as the expression pattern or activation state of integrins all contribute to the coordinated tissue responses to a particular microenvironment.

Signalling by Integrins

Binding of an ECM molecule to its integrin receptor initiates a series of intracellular signalling events which led to changes in cellular behaviour. Signalling via integrins often involves the recruitment of adaptor proteins and activation of a series of cytoplasmic protein kinases. (A detailed description of tyrosine kinase signalling is provided in the chapter on Signalling by Tyrosine Kinases.) The classical integrin-mediated signalling pathways originally described involved the cytoplasmic domains of the β integrin subunit undergoing a conformational change upon ligand binding. This conformational change in turn allows focal adhesion kinase (FAK) to phosphorylate itself. Phosphorylated FAK could then bind to promote the activation of a number of different downstream mediators including members of the Src family of protein kinases, the PI3 kinase pathway, as well as recruitment of adaptor proteins including $p130^{cas}$, Crk and the Grb2–Sos complex which recognize phosphorylated intermediates via their SH2 domains. In turn, assembly of these focal adhesion complexes and phosphorylated intermediates leads to the activation of Ras and Raf. Activated Raf then phosphorylates extracellular-related kinase (ERK). Phosphorylated ERK can then translocate to the nucleus and activate cell proliferation via its ability to modulate the activity of transcription factors such as TCF necessary for expression of cell cycle mediators including cyclin D1 and c-fos (**Figure 1**).

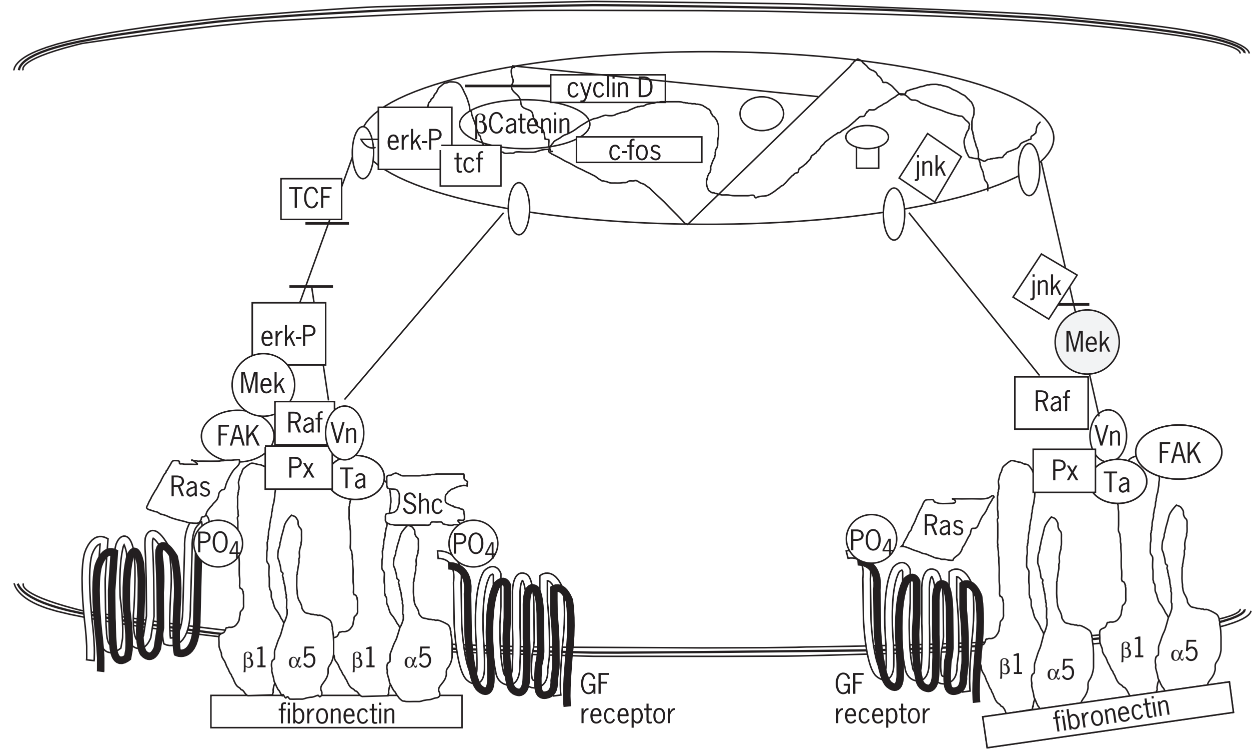


Figure 1 Intracellular signalling pathways which may participate in cell proliferation induced by adhesion to a fibronectin extracellular matrix. Integrins and growth factor receptors are clustered and form focal adhesion complexes which contain vinculin (Vn), paxillin (Px) and talin (Ta), which in turn bind to the cytoskeleton. The growth factor receptors are also phosphorylated (PO_4) and signalling initiated by the focal adhesion complexes allows ERK or JNK to become activated and phosphorylated and translocate to the nucleus where they act as transcription factors to allow expression of genes including *c-fos* or cyclin D1. β-Catenin can also move to the nucleus and form transcriptional complexes with TCF. Note that on this substratum that cells generally become flattened and elongated.

More recently, it has been shown that α integrin subunits could also specifically initiate signalling cascades. Rather than initiating signalling through phosphorylation of FAK, the α chains require the membrane-bound caveolin-1 complex to recruit the Shc adaptor molecule. Shc in turn can interact with the Grb2–Sos complex which then activates Ras and Raf and the MAPK signalling pathways to promote proliferation (Wary *et al*., 1996, 1998). Recent studies with *Drosophila* have indicated that Shc is required only for a subset of receptor tyrosine kinase activities, further suggesting that recruitment of Shc may impart some specificity to signalling cascades (Luschnig *et al*., 2000). What is not clear is how different cell types might differentially recruit signalling mediators as many of the α integrin subunits which recruit Shc are ubiquitously expressed.

Another issue which has been raised in studying ECM–integrin-mediated proliferation is the apparent lack of specificity of the signalling intermediates. For example, the signalling mediators employed by integrins to induce cell proliferation are essentially the same factors required for mitogen or cytokine-induced proliferation (Hill and Treisman, 1995). This has led to some debate as to whether integrins were simply permissive for growth factor-induced signalling rather than being capable of independently initiating these signalling cascades. More recent evidence suggests that ECM–integrin binding can independently initiate signalling responses but these signals are relatively transient (Chen *et al*., 1996). What is clear, however, is that the propagation or amplification of MAPK pathways arising from binding of soluble mitogens to their receptors is absolutely dependent upon cell adhesion to the ECM. For example, in isolated non-adherent fibroblasts, addition of the mitogen PDGF could only promote transient MAPK activation and cell cycle progression to the G1/S boundary. Further progression through the cell cycle including DNA synthesis and expression of cyclin E required that cells be attached to an ECM to sustain a threshold of MAPK activity compatible with cell cycle progression. Moreover, they noted that the shape the cells acquired upon adhesion to fibronectin, rather than adhesion *per se*, was critical for proliferative responses (Zhu and Assoian, 1995; Zhu *et al*., 1996).

Complex Matrices and Growth Arrest

Although the signalling pathways induced by the interactions of individual ECM ligands with a particular integrin receptor in homogeneous cell populations have been relatively well defined, it is not clear how signalling cascades are impacted by complex multiligand matrices interacting with multiple cell surface integrins. Whether complex interactions like these which normally occur *in vivo* give rise to qualitative, quantitative or spatially modified signalling processes has not been resolved.

Furthermore, many cells, in particular epithelial and endothelial cells, respond to complex BM ECMs by withdrawing from the cell cycle, acquiring a differentiated phenotype and initiating tissue-specific gene expression. Although MAPK activity is significantly attenuated in cells induced to growth arrest on these matrices, it is not yet clear how this is achieved as initial adhesion to these matrices via integrins also activates MAPK, albeit transiently. It may be possible that BM activates phosphatases which are capable of uncoupling the kinase-dependent activity of the MAPK pathway. For example, PTEN is a phosphatase which attenuates integrin-induced phosphorylation of FAK as well as the PI3-kinase pathway (for a review, see Tamura *et al*., 1999).

Another hypothesis that has been put forward to explain inhibition of cell proliferation by BM binding integrins including $\alpha3\beta1$, $\alpha6\beta1$ and $\alpha2\beta1$ is that they do not efficiently recruit Shc and are not capable of sustaining MAPK activity at a level which would support proliferation (Wary *et al*., 1996). However, other studies have shown that these same α integrins could not only sustain MAPK but also promote proliferation in both endothelial and epithelial cells (Aplin *et al*., 1999; Gonzales *et al*., 1999). Thus other factors which may or may not influence Shc recruitment must be involved in deciding whether cells proliferate or growth arrest using BM integrins.

Interestingly, when cells did proliferate using these same integrins, not only was it independent of Shc but, more importantly, it also appeared that the shape that the cells acquired was a dominant determinant of whether they could sustain MAPK activity and undergo proliferation (Aplin *et al*., 1999). Indeed, recent evidence suggests that the structure imposed on cells through their interaction with the ECM is responsible for modulating these signalling pathways and ultimately determines whether cells proliferate or undergo growth arrest and differentiate (Roskelley *et al*., 1994; Chen *et al*., 1997; Wang *et al*., 1998). Furthermore, evidence is emerging that ECM-induced changes in cell morphology and architecture may in fact influence the activity of phosphatases such as PTEN (Wu *et al*., 2000). The role of ECM in modulating the morphology of cells and its influence on mediating intracellular signalling is discussed in the following sections.

INTEGRATION OF CELL MORPHOLOGY AND SIGNALLING BY ECM

Integrin Clustering and Cell Shape

One of the most critical events required for propagation and maintenance of signals generated following binding of integrins by ECM is clustering of the integrin receptors, often at sites known as focal adhesions or focal contacts. Without the appropriate clustering, ligand occupation of the receptor is not sufficient to induce a full biological response (Miyamoto *et al*., 1996). The clustering is believed to facilitate interactions between the integrin cytoplasmic tails and adaptor proteins to allow focal adhesion complexes to assemble. Not surprisingly, immunoprecipitation of these FA complexes showed that the growth factor (GF) receptors are also found within these integrin-containing complexes. For example, both PDGF-BB and insulin receptors were immunoprecipitated in complexes isolated using antibodies against $\alpha v\beta3$ integrin (Schneller *et al*., 1997). The colocalization of GF receptors and integrins is believed to facilitate 'crosstalk' between ECM and GF receptors and coordinate or amplify the signals which may be independently generated by the ECM or soluble mitogens.

Clustering of integrins is not only required for integrin-induced migration or proliferation but is also essential for differentiation and tissue-specific gene expression in mammary epithelial cells on BM (Roskelley *et al*., 1994). This absolute dependence on clustering of integrins for either proliferation or differentiation underscores the importance of having structure imposed upon cytoplasmic signalling mediators.

In addition to structural changes at the level of the focal adhesion, adhesion to different ECMs can induce cells either to spread or to become rounded and polarized. It has become increasingly evident that cell-shape changes are a necessary and integral component of how cell–ECM interactions can generate tissue-specific architectures and gene expression, i.e. in order for ECM to evoke the appropriate response, cells must adopt an appropriate morphology (Roskelley *et al*., 1994). In general, cell spreading appears to be required for cells to proliferate while cell rounding is a prerequisite for growth arrest. For example, work by Ingber and colleagues has elegantly demonstrated that when endothelial cells are cultured on FN matrices they often adopt a spread morphology and proliferate. However, using micropatterned substrates which forced cells to become rounded while maintaining a similar degree of integrin mediated contact with the matrix, cells were unable to grow in the presence of mitogens (Chen *et al*., 1997). In contrast, cell rounding, which can be induced by BM-type ECM, is required for other functions such as the expression of the β-casein gene by mammary epithelial cells (Roskelley *et al*., 1994).

Binding of cells to the ECM not only can ligate and cluster integrins to initiate signalling cascades, but also provides the cells with a morphology to sustain the appropriate response.

What has not been directly established in these studies is how the cell shape impacts on intracellular signalling cascades. It is entirely possible that a cell's shape can determine whether integrins recruit signalling intermediates which interact with the growth-promoting MAPK pathways. For example, perhaps cell rounding, which generally suppresses growth, might preclude recruitment of membrane-associated mediators such as caveolin–Shc complexes and thereby attenuate proliferative signals, whereas cell spreading may support this effect. In order to understand how cellular geometry influences intracellular signalling it is necessary to understand the dynamics of integrin–cytoskeletal interactions which underlie these morphological changes. These are discussed briefly below.

Integrin–Cytoskeletal Connections

The observations that β integrin subunits extend into the cytoplasm has long generated speculation that the integrin cytoplasmic domains directly interact with cytoskeletal elements to bring about the changes in clustering and cell morphology. Functional linkages are supported by an extensive literature which shows that disrupting ECM–integrin interactions with function blocking antibodies leads to significant alterations in cell shape and cytoskeletal organisation. Similarly, directly disrupting the cytoskeleton with a variety of agents will impair integrin-mediated functions (for a review, see Schoenwaelder and Burridge, 1999).

It has now been established that integrins, in particular the $\beta1$ subunit, can directly interact with either microfilaments (actin) or intermediate filaments of the cell cytoskeleton. This is mediated by binding of proteins including α-actinin, talin, vinculin, filamen and paxillin. Many of these proteins directly interact with integrin cytoplasmic domains as well as actin filaments (for reviews, see Yamada and Geiger, 1997; Critchely, 2000). Furthermore, the ability of integrins to interact with these cytoskeletal proteins also requires that the integrins be clustered in focal adhesions (Calderwood *et al*., 1999). In addition to clustering, ligand occupancy is also required as the ability of integrins to bind these proteins is masked or suppressed in unoccupied integrins (Miyamoto *et al*., 1996).

Although integrins may directly interact with many of these cytoskeletal elements, many biological processes such as cell migration require that these interactions be dynamic. One means by which the ECM and integrins can dynamically reorganize the cytoskeleton is through activation of members of the Rac/Rho GTPase family which modulate actin cytoskeleton dynamics and interactions with integrins (for a review, see Bishop and Hall, 2000). For example, activated Rho can induce phosphatidylinositol-4,5-biphosphate (PtdInsP2), which in turn unmasks talin and actin binding sites in vinculin (Gilmore and Burridge, 1996). In addition to mediating dynamic changes in cell shape and motility, integrin-mediated changes in gene expression are also dependent upon Rho activation as induction of collagenase gene expression in fibroblasts by $\alpha5\beta1$ integrin ligation could not proceed when Rho was mutated (Kheradmand *et al*., 1998).

It is believed that these cytoskeletal rearrangements induced by ECM and integrins act to organize or compartmentalize signalling intermediates in such a way as to facilitate or enhance their interactions. Indeed, the attenuation of MAPK activity observed in suspended cells as compared with adherent cells supports this notion (Zhu and Assoian, 1995; Aplin *et al*., 1999). It has also been suggested that efficient signal transduction also requires that additional factors be recruited to stabilize these interacting protein complexes.

Nonintegrin–Cytoskeletal Linkages and Scaffolding Proteins

Analogous to the adaptor proteins mentioned earlier, scaffolding proteins can act to stabilize the relatively weak interactions between signalling mediators by binding multiple components which interact with each other, as well as with cytoskeleton components. For example, in *Drosophila*, the Ina D scaffold protein involved in photoreceptor signal transduction acts to bind and stabilize up to five partner proteins at a specific subcellular location resulting in the formation of a 'transducisome' (for a review, see Burack and Shaw, 2000). Many of the proteins which carry out these scaffolding functions contain one or more domains capable of facilitating protein–protein interactions. One particularly common domain found in many of these proteins are the PDZ domains (post-synaptic/discs large/Zo-1 domains) (for a review, see Dimitratos *et al*., 1999). In addition to facilitating interactions between proteins, many of these PDZ domain proteins have been shown to undergo extensive interactions with both the plasma membrane and the cytoskeleton, which probably contributes to their ability to localize these multiprotein complexes or receptors at specific cellular locations (Fanning *et al*., 1998; Hildebrand and Soriano, 1999).

Although a role for scaffolding proteins in mitogen-induced signalling is well established, there is as yet no direct evidence that integrin signalling pathways employ these proteins. Interestingly, there is evidence that non-integrin ECM receptors interact with several closely related proteins. Dystroglycan is a component of the dystrophin-associated protein complex found in muscle, neurons and epithelial cells. Dystroglycan not only binds to laminin in the extracellular space but also binds

dystrophin, which in turn binds actin in the cytoplasm (Ervasti and Campbell, 1993; Kachinsky *et al.*, 1999). Syndecan, a heparan sulfate proteoglycan, is another nonintegrin ECM receptor capable of binding laminin, while syndecan's cytoplasmic face can bind to CASK. CASK is a specialized type of PDZ protein or MAGUK (membrane associated guanylate kinase), which can also interact with protein 4.1 of the actin cytoskeleton to provide an additional link between the ECM and the cytoskeleton (Cohen *et al.*, 1998). Furthermore, when CASK's interaction with syndecan is disrupted, possibly by proteolytic cleavage, CASK can then shuttle directly to the nucleus and interact with T-brain, a T box transcription factor which induces expression of reelin, another ECM protein which is required for neuronal adhesion, migration and pathfinding (D'Arcangelo *et al.*, 1995; Hseuh *et al.*, 2000). Thus an intact ECM is required to prevent translocation of these factors.

Related members of this PDZ/MAGUK family of proteins have also been implicated in maintenance of cell polarity and cell–cell junctions, a process which is significantly enhanced by adhesion of cells to BM. The relationship between ECM, the establishment of cell junctions and recruitment of these and other proteins in growth arrest and differentiation is discussed in the following section.

Establishment of Polarity and Cell–Cell Junctions

When cells assume their tissue-specific 3D organisation, they also establish extensive cell–cell junctions and exhibit tissue polarity. Not surprisingly, the cytoskeletal and morphological changes induced by cells interacting with the ECM are necessary for the formation of appropriate cell junctions and polarity. Hemidesmosomes, adherens junctions and tight or occludens junctions found in epithelial tissues can be directly influenced by the ECM.

Adherens junction are multiprotein complexes present in most epithelial cells *in vivo*. These complexes are located between adjacent epithelial cells and contain, most notably, E-cadherin in an insoluble complex with α- and β-catenins (for a review, see Fuchs *et al.*, 1997). As mentioned, ECM and integrins play a critical role in assembling and establishing adherens junctions in epithelial tissues, as culturing epithelial cells on BM can promote the formation of polarized 3D spheroids with functional adherens junctions containing E-cadherin and α- and β-catenin (for reviews, see Gumbiner 1996; Weaver *et al.*, 1997). On the other hand, loss of ECM adhesion or disruption of appropriate cell–ECM contacts and 3D morphology interferes with the assembly of junctional complexes. Like most cell–ECM interactions, this relationship between integrins and adherens junctions is dynamic and reciprocal, as establishment of adherens junctions and recruitment of cadherin–catenin complexes can also feedback to down-regulate expression of the β1 integrin associated with growth in keratinocytes (Hodivala and Watt, 1994).

Although the exact mechanisms which link the ECM and adhesion junctions is not clear, both focal adhesions and adherens junctions contain many of the same molecules which interact with the cytoskeleton. For example, vinculin, which localizes to FA via its interactions with talin and β integrins, is also present in adherens junctional complexes. Vasp, a vinculin-binding protein capable of binding G actin and nucleating actin fibril assembly, is not only found in focal adhesions but also recruited to epithelial junctions along with vinculin (Vasioukhin *et al.*, 2000). The use of common cytoskeletal proteins which bind to integrin or cadherin complexes and in turn interact with the actin cytoskeleton suggests an integrated system of cytoskeletal fibres being pulled by potentially competing molecules to generate appropriate cell tension and shape.

What is clear is the requirement for ECM and functional adhesion junctions in maintaining a quiescent differentiated state in epithelial tissues. Interfering with cell-ECM interactions disrupts junctional complex assembly, and β-catenin is no longer retained in insoluble junctional complexes (Weaver *et al.*, 1997; Novak *et al.*, 1998). Free β-catenin can then migrate to the nucleus and form complexes with the transcription factor TCF/LEF, which in turn promotes expression of a number of genes which are incompatible with a quiescent differentiated phenotype including matrilysin, fibronectin and cyclin D1 and the oncogenic c-*myc* gene (He *et al.*, 1998; Crawford *et al.*, 1999; Shtutman *et al.*, 1999).

In addition to β-catenin, many of the MAGUK/PDZ-domain protein family members also figure prominently in the formation and maintenance of junctional complexes in polarized epithelial tissues. Elegant work on *Drosophila* has identified genes critical for epithelial polarization including the PDZ domain protein *scribble*, and a related MAGUK protein called discs large lethal (*dll*). When these genes were mutated, epithelial cells exhibited a loss of polarity and became round and multilayered, further emphasizing that tissue structure is essential to prevent tumorigenesis (Bilder *et al.*, 2000). Although the mammalian homologue of scribble is not yet known, the mammalian homologue of dll, DLG, has been identified as a protein which binds to APC, a junction-associated tumour-suppresser protein which also binds β-catenin to prevent its translocation to the nucleus (Matsumine *et al.*, 1996). Recent evidence has shown that the activity of another tumour suppressor, PTEN, a phosphatase which attenuates integrin-mediated FAK and PI3 kinase activity, could be enhanced by binding to MAGI-2, a MAGUK scaffolding protein related to DLG. Like dll and scribble, MAGI-2 is normally recruited to and anchored in membranes at epithelial tight junctions where it binds PTEN via its PDZ domains (Wu *et al.*,

2000). Although a direct link between the ECM and these particular junction-associated proteins has not been demonstrated, the ability of non-integrin ECM receptors to bind closely related proteins and the critical role of the ECM in organizing the cytoskeleton and cell junction assembly suggest that such interactions are likely.

ECM and integrins have also been directly linked to the formation of hemidesmosomes found on the basolateral surface of epithelial cells which contact the BM. The $\alpha6\beta4$ laminin-binding integrin is a major component of hemidesmosomes (Sonnenberg *et al*., 1991). In contrast to other β integrins, $\beta4$ has an unusually long cytoplasmic tail over 1000 amino acids in length. The $\beta4$ tail can directly bind to plectin, a large cytoskeletal protein capable of interacting with both the actin microfilament and intermediate filament cytoskeleton (for a review, see Steinbock and Wiche, 1999). In contrast to other integrins which require activation and clustering to interact with proteins linked to the cytoskeleton, the $\beta4$ integrin cytoplasmic tail appears capable of interacting with hemidesmosome components even in the absence of ligand occupancy (Nievers *et al*., 2000). A critical role for the $\beta4$ integrin in the formation and function of hemidesmosomes is evident in $\beta4$ integrin-deficient mutants, where blistering of the skin and epidermal detachment were observed due to the lack of hemidesmosome formation (Dowling *et al*., 1996; Van der Neut *et al*., 1996) Furthermore, interfering with basolateral localization of $\alpha6\beta4$ in breast epithelial cells not only disrupts hemidesmosome formation but also induces these cells to undergo apoptosis (V. M. Weaver and M. J. Bissell, unpublished work).

Together a picture emerges that beyond simple ligation of integrins and induction of signalling cascades, the ECM also directs the changes in cell and tissue architecture via cytoskeletal linkages from both integrin and nonintegrin ECM receptors. The cytoskeletal reorganisation helps to organize hemidesmosomes and adherens junctions and immobilize multidomain scaffolding proteins. This sequestration can either prevent untoward transcriptional activity or enhance the function of growth-suppressing genes such as *PTEN* or *APC*.

3D ORGANISATION AND INTEGRATED TISSUE RESPONSES

Mammary Gland

One model which has been invaluable for investigating and understanding how interactions between cells and complex ECMs can direct and maintain the functionally differentiated state of tissues is the adult mammary gland. At the onset of pregnancy, the adult mammary gland undergoes a series of morphological and functional changes which culminate in the establishment of organized acinar structures consisting of polarized epithelial cells which produce and secrete milk proteins. This organogenesis is intimately dependent upon the production and deposition of an intact basement membrane.

Many of the cellular and molecular events which contribute to this functional differentiation have been elucidated through the use of a 3D tissue culture model which mimics the normal postnatal mammary gland morphogenesis which occurs *in vivo* (Barcellos-Hoff *et al*., 1989). Either primary or immortalized mammary epithelial cells can be induced to undergo this morphological and functional differentiation by simply plating them on a reconstituted intact BM. The mammary epithelial cells immediately adhere to the BM and within a few days form 3D spheroids which resemble the acinar structure of the gland *in vivo* both in organisation and in size. Contact with the BM also induces cells to exit the cell cycle, become polarized and form adherens junctions. In contrast, when the same cells are plated on tissue culture plastic without the BM matrix, they adhere primarily to non-BM proteins which are present in the serum including fibronectin and vitronectin using the appropriate integrins. These cells will spread to form monolayers and are not able to form spheroids. These cells also cannot polarize, form proper adherens junctions or exit the cell cycle. Not surprisingly, these cells cannot express the tissue-specific milk protein genes (Barcellos-Hoff *et al*., 1989; Schmidhauser *et al*., 1990).

The BM-induced functional differentiation requires the contribution of several different laminin-binding receptors at different stages of this process. Initially the epithelial cells adhere to the BM via laminin-binding $\beta1$ integrins, most likely the $\alpha3$ or $\alpha6$ $\beta1$ heterodimers. Following the initial adhesion, cells begin to round in a manner which is not perturbed when either $\alpha6\beta4$ or other $\beta1$ integrins are blocked. Instead, the E3 fragment of laminin, which probably binds to the nonintegrin ECM receptors dystroglycan or syndecan, appears necessary for this process (J. Muschler and M. J. Bissell, unpublished work). Furthermore, although the E3-dependent rounding is a prerequisite for the cells to begin expressing milk proteins (Streuli *et al*., 1995), milk gene expression in rounded cells could also be blocked by antibodies against both the $\alpha6\beta4$ integrin and $\beta1$ integrins but not $\alpha1$, $\alpha5$ or αv integrins (Muschler *et al*., 1999). This 'division of labour' between both integrin and nonintegrin receptors and the dependence on cell shape emphasize the integrated nature of the response of cells to BM.

In addition to the role for ECM-binding proteins, the organisation of the BM and its presentation to the cells is also critical in eliciting these responses. As mentioned previously, when the BM is cross-linked or fixed so that it is no longer malleable, cells are unable to become rounded and acquire a polarized morphology and cannot express milk proteins despite the cells' ability to adhere strongly to this matrix (Streuli and Bissell, 1990). Furthermore, although the minor BM component nidogen alone is not capable of inducing milk gene expression in these cells,

addition of nidogen can enhance the laminin-dependent induction of milk gene expression (Pujuguet *et al*., 2000). Nidogen normally acts as a bridge between BM laminin and type IV collagen, and thus probably acts to organize the presentation of laminin to the cells (see **Figure 2**).

Furthermore, once these cells have undergone the necessary changes in cell shape, they must also undergo further changes in cell–cell organisation to form 3D alveolar structures, complete with functional adherens junctions and hemidesmosomes. The generation of these 3D structures is particularly critical for maintaining a functional differentiated state, as interfering with this organisation can induce the cells to undergo apoptosis (Boudreau *et al*., 1996). When this organisation is disrupted, the cells are not able to remain quiescent and begin to progress through the cell cycle, a process which often triggers apoptosis in normally quiescent cell (Evan *et al*., 1992). Furthermore, the 3D organisation is also essential for meaningful cross-talk between integrins and growth factor receptors (Wang *et al*., 1998).

Consequences of Disrupted Tissue Organisation

Based on the above, it is clear that 3D tissue-type morphology and the related intracellular organisation imposed by the ECM determine whether cells will differentiate, proliferate or undergo apoptosis. With respect to differentiation, it is clear that in addition to cellular rounding, establishing polarity and forming junctional complexes are equally critical for maintaining this state. It might be predicted

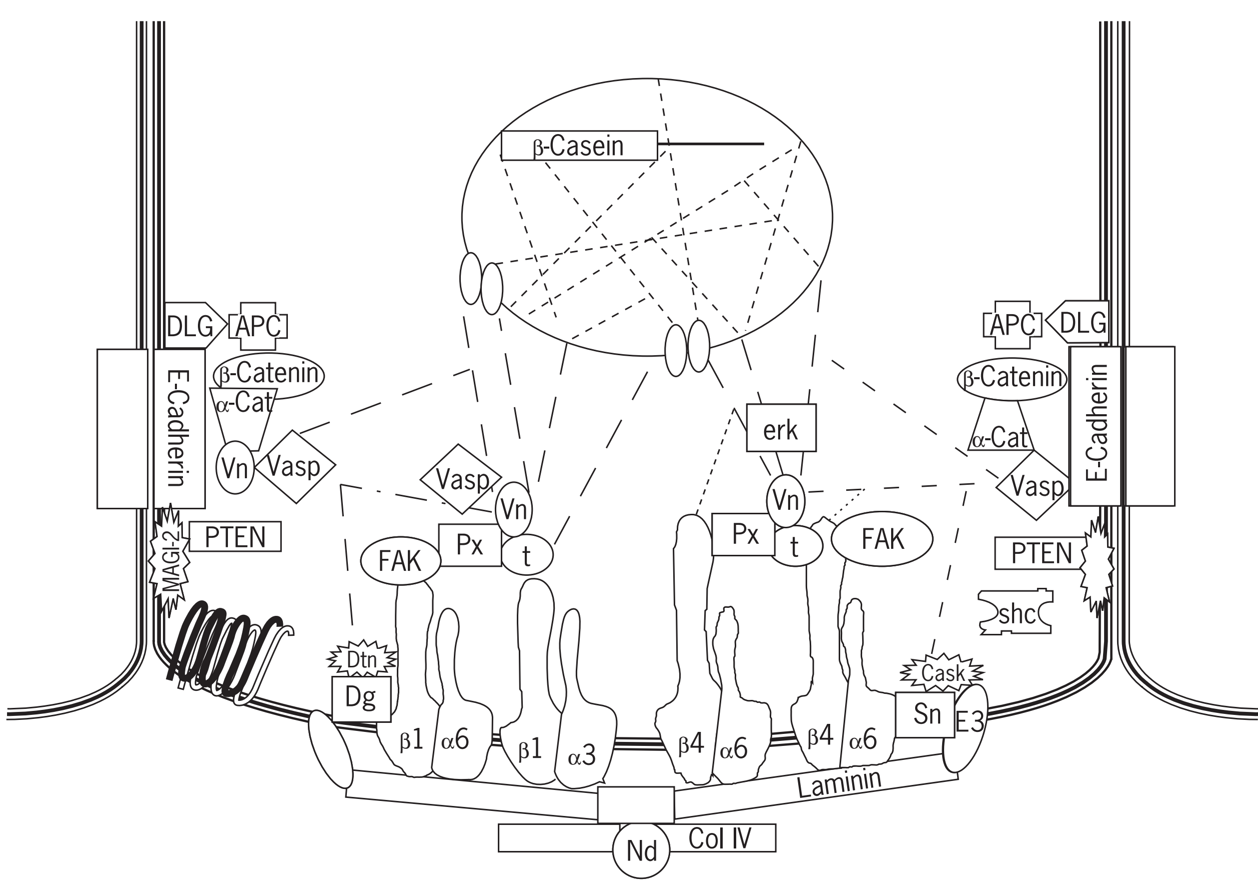


Figure 2 Intracellular organisation of polarized, quiescent epithelial cells in response to basement membrane extracellular matrix. The basement membrane matrix on the basolateral surface of the cells is comprised of several components including laminin, type IV collagen (Col IV) and nidogen (Nd). Laminin binds to $\alpha6\beta4$, $\alpha6\beta1$ and $\alpha3\beta1$ integrins and generates signals which allow transcription of the tissue-specific gene β-casein. Cells also employ non-integrin receptors including dystroglycan (Dg) and syndecan (Sn) to bind to the E3 domain of laminin which help cells to become rounded. Cask and dystrophin (Dtn), which associate with Sn and Dg receptively, can also directly bind the actin cytoskeleton (dashed lines). Additional cytoskeletal links are provided by vinculin (vn) and vasp which associate both with integrin complexes and with adherens junctions which form in the polarizing cells. The adherens junctions contain E-cadherin complexed with β-catenin and α-catenin. Other junction-associated proteins include MAGI-2 and DLG (discs large lethal), which in turn bind the tumour-suppressor proteins PTEN and APC to enhance their activity.

then that interfering with any component of this integrated 3D system not only compromises tissue function but also contributes to the process of deregulated growth control and ultimately tumorigenesis. Perhaps the most informative examples of how critical the ECM and an organized 3D structure are for maintaining differentiated tissue function and preventing tumour development and progression come from studies with tumour cells themselves which lack this organisation.

One system which very clearly demonstrates the strong correlation between loss of tissue organisation and tumour progression is a series of human breast epithelial cells which become increasingly tumorigenic as tissue architecture is progressively lost. The HMT-3522 cell series originated from a purified breast epithelial cell population isolated from a female with fibrocystic breast disease. Although these cells, referred to as S1, can remain relatively stable for over 500 passages, cells could alternatively be induced to undergo tumorigenic progression by selecting cells which would survive in the absence of essential growth factors. These factor-independent cells, designated T4-2, were found ultimately to form tumours when injected into nude mice, while the original S1 population which was maintained at a similar passage number but in defined media containing growth factors were stable and non-tumorigenic (Briand *et al*., 1987). More importantly, it was observed that whereas S1 cells maintained the ability to form organized, polarized alveolar structures which could undergo growth arrest in response to BM, the tumorigenic derivatives formed progressively disorganized and nonpolarized groups of cells which failed to growth arrest despite the presence of an intact BM. The disorganized T4-2 cells were found to have poorly organized nuclei, adherens junctions, f-actin and aberrant integrin localization and expression (Weaver *et al*., 1997) (see **Figure 3**). In fact, the extent to which epithelial cells respond to BM and elaborate organized structures is a relatively accurate means of predicting their degree of tumorigenicity with the disorganized cells being more tumorigenic (Petersen *et al*., 1992).

Aberrations in ECM–Integrins in Cancer

The disruption of normal architecture and morphology in the progressively tumorigenic cells described above could be directly related to the improper expression of high

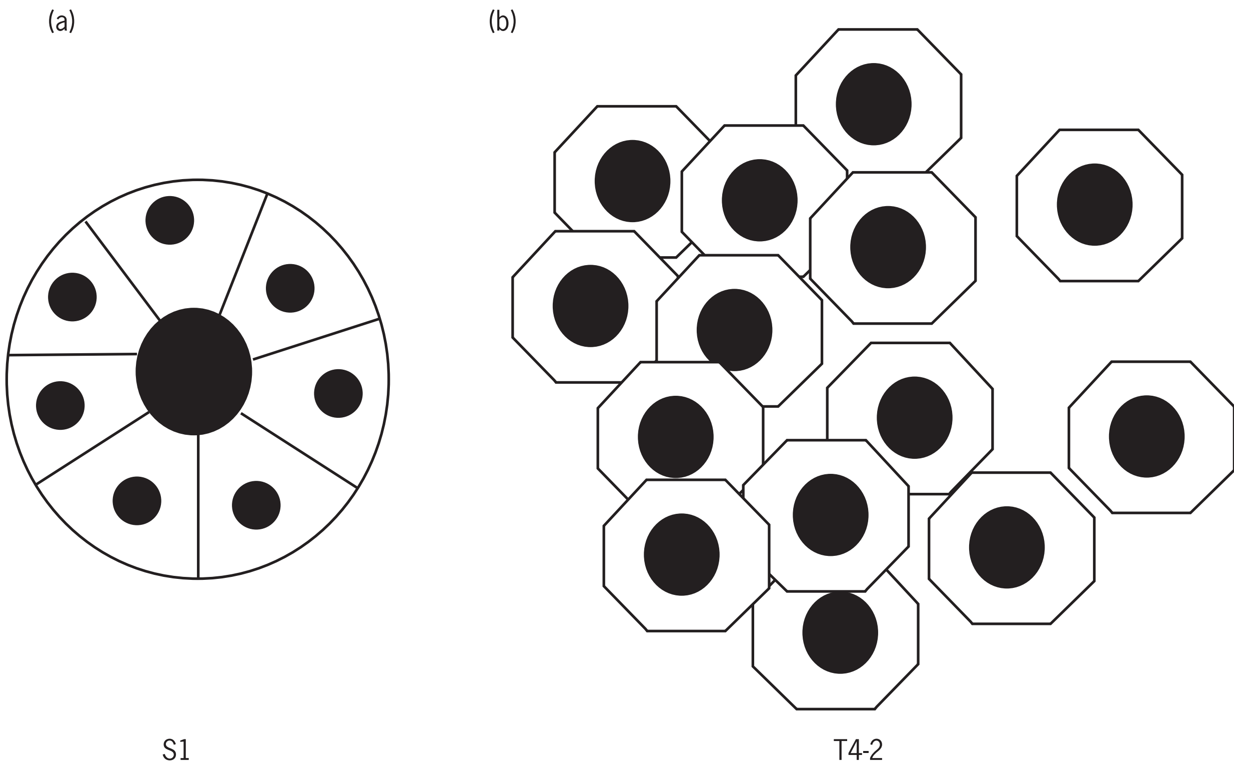


Figure 3 Cell morphology and tumorigenicity. (a) Schematic representation of normal, growth-arrested, breast epithelial cells (S1) organized into a acinar structures when cultured in a 3D basement membrane extracellular matrix. Note the basolateral localization of the nucleus in these polarized cells. (b) Tumorigenic derivatives of the normal breast epithelial cells (T4-2) form disorganized clusters when cultured in basement membrane extracellular matrix. The cells do not become polarized or form proper adherens junctions and continue to proliferate. (Adapted from Bissell *et al*., 1999.)

levels of a number of surface receptors including $\beta1$ integrin. In addition, these cells display an aberration in surface expression of $\alpha6\beta4$ integrin which would otherwise normally encourage growth arrest. Together these defects might be expected not only to disrupt the formation of hemidesmsomes, but also to impair the establishment of adherens junctions and recruitment of β-catenin (Weaver *et al*., 1997).

An extensive literature documents a wide range of abnormal cell–ECM interactions in tumour cells. Aberrant integrin expression has been described in a number of breast tumour cells (see review by Zutter *et al*., 1998). In addition to overexpressing integrins, many tumour cells also lack expression of integrin receptors which might suppress growth, and studies aimed at reintroducing the missing integrins can often result in restoring growth arrest. For example, in tumorigenic intestinal epithelial cells, which lack expression of $\beta4$ integrin, transfection of the $\beta4$ integrin induced expression of the Cdk inhibitor $p21^{WAF1}$ and led to growth arrest (Clarke *et al*., 1995). Similarly, introduction of the $\alpha2\beta1$ collagen/laminin receptor missing from a poorly differentiated breast tumour line also induced growth arrest (Zutter *et al*., 1995).

Although this approach may compensate for a lack of a functional integrin, often integrins are expressed but not properly localized in the tumour cells. In the T4-2 cell series described above, $\beta4$ integrin was expressed but not at the cell surface (Weaver *et al*., 1997) (**Figure 4**). It might be expected that integrins localized to different cellular locations may not be able to access their ligands, particularly if the ECM ligands have a restricted distribution, such as BM laminin. In addition, changing the localization of an integrin often results in the receptor adopting a different conformation, which might also prevent binding even if the ligand is accessible (Bishop *et al*., 1998).

In addition to deficiencies in ECM receptors, tumour cells often have altered ECMs. For example, many tumour cells are not able properly to synthesize or assemble BMs, which in turn are required for growth arrest (Petersen *et al*., 1992; Howlett *et al*., 1994). One of the more common reasons for a loss of organized ECM surrounding tumour cells is the finding that many tumour cells express high

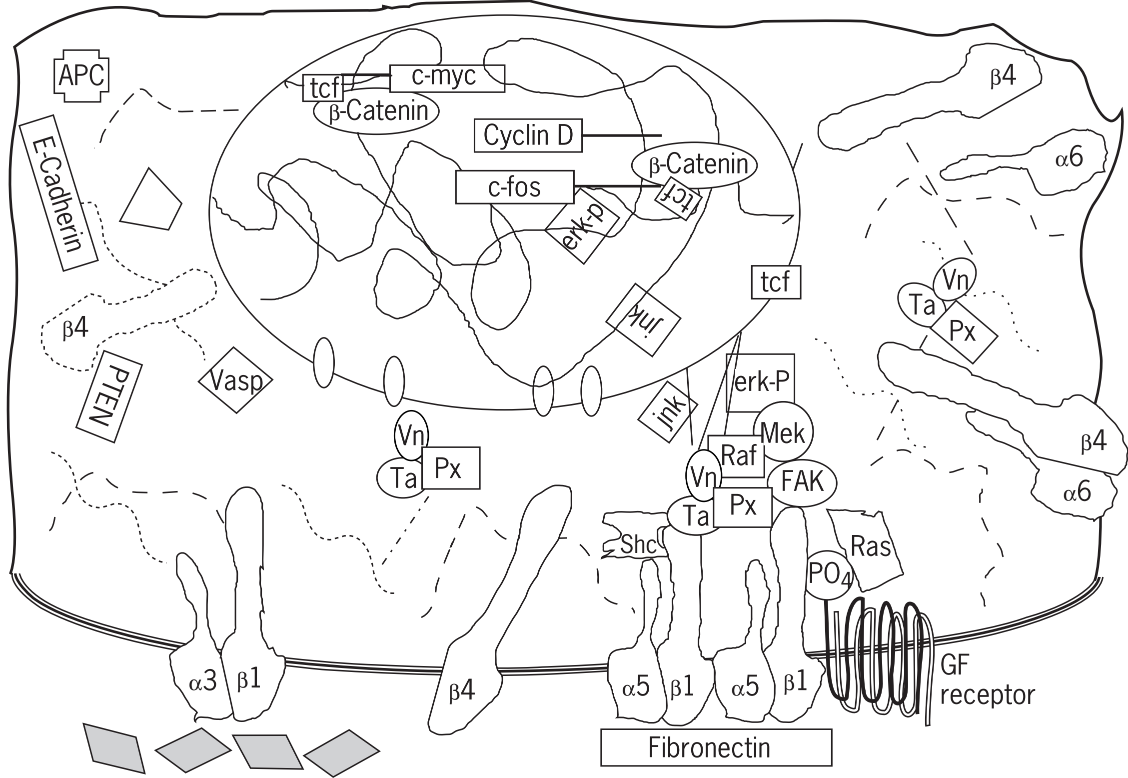


Figure 4 Structural disorganisation of tumour cells. Expression and subcellular localization of integrin subunits including $\alpha6$ and $\beta4$ is disrupted in tumorigenic epithelial cells. In addition the basement membrane may be degraded (grey diamonds) by high levels of matrix-degrading protease activity. Structural connections with the cytoskeleton (dashed lines) are also diminished. Polarity is absent and adherens junctions are not properly formed, allowing free β-catenin to translocate to the nucleus and activate transcription of growth-promoting genes including c-*myc*. Growth-promoting $\beta1$ integrin heterodimers ($\alpha5\beta1$ is shown here) are present and continue to generate proliferative signals including phosphorylated erk (erk-P), which may translocate to the nucleus and activate additional growth-promoting genes. Growth factor receptors are also active and phosphorylated.

levels of various matrix-degrading proteinases (Barsky *et al*., 1983; Liotta and Stetler-Stevenson, 1990).

More recent evidence has shown that even in normal cells excessive and sustained proteolytic activity and BM degradation can directly contribute to tumour development and progression. Loss of BM by excessive proteolysis resulted in the transformation of epithelial cells to a mesenchymal phenotype (Lochter *et al*., 1997). The loss of the syndecan which also binds BM laminin also resulted in normal epithelial cells undergoing a similar transition to an anchorage-independent mesenchymal phenotype (Kato *et al*., 1995). These results together emphasize how BM not only promotes but also maintains and stabilizes the epithelial cell phenotype. Furthermore, in transgenic mice engineered to overexpress MMP-3 in differentiated mammary epithelial cells, an unusually high incidence of adenocarcinoma was observed in the BM-compromised mammary epithelial cells (Sternlicht *et al*., 1999). The chronic degradation of the BM in turn was permissive for a number of other genetic alterations which contributed to the tumorigenic phenotype. These findings again emphasize the requirement for BM in maintaining functionally differentiated tissues and for stabilizing the genetic information in the cells.

Apoptosis

It should be emphasized that loss of normal cell–ECM interaction does not immediately result in oncogenic transformation of most cells. The tumours which arise in the MMP-overproducing mice described above require chronic or sustained activity of the protease and loss of basement membrane before overt tumorigenesis is observed. As mentioned earlier, normal mammary epithelial cells initially respond to the loss or degradation of their ECM by undergoing programmed cell death or apoptosis within 48 h (Boudreau *et al*., 1995). This is viewed as a protective mechanism to help eliminate potentially dangerous cells which have lost the growth suppressive signals normally supplied by the BM.

Although complete loss of adhesive contact with any type of ECM can induce apoptosis in normally adhesive cells (Meredith *et al*., 1993; Frisch and Francis, 1994), the ability of a cell to respond to more subtle changes in their microenvironment provides a more effective means to eliminate potentially harmful cells. In fact, normal adherent epithelial cells can be induced to undergo apoptosis by simply disrupting their three-dimensional organisation (Boudreau *et al*., 1996). Even in polarized MDCK epithelial cells, a collagen gel overlay was found to induce apoptosis by simply disorientating cells (Tang *et al*., 1998). Unfortunately, many tumour cells that have compromised interactions with the ECM fail to recognize alterations in cell morphology and orientation, and are often resistant to apoptosis, even in the complete absence of adhesion, a phenomenon referred to as anchorage independence. Attempts to restore their sensitivity to apoptotic signals is a major anti-tumorigenic strategy (see the chapter on *Apoptosis*).

It is interesting that not all normal cells immediately respond to loss or degradation of their BM by undergoing apoptosis, nor do they proliferate uncontrollably. In cells which retain a high regenerative potential, such as vascular endothelium and hepatocytes, the loss of BM allows cells to re-enter the cell cycle and divide. Loss of BM will induce endothelial cells to undergo angiogenesis, while hepatocytes can be induced to proliferate and regenerate portions of the liver following a partial hepatectomy (Fausto, 2000). In these cases, however, the amount or degree of cell proliferation which occurs is limited because proliferating endothelial cells or hepatocytes immediately begin to resynthesize and deposit a new BM, which in turn induces the cells to re-establish a 3D tissue organisation and subsequently withdraw from the cell cycle (Kubota *et al*., 1988; Boudreau *et al*., 1997). Therefore, cells which do not undergo apoptosis when they re-enter the cell cycle can be directed to growth arrest and differentiate by re-establishing and responding to a functional BM.

Restoring a Cell's Balance of Surface Receptors and Interaction with the BM Can Reverse Tumorigenesis

Although attempting to restore a normal dialogue with the ECM might at first seem somewhat naive to apply to genetically destabilized tumour cells, in fact re-establishing a functional BM and 3D organisation, or restoring a tumour cell's ability to recognize and respond to BM, can in fact induce growth arrest in tumorigenic cells. Transfection of the tumorigenic breast epithelial cell line MDA-435 with the NM-23 tumour suppressor resulted in the ability of the cells to resynthesize and deposit a 3D BM and to form organized, acinar structures leading to growth arrest (Howlett *et al*., 1994). In the case of the progressively tumorigenic human breast cells mentioned earlier, blocking the excessive β1 integrin signalling allowed cells to revert completely to a nontumorigenic phenotype when they were cultured in the presence of a complete basement membrane. This reversion was accompanied by reorganisation into acinar-type structures, establishment of normal adherens junctions containing E-cadherin and associated α- and β-catenins, reorganisation of the actin cytoskeleton and growth arrest. Furthermore, blocking β1 integrin also attenuated the high levels of expression and activity of the EGF receptor normally found in these tumorigenic cells. Interestingly, simply blocking β1 integrin was not sufficient to downregulate EGF receptor expression or activity if the cells were also prevented from acquiring the corresponding 3D organisation, underscoring the contribution of the basement membrane to this reversion process (Wang

et al., 1998). Perhaps the most striking observation made using this system was that when $\beta 1$ integrin was blocked and cells were allowed to resume their normal interaction with the basement membrane, the resulting 3D structure was sufficient to override a host of genetic alterations that had accumulated as cells progressed toward increasing tumorigenicity (Weaver *et al*., 1997).

CONCLUSION

Simply stated, cancer is a problem of tissue organisation. Although it has long been recognized that one of the hallmarks in diagnosing tumour cells is their aberrant morphology, both at the tissue, cell and nuclear level, the molecular consequences of this disorganisation are now beginning to be appreciated. By maintaining a proper 3D organisation through dynamic reciprocal interactions with their microenvironment, tissue structure can act in a dominant manner to override a host of genetic aberrations that may otherwise compromise normal tissue function.

A recent large-scale study emphasized an environmental rather than genetic basis of cancer based on findings that the incidence of identical twins developing similar cancers was almost nondetectable, whereas genetically unrelated individuals exposed to similar environments were more likely to do so (Lichtenstein *et al*., 2000). Although these observations apply to the relationship of an entire organism to its environment, it would appear that the same principles may also apply to tumour development at the cell and molecular level. Thus although cells may harbour genetic alterations from birth or acquire distinct genetic changes over a life span, it is how the cells interact with their immediate extracellular environment which appears to govern cell form, function and fate.

ACKNOWLEDGEMENTS

This work was supported by the Director, Office of Biological and Environmental Research (contract DE-AC03-76SF00098) of the U.S. Department of Energy and by NIH grants CA-57621 and CA-64786 to MJB, and by NIH grant CA-85249 to NB.

REFERENCES

Aplin, A. E., *et al*. (1999). Anchorage-dependent regulation of the Mitogen-activated protein kinase cascade by growth factors is supported by a variety of integrin alpha chains. *Journal of Biological Chemistry*, **274**, 31223–31228.

Barcellos-Hoff, M. H., *et al*. (1989). Functional differentiation and alveolar morphogenesis of primary mammary cultures on reconstituted basement membrane. *Development*, **105**, 223–235.

Barsky, S. H., *et al*. (1983). Loss of basement membrane components by invasive tumors but not by their benign counterparts. *Laboratory Investigation*, **49**, 140–147.

Bernfield, M., *et al*. (1999). Functions of cell surface heparan sulfate proteoglycans. *Annual Review of Biochemistry*, **68**, 729–777.

Bilder, D., *et al*. (2000). Cooperative regulation of cell polarity and growth by drosophila tumor suppressors. *Science*, **289**, 113–116.

Bishop, A. L. and Hall, A. (2000). Rho GTPases and their effector proteins. *Biochemical Journal*, **348**, 241–255.

Bishop, L. A., *et al*. (1998). Lack of intrinsic polarity in the ligand-binding ability of keratinocyte beta1 integrins. *Experimental Dermatology*, **7**, 350–361.

Bissell, M. J., *et al*. (1999). Tissue structure, nuclear organization and gene expression in normal and malignant breast. *Cancer Research (Supplement)*, **59**, 1757–1764.

Boudreau, N. and Bissell, M. J. (1996). Regulation of gene expression by the extracellular matrix. In: Comper, W. E. (ed.), Extracellular Matrix. Volume 2. Molecular Components and Interactions, 246–261 (Harwood Academic Publishers: Amsterdam).

Boudreau, N., *et al*. (1995). Suppression of ICE and apoptosis in mammary epithelial cells. *Science*, **267**, 891–893.

Boudreau, N., *et al*. (1996). Suppression of apoptosis by basement membrane requires three-dimensional tissue organization and withdrawal from the cell cycle. *Proceedings of the National Academy of Sciences of the USA*, **93**, 3509–3513.

Boudreau, N., *et al*. (1997). Hox D3 induces an angiogenic phenotype in endothelial cells. *Journal of Cell Biology*, **139**, 257–264.

Briand, P., *et al*. (1987). A new diploid non-tumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium. *In Vitro Cell Developmental Biology*, **23**, 181–188.

Burack, W. R. and Shaw, A. S. (2000). Signal transduction; hanging on a scaffold. *Current Opinion in Cell Biology*, **12**, 211–216.

Calderwood, D. A., *et al*. (1999). The talin head domain binds to integrin beta subunit cytoplasmic tails and regulates integrin activation. *Journal of Biological Chemistry*, **274**, 28071–28074.

Chen, Q., *et al*. (1996). Integrin-mediated activation of MEK and mitogen-activated protein kinase is independent of Ras. *Journal of Biological Chemistry*, **271**, 18122–18127.

Chen, C. S., *et al*. (1997). Geometric control of cell life and death. *Science*, **276**, 1425–1428.

Clarke, A. S., *et al*. (1995). Activation of the p21 pathway of growth arrest and apoptosis by the $\beta 4$ integrin cytoplasmic domain. *Journal of Biological Chemistry*, **270**, 22673–22676.

Cohen, A. R., *et al*. (1998). Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *Journal of Cell Biology*, **142**, 129–138.

Crawford, H. C., *et al*. (1999). The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. *Oncogene*, **18**, 2883–2891.

Critchley, D. R. (2000). Focal adhesions – The cytoskeletal connection. *Current Opinion in Cell Biology*, **12**, 133–139.

D'Arcangelo, G., *et al*. (1995). A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. *Nature*, **376**, 719–723.

Dimitratos, S. D., *et al*. (1999). Signaling pathways are focused at specialized regions of the plasma membrane by scafolding proteins of the MAGUK family. *Bioessays*, **21**, 912–921.

Dowling, J., *et al*. (1996). Beta 4 integrin is required for hemidesmosome formation, cell adhesion and cell survival. *Journal of Cell Biology*, **134**, 559–572.

Ekblom, P. (1996). Receptors for laminins during epithelial morphogenesis. *Current Opinion in Cell Biology*, **8**, 700–706.

Ervasti, J. M., and Campbell, K. P. (1993). A role for the dystrophin–glycoprotein complex as a transmembrane linker between laminin and actin. *Journal of Cell Biology*, **122**, 809–823.

Evan, G. I., *et al*. (1992). Induction of apoptosis in fibroblasts by c-myc protein. *Cell*, **69**, 119–128.

Fanning, A. S., *et al*. (1998). The tight junction protein ZO-1 establishes a link between the transmembrane protein occludin and the actin cytoskeleton. *Journal of Biological Chemistry*, **273**, 29745–29753.

Fausto, N. (2000). Liver regeneration. *Journal of Hepatology,* **32**, 19–31.

Frisch, S. M. and Francis, H. (1994). Disruption of epithelial cell–matrix interactions induces apoptosis. *Journal of Cell Biology*, **124**, 619–626.

Fuchs, E., *et al*. (1997). Integrators of epidermal growth and differentiation; distinct functions for β1 and β4 integrins. *Current Opinion in Genetics and Development*, **7**, 672–682.

Gianelli, G., *et al*. (1997). Induction of cell migration by matrix metalloproteinase-2 cleavage of laminin-5. *Science*, **277**, 225–228.

Gilmore, A. P. and Burridge, K. (1996). Regulation of vinculin binding to talin and actin by phosphatidyl-inositol-4,5-bisphosphate. *Nature*, **381**, 531–535.

Gonzales, M., *et al*. (1999). A cell signal pathway involving laminin-5, alpha 3 beta 1 integrin and mitogen activated protein kinase can regulate epithelial cell proliferation. *Molecular Biology of the Cell*, **10**, 259–270.

Gumbiner, B. M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell*, **84**, 345–357.

He, T. C., *et al*. (1998). Identification of c-MYC as a target of the APC pathway. *Science*, **281**, 1509–1512.

Hildebrand, J. D. and Soriano, P. (1999). Shroom, a PDZ domain-containing actin-binding protein, is required for neural tube morphogenesis in mice. *Cell*, **99**, 485–497.

Hill, C. S. and Treisman, R. (1995). Transcriptional regulation by extracellular signals: mechanisms and specificity. *Cell*, **80**, 199–211.

Hodivala, K. J. and Watt, F. M. (1994). Evidence that cadherins play a role in the downregulation of integrin expression that occurs during keratinocyte terminal differentiation. *Journal of Cell Biology*, **124**, 589–600.

Howlett, A. R., *et al*. (1994). A novel function for the nm-23H1 gene; overexpression in human breast carcinoma cells leads to the formation of basement membranes and growth arrest. *Journal of National Cancer Institute*, **86**, 1838–1844.

Hseuh, Y.-P., *et al*. (2000). Nuclear translocation and transcription regulation by the membrane associated guanylate cycles kinase CASK/LIN-2. *Nature*, **404**, 298–302.

Hughes, P. E. and Pfaff, M. (1998). Integrin affinity modulation. *Trends in Cell Biology*, **8**, 359–364.

Hughes, P. E., *et al*. (1997). Suppression of integrin activation; a novel function of a Ras/Raf initiated MAP-kinase pathway. *Cell*, **88**, 521–530.

Hynes, R. O. and Zhao, Q. (2000). The evolution of cell adhesion. *Journal of Cell Biology*, **150**, F89–F86.

Joshi, P., *et al*. (1993). Endothelial cells adhere to the RGD domain and the fibrinogen-like terminal knob of tenascin. *Journal of Cell Science*, **106**, 389–400.

Kachinsky, A. M., *et al*. (1999). A PDZ-containing scaffold related to the dystophin complex at the basolateral membrane of epithelial cells. *Journal of Cell Biology*, **145**, 391–402.

Kato, M., *et al*. (1995). Loss of cell surface syndecan-1 causes epithelia to transform into anchorage-independent mesenchyme-like cells. *Molecular Biology of the Cell*, **6**, 559–576.

Kheradmand, F., *et al*. (1998). Role of Rac1 and oxygen radicals in collagenase-1 expression induced by cell shape change. *Science*, **280**, 898–902.

Kubota, Y., *et al*. (1988). Role of laminin and basement membrane in the morphological differentiation of human endothelial cells into capillary-like structures. *Journal of Cell Biology*, **107**, 1589–1598.

Lichtenstein, P., *et al*. (2000). Environmental and heritable factors in the causation of cancer – Analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine*, **343**, 78–85.

Liotta, L. A. and Stetler-Stevenson, W. G. (1990). Metalloproteinases and cancer invasion. *Seminars in Cancer Biology*, **1**, 99–106.

Lochter, A. *et al*. (1997). Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells. *Journal of Cell Biology*, **139**, 1861–1872.

Luschnig, S., *et al*. (2000). The *Drosophila* SHC adaptor protein is required for signalling by a subset of receptor tyrosine kinases. *Molecular Cell*, **5**, 231–241.

Madri, J. A. and Williams, S. K. (1983). Capillary endothelial cell culture: phenotypic modulation by matrix components. *Journal of Cell Biology*, **97**, 153–165.

Matsumine, A., *et al*. (1996). Binding of APC to the human homolog of the *Drosophila* discs large tumor suppressor protein. *Science*, **272**, 1020–1023.

Meredith, J. E., Jr, *et al*. (1993). The extracellular matrix as a cell survival factor. *Molecular Biology of the Cell*, **4**, 953–961.

Miyamoto, S., *et al*. (1996). Integrins can collaborate with growth factors for phosphorylation of receptor tyrosine kinases and MAP kinase activation: roles of integrin aggregation and occu-pancy of receptors. *Journal of Cell Biology*, **135**, 1633–1642.

Muschler, J., *et al*. (1999). Division of labor among the alpha6-beta4 integrin, beta1 integrins, and an E3 laminin receptor to signal morphogenesis and beta-casein expression in mammary epithelial cells. *Molecular Biology of the Cell*, **10**, 2817–2828.

Nievers, M. G., *et al*. (2000). Formation of hemidesmosome-like structures in the absence of ligand binding by the (alpha)6 (beta)4 integrin requires binding of HD1/plectin to the cytoplasmic domain of the (beta)4 integrin subunit. *Journal of Cell Science*, **113**, 963–973.

Novak, A., *et al*. (1998). Cell adhesion and the integrin-linked kinase regulate the LEF-1 and beta-catenin signalling pathways. *Proceedings of the National Academy of Sciences of the USA*, **95**, 4374–4379.

Petersen, O. W., *et al*. (1992). Interaction with the basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. *Proceedings of the National Academy of Sciences of the USA*, **89**, 9064–9068.

Pierschbacher, M. D. and Ruoslahti, E. (1984). Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature*, **309**, 30–33.

Poznansky, M. C., *et al*. (2000). Efficient generation of human T cells from a tissue-engineered thymic organoid. *Nature Biotechnology*, **18**, 729–734.

Prockop, D. J. and Kivirikko, K. I. (1995). Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry*, **64**, 403–434.

Pujuguet, P., *et al*. (2000). Nidogen-1 regulates laminin-1-dependent mammary-specific gene expression. *Journal of Cell Science*, **113**, 849–858.

Roskelley, C. D., *et al*. (1994). Extracellular matrix-dependent tissue-specific gene expression in mammary epithelial cells requires both physical and biochemical signal transduction. *Proceedings of the National Academy of Sciences of the USA*, **91**, 12378–12382.

Saksela, O. and Rifkin, D. B. (1990). Release of basic fibroblast growth factor–heparan sulfate complexes from endothelial cells by plasminogen activator-mediated proteolytic activity. *Journal of Cell Biology*, **110**, 767–775.

Schmidhauser, C., *et al*. (1990). Extracellular matrix and hormones transcriptionally regulate bovine 5′ sequences in stably transfected mouse mammary cells. *Proceedings of the National Academy of Sciences of the USA*, **87**, 9118–9122.

Schneller, M., *et al*. (1997). Alphav beta3 integrin associates with activated insulin and PDGF beta receptors and potentiates the biological activity of PDGF. *EMBO Journal*, **16**, 5600–5607.

Schoenwaelder, S. M. and Burridge, K. (1999). Bidirectional signalling between the cytoskeleton and integrins. *Current Opinions in Cell Biology*, **11**, 274–286.

Schwarzbauer, J. (1999). Basement membranes: putting up the barriers. *Current Biology*, **9**, R242–R244.

Shtutman, M., *et al*. (1999). The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proceedings of the National Academy of Sciences of the USA*, **96**, 5522–5527.

Sonnenberg, A., *et al*. (1991). Integrin alpha 6 beta 4 complex is located in hemidesmosomes suggesting a major role in epidermal cell-basement membrane adhesion. *Journal of Cell Biology*, **113**, 907–917.

Steinbock, F. A. and Wiche, G. (1999). Plectin; a cytolinker by design. *Biological Chemistry*, **380**, 151–158.

Sternlicht, M. D., *et al*. (1999). The stromal proteinase MMP3/ stromelysin-1 promotes mammary carcinogenesis. *Cell*, **98**, 137–146.

Streuli, C. H. and Bissell, M. J. (1990). Expression of extracellular matrix components is regulated by substratum. *Journal of Cell Biology*, **110**, 1405–1415.

Streuli, C. H., *et al*. (1991). Control of mammary epithelial differentiation: basement membrane induces tissue-specific gene expression in the absence of cell–cell interaction and morphological polarity. *Journal of Cell Biology*, **115**, 1383–1395.

Streuli, C. H., *et al*. (1995). Laminin mediates tissue-specific gene expression in mammary epithelia. *Journal of Cell Biology*, **129**, 591–603.

Tamura, M., *et al*. (1999). PTEN gene and integrin signalling in cancer. *Journal of the National Cancer Institute*, **91**, 1820–1828.

Tang, M. J., *et al*. (1998). Collagen gel overlay induces apoptosis of polarized cells in cultures; disoriented cell death. *American Journal of Physiology*, **275**, C921–C931.

Van der Neut, R., *et al*. (1996). Epithelial detachment due to absence of hemidesmosomes in integrin beta 4 null mice. *Nature Genetics*, **13**, 366–369.

Vasioukhin, V., *et al*. (2000). Directed actin polymerization is the driving force for epithelial cell–cell adhesion. *Cell*, **100**, 200–219.

Vu, T. H., *et al*. (1998). MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. *Cell*, **93**, 411–422.

Wang, F., *et al*. (1998). Reciprocal interactions between beta 1 integrin and epidermal growth factor receptor in three-dimensional basement membrane breast cultures; a different perspective in epithelial biology. *Proceedings of the National Academy of Sciences of the USA*, **95**, 14821–14826.

Wary, K. K., *et al*. (1996). The adaptor protein Shc couples a class of integrins to the control of cell cycle progression. *Cell*, **87**, 733–743.

Wary, K. K., *et al*. (1998). A requirement for caveolin-1 and associated kinase Fyn in integrin signalling and anchorage-dependent cell growth. *Cell*, **94**, 625–634.

Weaver, V. M., *et al*. (1996). The importance of the micro-environment in breast cancer progression: recapitulation of

mammary tumorigenesis using a unique human mammary epithelial cell model and a three-dimensional culture assay. *Biochemistry and Cell Biology*, **74**, 833–851.

Weaver, V. M., *et al*. (1997). Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies. *Journal of Cell Biology*, **137**, 231–245.

Woodley, D. T., *et al*. (1983). Interactions of basement membrane components. *Biochimica Biophysica Acta*, **761**, 278–283.

Wu, X., *et al*. (2000). Evidence for regulation of the PTEN tumor suppressor by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2. *Proceedings of the National Academy of Sciences of the USA*, **97**, 4233–4238.

Yamada, K. M. and Geiger, B. (1997). Molecular interactions in cell adhesion complexes. *Current Opinion in Cell Biology*, **9**, 76–85.

Yokosaki, Y., *et al*. (1994). The integrin alpha 9 beta 1 mediates cell attachment to a non-RGD site in the third fibronectin type III repeat of tenascin. *Journal of Biological Chemistry*, **269**, 26691–26696.

Zhu, X. and Assoian, R. K. (1995). Integrin-dependent activation of MAP kinase: a link to shape-dependent cell proliferation. *Molecular Biology of the Cell*, **6**, 273–282.

Zhu, X., *et al*. (1996). Adhesion-dependent cell cycle progression linked to the expression of cyclin D1, activation of cyclinE–cdk2, and phosphorylation of the retinoblastoma protein. *Journal of Cell Biology*, **133**, 391–403.

Zutter, M. M., *et al*. (1995). Re-expression of the alpha 2 beta 1 integrin abrogates the malignant phenotype of breast carcinoma cells. *Proceedings of the National Academy of Sciences of the USA*, **92**, 7411–7415.

Zutter, M. M., *et al*. (1998). Altered integrin expression and the malignant phenotype: the contribution of multiple integrated integrin receptors. *Journal of Mammary Gland Biology and Neoplasia*, **3**, 191–200.

FURTHER READING

Bissell, M. J., *et al*. (1982). Dynamic reciprocity – How does the extracellular matrix direct gene expression? *Journal of Therapeutic Biology*, **99**, 31–68.

Bissell, M. J., *et al*. (1999). Tissue structure, nuclear organization and gene expression in normal and malignant breast. *Cancer Research (Supplement)*, **59**, 1757–1764.

Boudreau, N. J. and Jones, P. L. (1999). Extracellular matrix and integrin signalling: the shape of things to come. *Biochemical Journal,* **339**, 481–488.

Erickson, A. C. and Couchman, J. R. (2000). Still more complexity in mammalian basement membranes. *Journal of Histochemistry and Cytochemistry*, **48**, 1291–1306.

Giancotti, F. and Ruoslahti, E. (1999). Integrin signaling. *Science*, **285**, 28–32.

Werb, Z. (1997). ECM and cell surface proteolysis: regulating cellular ecology. *Cell*, **91**, 439–442.

Chapter 16

Invasion and Metastasis

Lance A. Liotta and Cloud P. Paweletz
National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

CONTENTS

- Introduction
- Angiogenesis
- Invasion
- Cellular Adhesion
- Motility
- Extracellular Matrix Degradation during Invasion
- Coordination of the Machinery of Invasion at the Cell Surface

INTRODUCTION

We usually diagnose and treat cancer when it is too late for local therapeutic strategies and most patients already harbour occult or overt metastasis. In fact, 30% of patients are diagnosed with overt metastases, while an additional 30–40% appear metastasis free during initial diagnosis, but harbour occult metastasis instead. Unfortunately for the cancer patient, the existence of metastasis greatly reduces the success of current surgical, chemotherapy and radiotherapy strategies (Astrow, 1994).

During the course of the disease most patients suffer from metastases at multiple sites, not all of which may be occurring at the same time. Furthermore, metastases have the potential to metastasize further: the presence of large identifiable metastases in a given organ is frequently accompanied by a greater number of micrometastases. And lastly, formation of metastatic colonies is a continuous process that commences early in the growth of the tumour and increases with time.

Cancer metastasis is a highly complex process that involves the deregulation of interacting proteins and genes that are responsible for invasion, angiogenesis, circulation of tumour cells in blood vessels, colonization at secondary organ sites, and finally evasion of host defence systems (**Figure 1**). Metastatic dissemination via the lymphatics and the vascular systems is the culmination, and end stage, of a disease process that evolves over 5–20 years. During most of that time period, microscopic lesions are progressing through a series of hyperproliferative and premalignant states through to carcinoma *in situ* presenting a phenomenally long time period for initial screening and treatment of cancer (**Figure 2**). We know that for most types of human solid neoplasia, microinvasive carcinoma emerges from a carcinoma *in situ* precursor lesion (Gallager and Martin, 1969; Zhuang *et al*., 1995). Preventing the transition from premalignant to invasive/metatastic carcinoma is a major goal for cancer chemoprevention. In order to reach this goal, it will be important to identify molecular targets that are causally associated with the acquisition of cancer invasion.

The malignant tumour is a state that emerges from a host microecology that actually participates in the selection and expansion of the most aggressive neoplastic cells (Aboseif *et al*., 1999; Hanahan and Weinberg, 2000) (**Figure 3**). Instead of being autonomous, malignant cells communicate extensively with other cells and the extracellular matrix. Sustained proliferation is not unique to cancer cells. In fact, growth pressure alone will not cause a neoplasm to be malignant. Malignant tumour cells migrate across tissue boundaries and have the capacity to survive and grow among ‘foreign’ cell populations. The true life-threatening behaviour of malignant cancer cells is their propensity to infiltrate and usurp the ‘sovereignty’ of host tissue societies.

Normal cellular physiology is a tightly regulated process with positive and negative feedback loops that decides whether a cell should differentiate, divide, adapt or commit apoptosis. Genetic changes, such as activation of oncogenes, increased production of growth factors, loss of growth inhibitory cytokines or loss of function of tumour-suppressor genes may result in an imbalance of growth regulation, leading to uncontrolled proliferation. However, unrestrained growth by itself does not cause metastasis, and additional genetic mutations over and above those that cause uncontrolled proliferation are needed. Genetic defects in the cancer cell translate into proteomic derangements in signal transduction pathways. The result of such derangements is a persistent pathological communication state between the tumour cell and the host. Tumour cells that successfully invade and metastasize are selected out because somatic genetic progression has resulted in an altered communication circuit that continues to call up and

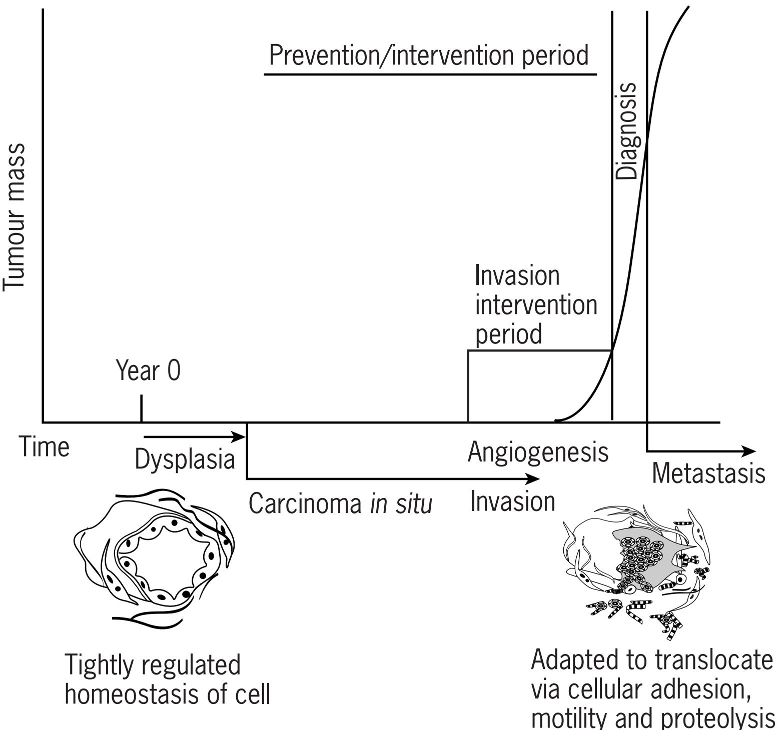


Figure 1 Intervention period for carcinogenesis. Theoretical improvement for intervention strategies as a function of time. Diagnosis of cancer usually is performed late into the disease, when most patients already present with occult metastasis. However, progression of cancer from dysplasia to metastatic dissemination may extend as far back as 10 years, providing a much larger window for intervention strategies before metastasis occurs.

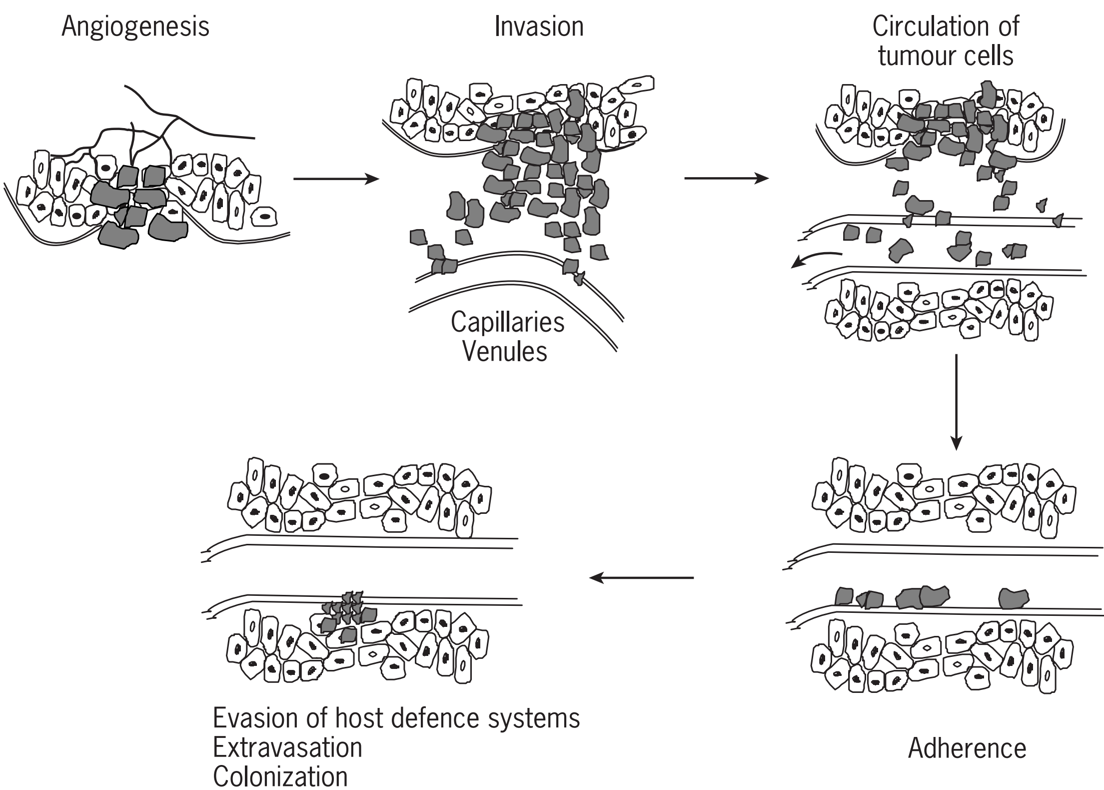


Figure 2 Pathogenesis of metastasis. Cancer metastasis is a highly complex process that involves the deregulation of interacting proteins and genes that are responsible for angiogenesis, invasion, circulation of tumour cells in blood vessels, colonization at secondary organ sites and finally evasion of host defence systems.

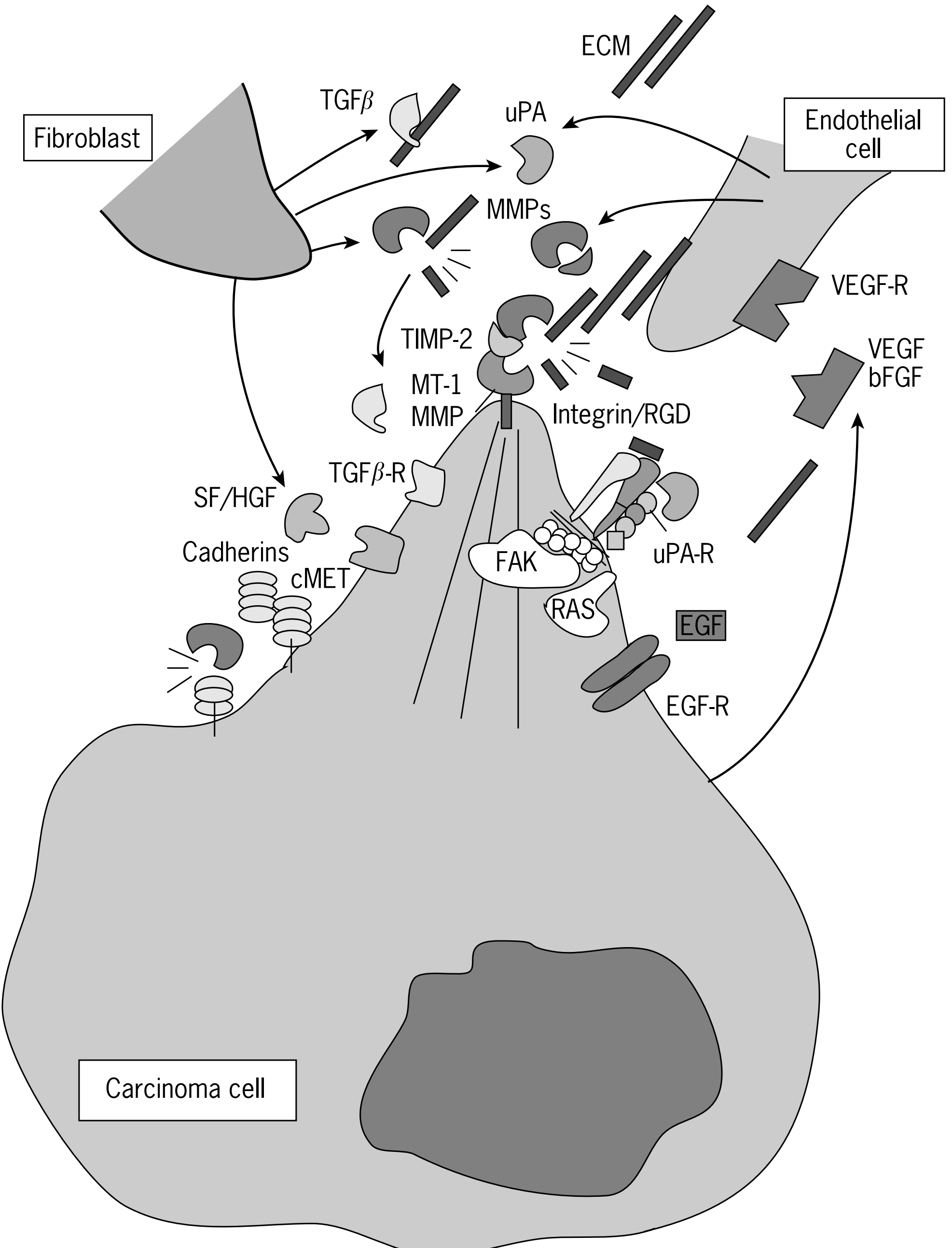


Figure 3 Microecology at the invasion front. Example mediators are shown. Motility and invasion is a bidirectional process. Fibroblasts produce chemoattractants including scatter factor/hepatocyte growth factor (SF/HGF) which stimulates motility of tumour cells by binding to the Met receptor (c-Met). Tumour cells produce angiogenesis factors including vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF), which bind to receptors on stromal vascular cells causing increased vascular permeability, endothelial proliferation, migration, and invasion. Fibroblasts and endothelial stromal cells elaborate latent enzymes including matrix metalloproteinases (MMPs) and urokinase plasminogen activator (uPA) which dock on the surface of the carcinoma invadopodia and become activated, thereby degrading the ECM, and clearing a pathway. ECM degradation releases bound growth factors such as transforming growth factor beta (TGF-β) and epidermal growth factor (EGF), which bind to cognate receptors (TGFβ-R and EGF-R) on the carcinoma cell. ECM proteolysis also exposes cryptic RGD sites which are recognized by integrins (integrin/RGD). Cross-talk between signal pathways within the carcinoma cells links motility, proliferation and prosurvival.

support invasion and survival. At the biochemical level, the mechanism for initial invasion may parallel, or be similar to, that used by nonmalignant cells that traverse tissue boundaries. Using specialized cell models and new array technology, investigators are uncovering the interplay of specific signal transduction molecules that mediate the malignant state (Clark *et al*., 2000; Paweletz *et al*., 2001). Using protein microarray technology, the authors demonstrated that activation of PI3 kinase substrates and suppression of apoptosis are early events in the microenvironment of prostate cancer evolution. This analysis provided direct quantitative evidence that suppression of apoptosis in human PIN and invasive prostate cancer may be associated with phosphorylation of Akt and its substrate GSK3-β. Moreover, the authors verified that downstream components of the apoptotic cascade (cleaved and noncleaved caspase-7, and also cleaved and noncleaved PARP) are shifted toward prosurvival messages at the cancer invasion front. High-grade PIN exhibited a lower level of phospho-ERK compared with normal-appearing epithelium. Invading carcinoma cells contained phospho-ERK levels that were even more reduced compared with PIN. These data are in keeping with known prosurvival pathways, which emanate from Akt through its substrates. Augmentation of the ratio of phosphorylated Akt to total Akt will suppress downstream apoptosis pathways through intermediate substrates such as GSK3β. Reduction in apoptosis will shift the balance of cell birth and death rates favouring the observed accumulation of cells within the epithelial gland. Prosurvival messages are required for migrating cells to resist the proapoptotic signals that take place during the disruption of integrin-mediated adhesion to extracellular matrix molecules. In parallel, transient ERK activation and augmentation of prosurvival pathways may be associated with cellular migration. Activation of Akt, a substrate of PI3K, can therefore promote cell motility and survival as the invading cancer cells leave the gland, invade the stroma and metastasize.

As the tumour cell invader enters foreign soil, it appropriates the local growth signals and ignores its instructions to undergo apoptosis. Tumour cells escaping the primary tumour mass respond to host signals that call up the capacity for motility (Jo *et al*., 2000), survival (Frisch and Francis, 1994), and proliferation (Brown and Giavazzi, 1995; Kohn and Liotta, 1995). Host cells contribute enzymes and cytokines that aid the tumour cell. For example, tumour cells can penetrate host cellular and extracellular barriers with the help of degradative enzymes produced by the host cells, but locally activated by the tumour (Chambers and Matrisian, 1997; Werb, 1997). The presence of the new malignant cells within the invaded host tissue is associated with a local reorganisation of the stroma, blood vessels, lymphatics and epithelial morphology. We may incorrectly assume that the host response to the tumour is designed to repel an invader. Instead, the host reaction to the tumour cells may be simply an indifferent accommodation process (Wernert, 1997). Regardless, the multifocal disruption and damage associated with the modified microenvironment are ultimately lethal to the host.

Distributions of metastases vary widely with histological type and anatomical location of the primary tumour (**Table 1** and **Table 2**). For some tumour types a frequent

Table 1 Organ preferences of metastasis in some human and animal selected models[a]

Tumour system subline	Lung	Liver	Brain	Ovary	Spleen	Lymph node
Murine B16 melanoma (i.v. or i.c.)						
B16–F1	+	+/−	+/−	+/−	−	+/−
B16–F10	++++	+/−	−	+	+/−	+/−
B16–F15b	+++	−	+++	+	+/−	+/−
Murine RAW117 large cell lymphoma (i.v. or s.c.)						
Raw117-P	+/−	+/−	−	−	+/−	−
Raw117-H10	+/−	++++	−	−	++	−
Murine MTI mammary carcinoma (i.v. or s.c.)						
TC3	+++	+	−	+/−	−	+
Chicken MD lymphoma (i.v.)						
AL-2	−	++++	−	+/−	−	−
AL-3	−	+	−	++++	−	−
Human A375 melanoma (i.v. in nude mice)						
A375-P	+/−	+/−	−	−	−	+/−
A375-SM	++	+/−	−	+/−	−	+
Human PC-3 prostatic carcinoma (i.v. in athymic mice)						
PC-3-125-IN	++++	−	−	−	−	−
PC-3-1-LN	++++	+	−	+	+/−	+++

[a]Metastasis: −, none; +/−, sometimes; +, few; ++, moderate; +++, many; ++++, large numbers and heavy tumour burden; i.v., intravenous; i.c., intracaecum. (Adapted from Nicolson, 1998.)

Table 2 Frequency of metastatic sites

Site	10%	10–30%	30–50%	50–70%	70%
Breast		Kidney, skin, brain	Adrenal	Liver, bone, lung	Lymph nodes
Bladder	Brain, skin	Kidney,bone	Adrenal, lung		
Cervix	Brain, skin	Kidney, bone	Adrenal, lung		
Colorectum	Skin	Brain, kidney, lung	Bone, adrenal, liver	Lymph nodes	
Kidney	Skin, bone	Brain, kidney	Liver	Lung	
Lung	Lung	Kidney, distant nodes	Adrenal, brain	Bone	Liver, local lymph nodes
Melanoma		Kidney	Adrenal, brain, bone, skin	Lung, liver nodes	
Ovary	Brain, skin, kidney	Bone, adrenal	Lung, liver nodes		
Prostate	Brain, skin	Kidney, adrenal, liver, lung	Bone, nodes		

(Adapted from Weiss, 1992, and references therein.)

organ location of distant metastases appears to be the first capillary bed encountered. Lung metastases from sarcoma or colorectal cancer dissemination to the liver can be considered examples of this kind of metastasis. In the gynaecological tumours, distant metastases are seen in two forms: serosal dissemination, such as liver capsule metastases from ovarian cancer, and capillary-associated dissemination, such as lung parenchymal disease. However, not all metastases can be explained by anatomical considerations alone, such as metastasis to the ovary from breast carcinoma or dissemination to the liver from ocular melanomas, and hence must be considered as organ tropism.

The organ preference for metastatic colonization is heavily influenced by communications between the circulating tumour cell and the target host tissue. Various molecular mechanisms attempt to explain preferential organ distribution during metastasis. First, cancers shed equal numbers of tumour cells into the vascular system, and thus tumour cells disseminate equally to all organs, but only grow preferentially in some specific organs. For example, the insulin-like-growth factors are present in liver and lung and have been implicated in growth and motility for breast and lung carcinoma. Second, circulating tumour cells may adhere preferentially to the endothelial luminal surface. Nicolson *et al*. have identified endothelial surface antigens that may mediate preferential adhesion of circulating tumour cells. Lastly, circulating tumour cells may respond to soluble factors diffusing locally out of target organs. Chemokines are growth factor-like molecules which bind to G-protein coupled receptors. Circulating leukocytes and stem cells are known to use chemokine mechanisms to home in on specific organs. They induce leucocytes to adhere tightly to endothelial cells and migrate toward the highest concentration of chemokine. Since this behaviour seemed identical with that required for metastatic tumour cells, Mueller *et al*. hypothesized that tumour cells may co-opt the same chemokines to direct metastatic organ preference. They conducted a comprehensive survey of known chemokines and found a receptor–ligand pair (CCR4 and CXCL12) which fit the profile expected for breast cancer metastasis homing to bone, lung and liver. *In vitro*, the CXCL12 ligand stimulated breast cancer cells to carry out the basics of invasion: pseudopodial protrusion, directed migration and penetration of extracellular matrix barriers. *In vivo*, using animal models, the authors blocked metastasis to CXCL12-rich lung tissue by treatment with a neutralizing antihuman CXCR4 monoclonal antibody.

ANGIOGENESIS

The transition from normal epithelium to invasive carcinoma is preceded by, or is concomitant with, activation of local host vascular channels and stromal fibroblasts. Stromal cell activation and recruitment by the tumour cell promotes premalignant cell transformation and malignant invasion. For example, during the transition from *in situ* to invasive carcinoma, disorganisation and disruption of the periglandular basement membrane and a local neovascular 'blush' can precede frank malignant conversion (Guidi *et al*., 1997). Neovascularization offers a portal for dissemination. Locally activated vascular channels at the invasive edge of the tumour are highly permeable and offer a reduced barrier for intravasation (Dvorak *et al*., 1995). A variety of molecules have been found to mediate angiogenesis *in vitro* and *in vivo*. Among these are basic fibroblast growth factor (bFGF), angiogenin, vascular permeability factor (VPF) and tumour necrosis factor α and β (TNF-α, TNF-β) (Folkman and Klagsbrun, 1987).

Neovascularization is a form of physiological invasion (Fidler and Ellis, 1994; Folkman, 1995). Endothelial cells migrate, elaborate degradative enzymes and traverse extracellular matrix barriers along a chemotropic gradient emanating from the tumour cells. Physiological and malignant invasion employ similar molecular mechanisms.

The difference is that malignant invasion persists. Neovascularization, wound healing and neurite outgrowth during embryogenesis are examples of physiological invasion. In response to trophic signals, vascular cells, wounded epithelial sheets or neurites will migrate, penetrate tissue barriers and establish appropriate new anastomoses (Kohn and Liotta, 1995; Carmeliet and Jain, 2000). However, when the trophic signal is removed or the injury is repaired, physiological invasion ceases. Malignant cells perpetually stimulate host stromal and vascular cells to conduct physiological invasion. Promotion of the local invasive environment creates a permissive field for the malignant cell.

Rapid-growing tumours are capable of shedding up to millions of tumour cells into the vascular circulation by angiogenesis and invasion alone (Liotta *et al*., 1974). Furthermore, experimental studies show that less than 0.05% of circulating tumour cells are successful in initiation of metastatic colonies, making metastatic dissemination a highly inefficient process (Liotta *et al*., 1974; Nicolson, 1991). These studies are also clinically validated by the observation that circulating tumour cells are detected in patients who never form a metastasis.

INVASION

Invasion is the active translocation of neoplastic cells across tissue boundaries and through host cellular and extracellular matrix barriers. Invasion is dependent on the coordinated activity of a series of interacting proteins extending from the inside of the cell to the cell surface and the adjacent host cellular and extracellular microenvironment (**Figure 3**). Cellular adhesion, local proteolysis and motility are the triad of necessary functions that mediate invasion. While invasion is not directly caused by growth pressure, nevertheless the genetic and proteomic deregulation that causes invasion can effect neoplastic proliferation indirectly by promoting cancer cell survival.

CELLULAR ADHESION

Normal tissue morphology and organ architecture are tightly regulated by a communication reciprocity between the tissue cells and the extracellular matrix (ECM) and/or basement membrane. Adhesion is more than just anchoring – it is dynamic solid phase signal transduction (Fashena and Thomas, 2000). Receptors sensing changes in the cell-ECM state provide extracellular signals, which trigger corresponding intracellular signal transduction pathways that regulate proliferation, differentiation and migration. Receptors involved in sensing the ECM include growth factor and hormone receptors which recognize ligands solubilized from the ECM and receptors which directly bind to the solid-phase molecules of the ECM. The latter include the integrins, the cell adhesion molecules (CAMs) and cadherins (**Figure 3**).

The integrin family of cell surface extracellular matrix proteins consists of heterodimeric units, designated α (140 kDa) and β (95 kDa). An important aspect of this family is that integrins can exist in a binary 'on' or 'off' state, thereby selectively changing affinity for corresponding ligands (Juliano and Haskill, 1993). Activation of integrins has been shown to be involved in cell migration, cell proliferation and metastatic dissemination. Loss of sustained integrin stimulation (Frisch and Ruoslahti, 1997) has been associated with apoptosis. Integrin interactions with intracellular and extracellular molecules determines function. Such interactions are dictated by the context of the intergrins in the cell function, not the absolute levels of integrins. For example, increased expression of integrin receptors on cell surfaces has been associated with an invasive phenotype of melanoma and squamous carcinoma of the head and neck, whereas loss of integrin expression status has been shown in cancers of breast, prostate and colon (Chammas and Brentani, 1991).

A wide variety of cell–cell adhesion receptors (CAMs) have been studied for their role in cancer invasion. These include, but are not limited to, intercellular adhesion molecules (ICAMs), L-, E- and P-selectins, vascular cell adhesion molecules (VCAMs), neural cell adhesion molecules (NCAMs) and neuroglial cell adhesion molecules (NG-CAMs). Unlike other receptors that bind proteins, selectins bind carbohydrate ligands on endothelial cells. The cadherins comprise a family of transmembrane glycoproteins that mediate Ca^{2+}-dependent cell–cell adhesion (Takeichi, 1991). Special intracellular proteins, the catenins, form zipper-like structures constituting extracellular cell–cell bonds with the cell cytoplasm. These interactions are regulated by tyrosine phosphorylation as well as additional cell to cell communications. Down-regulation of epithelial cadherin transcription, E-cadherin, has been shown to correlate with an aggressive cancer cell phenotype (Frixen *et al*., 1991; Vlemininckx *et al*., 1991). For example, transforming Madin–Darbey canine kidney (MDCK) cells by H-*ras* not only diminshed E-cadherin expression, but also increased the invasive behaviour of these cells. This effect could be reversed by transfecting E-cadherins back into the transformed cells (Vleminckx *et al*., 1991). Furthermore, overexpression of E-cadherin in highly invasive tumour types (bladder, breast, lung and pancreas) caused loss of invasiveness. Surface receptors that participate in cell–cell adhesion and interaction can activate signalling pathways responsible for maintaining normal cell and tissue architecture. Deregulation of these receptors in cancer can promote invasion by (a) reducing cell–cell adhesion which prevents shedding of tumour cells and (b) failing to suppress the inappropriate mixing of tumour cell and host cell populations during invasion.

MOTILITY

Translocation of individual cells across tissue boundaries is a necessary component of invasion. Cell motility and migration are not unique to tumour cells. This process is essential for normal immune cell function and for embryological development, organogenesis and gastrulation. Deregulation and persistence of motility may distinguish carcinoma cells from their normal epithelial counterparts (Nabeshima *et al.*, 1997). The direction of tumour cell motility is controlled by a multitude of chemoattractants, including cytokines (hepatocyte growth factor), collagen peptides, formyl peptides and autocrine growth factors (e.g. Autotaxin) (Anzano *et al.*, 1983). These agents may stimulate both the initiation and maintenance of tumour cell motility and the directness of that migration. Chemoattractants can be secreted by host stromal cells or the tumour cells themselves, or be released from the extracellular matrix.

An early event in motility is cytoskeletal remodelling causing extension of a dominant pseudopod toward the direction of movement. This is followed by translocation of the whole cell body (Stossel, 1993; You *et al.*, 1996). The pseudopod of the invading cell has been renamed an 'invadopodia' because it may direct local proteolytic machinery (Bowden *et al.*, 1999) literally to create a tunnel in the extracellular matrix in front of the cell. Protruding invadopodia, in response to chemoattractants, may serve to sense organs for the migrating cell to locate directional clues, to secrete motility-stimulating factors, to promote propulsive traction for locomotion and induce matrix proteolysis. The complexity of tumour cell migration requires that more than one agent is involved in the direction, location, and magnitude of the migratory response. During the course of invasion, the tumour cell must interact with the extracellular matrix components and be exposed to host-derived factors. Tumour cells have receptors for many of these potential attractants. Therefore, the response of tumour cells to autocrine motility stimulation and also endo- or paracrine stimulation by matrix components and host-derived factors is important to tumour motility.

EXTRACELLULAR MATRIX DEGRADATION DURING INVASION

Proteolytic modification of the cell surface and the extracellular matrix is believed to be an essential component of invasion (Liotta *et al.*, 1980), both neoplastic and physiological. The major enzymes that degrade the ECM and cell-associated proteins are (1) the matrix metalloproteinases (MMPs), a family of secreted and membrane anchored proteinases, (2) the adamalysin-related membrane proteinases, (3) the bone morphogenetic protein 1 type metalloproteinases and (4) tissue serine proteinases including tissue plasminogen activator, urokinase, thrombin and plasmin (Werb, 1997). Major ECM barrier substrates for degradative enzymes include collagens (more than 13 types), proteoglycans, laminin, fibronectin and vitronectin. Each compartment of the ECM contains a different complement of matrix molecules. Collagens I and III are examples of collagens preferentially localized to stroma, while collagens IV and V are predominant in the basement membrane, which forms the border between epithelium and stroma. Proteolysis of the ECM is observed in trophoblast implantation, embryo morphogenesis, wound healing, tissue remodelling and angiogenesis. An imbalance in the ratio of proteinases to protease inhibitor can regulate vascular morphogenesis and invasion (Ura *et al.*, 1989). All classes (serine, aspartyl, cysteinyl and metallo) of matrix-degrading proteinases participate, and coactivate each other, in the tumour–host invasion field (Nakajima *et al.*, 1987; Ostrowski *et al.*, 1988; Reich *et al.*, 1988). Evidence also exists that proteases inside the cell may also be involved during invasion (Koblinski *et al.*, 2000).

A large body of literature exists correlating degradative enzyme activity with cancer invasion and metastasis. The most studied proteases include tissue-type plasminogen activator (tPA), plasmin, cathepsin-D, -B, -L and -G, the urokinase plasminogen activator (uPA), metalloproteinases and the heparanases. Urokinase plasminogen activator, a serine protease, has been shown to correlate with a metastatic phenotype of cells. Antibodies against uPA block human HEP-3 cell invasion and murine B16–F10 melanoma cell metastasis after tail vain injection (Ossowski and Reich 1983; Esheicher *et al.*, 1989). Moreover, overexpression of uPA in H-*ras* transformed cell lines enhance lung metastases (Axelrod *et al.*, 1989). Inhibition of metalloproteinases has been demonstrated to inhibit cell invasion (DeClerck *et al.*, 1991). MMPs can be divided into three general classes: (1) interstitial collagenases, (2) stromelysins and (3) gelatinases. Interstitial collagenase degrades type I, II, III and VII collagens. Stromelysins degrade type I, III, IV, V and IX collagens, laminin, fibronectin, and gelatin. The third group of the MMP family, the gelatinases (MMP-2 and MMP-9), can degrade collagen type I, II, III, IV, V, VII, IX and X and fibronectin (Emonard and Grimaud, 1990). Association of MMP-2 and MMP-9 with the invasive phenotypes is abundant in the literature. Inhibition of MMP-2 by TIMP-1 reduces cellular invasion *in vitro* and *in vivo*. Induction of H-*ras* oncogene enhances expression of MMP-2 and MMP-9. Invasive colonic, gastric, ovarian and thyroid adenocarcinomas showed positive immunoreactivity for MMP-2, whereas normal colorectal, gastric mucosa and benign ovarian cysts showed reduced or negative staining (Monteagudo *et al.*, 1990; Levy *et al.*, 1991). A delicate balance between TIMPs and MMPs may act as a positive and negative feedback control regulating vascular morphogenesis and invasion (Mignatti *et al.*, 1986). MMPs and TIMPs have direct, and indirect, effects on angiogenesis, which are separate from their proteolytic functions (Chambers *et al.*, 1997; Hoegy *et al.*, 2001). Heparan sulfate

proteoglycans (HSPGs), major and ubiquitous components of the ECM, are substrates for heparanases, which cleave heparan sulfate glycosaminoglycan side chains. Augmentation of heparanase activity has been associated with tumour aggressiveness (Nakajima *et al.*, 1988; Vlodavsky *et al.*, 1995). Heparin and similar polysaccharides inhibit metastasis (Parish *et al.*, 1987). Transfection of nonmetastatic murine T-lymphoma Eb cell lines with full-length human heparanase cDNA (Vlodavsky *et al.*, 1999) enhances the metastatic phenotype in animal models. (See the chapters on *Models for Tumour Cell Adhesion and Invasion* and *Tumour Metastasis Models*.)

COORDINATION OF THE MACHINERY OF INVASION AT THE CELL SURFACE

Significant progress has been made in our understanding of the molecular cross-talk between tumour cells and host cells at the invasion front. A cascade of cytokines, motility factors, matrix receptors, enzymes and enzyme inhibitors simultaneously carries out the regulation, steering, proteolysis, traction and locomotion required for invasion (**Figure 3**).

Remodelling of the extracellular matrix, within the immediate pericellular environment of the cell, appears to be a necessary step in local invasion (Liotta *et al.*, 1991; Werb, 1997). The complement of enzyme classes is tightly and exquisitely regulated by a series of activation steps and specific inhibitors. In a striking demonstration of host–tumour interdependence, a majority of the enzymes and inhibitors complexed at the invasion front are contributed by host cells, not by the invading tumour cells (Nakahara *et al.*, 1997; Bowden *et al.*, 1999; Coussens *et al.*, 2000).

The enzyme machinery is confined to the cell surface at the point of invading pseudopodia ‘invadopodia’ by binding the enzymes to adhesion sites, cell surface receptors and adjacent ECM molecules (Nakahara *et al.*, 1997; Bowden *et al.*, 1999; Hoegy *et al.*, 2001). MT1-MMP, a membrane-anchored ECM-degrading enzyme, contains a transmembrane–cytoplasmic sequence that confines it to microinvasion sites on the tumour cell invadopodia surface (**Figure 3**). In complex with one of the tissue inhibitors of metalloproteinases (TIMP-2) it becomes a receptor and activator of MMP-2 (Nakahara *et al.*, 1997), a soluble MMP produced by stromal fibroblasts and endothelial cells. The serine proteinase uPA is confined to the invading pseudopodia through a cooperation between integrins and the uPA receptor (uPA-R) (Andreasen *et al.*, 1997). uPA-R is an adhesion receptor for vitronectin, and also interacts laterally with integrin β chains. Proteolysis of ECM proteins modifies integrin mediated anchorage, focal adhesions and cytoskeletal architecture and triggers signalling molecules such as focal adhesion kinase (FAK) (Braga, 2000; Fashena and Thomas, 2000). Such heterotypic complexes direct and confine the enzymatic field at the forward edge of the invading cell, leaving intact the peripheral and distal attachment sites required for traction. As the invading cell moves forward through ECM barriers, the leading edge complex of enzymes, inhibitors and receptors molecules cycle through adhesion, deadhesion and proteolysis. The direction of tumour cell invasion and migration can be influence by chemoattractants and by marking of preferred adhesion pathways. Local attractants include (1) hepatocyte growth factor/scatter factor, which binds to the Met (c-Met) receptor (Wernert, 1997; Jo *et al.*, 2000), (2) proteolysed matrix fragments which are recognized by integrins (Varner and Cheresh, 1996) or (3) cytokines and growth factors, such as EGF and TGF-β released from the degraded matrix (Roberts *et al.*, 1992). Cryptic RGD sites exposed by proteolysis (Davis, 1992; Fukai *et al.*, 1995; Varner *et al.*, 1995; Varner and Cheresh, 1996) may guide the path in front of the invading tumour cells.

The combination of microdissection and protein microarrays has been successfully applied to the micro-world of early stage cancer (Clark *et al.*, 2000; Paweletz *et al.*, 2001). Protein lysate microarrays consist of very small mass quantities (picograms) of protein lysates from cell lines, whole lysed tissue or microdissected subpopulations of lysed tissue cells immobilized and arrayed on a solid phase. The array can be probed with antibodies recognizing phosphorylated forms of signal proteins. Detection is highly sensitive, quantitative and precise, so that the state of signal pathways may be profiled. Individual subpopulations of host and tumour tissue cells within a microscopic field of invasion or premalignant transition can be microdissected and individually studied.

In conclusion, the process of cancer invasion is a coordinated effort by tumour cells and host cells within a microinvasion field. Within this field the tumour cells exchange cytokines, enzymes, inhibitors and growth factors which promote invasion by all cells involved. Pericellular remodelling of the ECM is commensurate with invasion. The different events of the metastatic cascade of angiogenesis, adhesion, proteolysis, motility and proliferation may provide useful and novel therapeutic targets. Investigators have identified some of the critical molecules involved in the extracellular cross-talk taking place among and between cells in the invasion field. This synthesis provides strategies for a new therapy concept ‘stromal therapy’ which targets the tumour–host communication interface. (See the chapter *Targeting the Extracellular Matrix*.)

REFERENCES

Aboseif, S., *et al.* (1999). Mesenchymal reprogramming of adult human epithlelial differentiation. *Differentiation*, **65**, 113–118.

Andreasen, P., *et al*. (1997). The urokinase type plasminogen activator system in cancer metastasis: a review. *International Journal of Cancer*, **71**, 1–22.

Anzano, M. A., *et al*. (1983). Sarcoma growth factor from conditioned medium of virally transformed cells is composed of both type alpha and type beta transforming growth factors. *Proceedings of the National Academy of Sciences of the USA*, **80**, 6264–6268.

Astrow, A. B. (1994). Commentary: rethinking cancer. *Lancet*, **343**, 494–495.

Axelrod, J. H., *et al*. (1989). Expression of human recombinant plasminogen activators enhance invasion and experimental metastasis of Ha-RAS-transformed NIH 3T3 cells. *Molecular Cell Biology*, **9**, 2133–2141.

Bowden, E., *et al*. (1999). An invasion related complex of cortactin, paxillin, and PKC u associates with invadopodia at sites of extracellular matrix degradation. *Oncogene*, **18**, 4440–4449.

Braga, V. (2000). The crossroads between cell–cell adhesion and motility. *Nature Cell Biology*, **2**, E182–E184.

Brown, P. D. and Giavazzi, R. (1995). Matrix metalloproteinase inhibition: a review of anti-tumour activity. *Annals of Oncology*, **6**, 967–974.

Carmeliet, P. and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature*, **407**, 249–257.

Chambers, A. F. and Matrisian, L. M. (1997). Changing views of the role of matrix metalloproteinases in metastasis. *Journal of the National Cancer Institute*, **89**, 1260–1270.

Chammas, R. and Brentani, R. (1991). Integrins and metastases: an overview. *Tumour Biology*, **12**, 309–320.

Clark, E. A., *et al*. (2000). Genomic analysis of metastases reveals and essential role for RhoC. *Nature Cell Biology*, **406**, 532–535.

Coussens, L. M., *et al*. (2000). MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. *Cell*, **103**, 481–490.

Davis, G. E. (1992). Affinity of integrins for damaged extracellular matrix: alpha v beta 3 binds to denatured collagen type I through RGD sites. *Biochemical and Biophysical Research Communications*, **182**, 1025–1031.

DeClerck, Y. A., *et al*. (1991). Inhibition of tumor cell invasion of smooth muscle cell layers by recombinant human metalloproteinase inhibitor. *Cancer Research*, **51**, 2151–2157.

Dvorak, H. F., *et al*. (1995). Vascular permeability factor/vascular endothelial growth factor, microvascular hyperpermeability, and angiogenesis. *American Journal of Pathology*, **146**, 1029–1039.

Emonard, H. and Grimaud, J. A. (1990). Matrix metalloproteinases: a review. *Cell Molecular Biology*, **36**, 131–153.

Esheicher, A., *et al*. (1989). Characterization of the cellular binding site for the urokinase type plasminogen activator. *Journal of Biological Chemsitry*, **264**, 1180.

Fashena, S. and Thomas, S. M. (2000). Signaling by adhesion receptors. *Nature Cell Biology*, **2**, E225–E229.

Fidler, I. J. and Ellis, L. M. (1994). The implication of angiogenesis for the biology and therapy of cancer metastasis. *Cell*, **79**, 185–188.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Medicine*, **1**, 27.

Folkman, J. and Klagsbrun, M. (1987). Angiogenic factors. *Science*, **235**, 442–447.

Frisch, S. and Francis, H. (1994). Disruption of epithelial cell–matrix interactions induces apoptosis. *Journal of Cell Biology*, **124**, 619–626.

Frisch, S. M. and Ruoslahti, E. (1997). Integrins and anoikis. *Current Opinions in Cell Biology*, **9**, 701–706.

Frixen, U. H., *et al*. (1991). E-cadherin mediated cell–cell adhesion prevents invasiveness of human carcinoma cells. *Journal of Cell Biology*, **113**, 173–185.

Fukai, F., *et al*. (1995). Release of biologically activities from quiescent fibronectin by conformational change and limited proteolysis by matrix metalloproteinases. *Biochemistry*, **34**, 11453–11459.

Gallager, H. S. and Martin, J. E. (1969). The study of mammary carcinoma by mammography and whole organ sectioning, early observation. *Cancer*, **23**, 855–873.

Guidi, A. J., *et al*. (1997). Vascular permeability factor (vascular endothelial growth factor) expression and angiogenesis in patients with ductal carcinoma in situ of the breast. *Cancer*, **80**, 1945–1953.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Hoegy, S., *et al*. (2001). Tissue inhibitor of metalloproteinases-2 (TIMP-2) suppresses TKR-growth factor signaling independent of metalloproteinase inhibition. *Journal of Biological Chemistry*, **276**, 3203–3214.

Jo, M., *et al*. (2000). Cross talk between epidermal growth factor receptor and c-Met signal pathways in transformed cells. *Journal of Biological Chemistry*, **275**, 8806–8811.

Juliano, R. L. and Haskill, S. (1993). Signal transduction from the extracellular matrix. *Journal of Cell Biology*, **120**, 577–585.

Koblinski, J. E., *et al*. (2000). Unraveling the role of proteases in cancer. *Clinica Chimica Acta*, **291**, 113–135.

Kohn, E. C. and Liotta, L. A. (1995). Molecular insights into cancer invasion: strategies for prevention and intervention. *Cancer Research*, **55**, 1856–1862.

Levy, A., *et al*. (1991). Increased expression of the 72 kDa type IV collagenase in human colonic adenocarcinoma. *Cancer Research*, **51**, 439–444.

Liotta, L. A., *et al*. (1974). Quantitative relationships of intravascular tumor cells: tumor vessels and pulmonary metastases following tumor implantation. *Cancer Research*, **34**, 997.

Liotta, L. A., *et al*. (1980). Metastatic potential correlates with enzymatic degradation of basement membranes. *Nature*, **284**, 67–68.

Liotta, L. A., *et al*. (1991). Cancer metastasis and angiogenesis: an imbalance of positive and negative regulation. *Cell*, **64**, 327–336.

Mignatti, P., *et al*. (1986). In vitro angiogenesis on the human amniotic membrane: requirement for basic fibroblast growth factor-induced proteinases. *Journal of Cell Biology*, **108**, 671–682.

Monteagudo, C., *et al*. (1990). Immunohistologic distribution of type IV collagenases in normal, benign, and malignant breast tissue. *American Journal of Pathology*, **136**, 585–592.

Mueller, A., *et al*. (2001). Involvement of chemokine receptors in breast cancer metastasis. *Nature*, **403**, 50–56.

Nabeshima, K., *et al*. (1997). Cohort migration of cancer cells. *Connective Tissue*, **29**, 199.

Nakahara, H., *et al*. (1997). Transmembrane/cytoplasmic domain mediated membrane type 1-matrix metalloproteinase docking to invadopodia is required for cell invasion. *Proceedings of the National Academy of Sciences of the USA*, **94**, 7959–7964.

Nakajima, M., *et al*. (1987). Degradation of basement membrane type IV collagen and lung subendothelial matrix by rat mammary adenocarcinoma cell clones of differing metastatic potentials. *Cancer Research*, **47**, 4869–4876.

Nakajima, M., *et al*. (1988). Heparanase and tumor metastasis. *Journal of Cell Biochemistry*, **36**, 157–167.

Nicolson, G. L. (1991). Gene expression, cellular diversification and tumor progression to the metastatic phenotype. *Bioessays*, **13**, 337–342.

Ossowski, L. and Reich, E. (1983). Antibodies to plasminogen activator inhibit human tumor metastasis. *Cell*, **35**, 611–619.

Ostrowski, L. E., *et al*. (1988). Expression pattern of a gene for a secreted metalloproteinase during late stages of tumor progression. *Molecular Carcinogenesis*, **1**, 13–19.

Parish, C. R., *et al*. (1987). Evidence that sulphated polysaccharides inhibit tumor metastasis by blocking tumor cell-derived heparanase. *International Journal of Cancer*, **40**, 511–518.

Paweletz, C. P., *et al*. (2001). Reverse phase protein microarrays which capture disease progression shown activation of pro-survival pathways at the cancer invasion front. *Oncogene*, **20**, 1981–1989.

Reich, R., *et al*. (1988). Effects of inhibitors of plasminogen activator, serine proteases, and collagenase IV on the invasion of basement membranes by metastatic cells. *Cancer Research*, **48**, 3307–3312.

Roberts, A. B., *et al*. (1992). TGF-beta: regulation of extracellular matrix. *Kidney International*, **41**, 557–559.

Stossel, T. P. (1993). On the crawling of animal cells. *Science*, **260**, 1086–1094.

Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*, **251**, 1451–1455.

Ura, H., *et al*. (1989). Expression of type IV collagenase and procollagen genes and its correlation with tumorigenic, invasive, and metastatic abilities of oncogene transformed human bronchial cells. *Cancer Research*, **49**, 4615–4621.

Varner, J. A. and Cheresh, D. A. (1996). Integrins and cancer. *Current Opinions in Cell Biology*, **8**, 724–730.

Varner, J. A., *et al*. (1995). The integrin α v β3: angiogenesis and apoptosis. *Cell Adhesion Communications*, **3**, 367–374.

Vleminckx, K., *et al*. (1991). Genetic manipulation of E-cadherin by epithelial tumor cells reveals an invasion suppressor role. *Cell*, **66**, 107–119.

Vlodavsky, I., *et al*. (1995). Inhibition of tumor metastasis by heparanase inhibiting species of heparin. *Invasion and Metastasis*, **14**, 290–302.

Vlodavsky, I., *et al*. (1999). Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nature Medicine*, **5**, 793–802.

Weiss, L. (1992). Comments on hematogenous metastatic patterns in humans as revealed by autopsy. *Clinical and Experimental Metastasis*, **10**, 191–199.

Werb, Z. (1997). ECM and cell surface proteolysis: regulating cellular ecology. *Cell*, **91**, 439–442.

Wernert, N. (1997). The multiple roles of tumor stroma. *Virchows Archives*, **430**, 433–443.

You, J., *et al*. (1996). Responses of tumor cell pseudopod protrusion to changes in medium osmolality. *Journal of Cell Physiology*, **167**, 156–163.

Zhuang, Z., *et al*. (1995). Identical allelic loss on chromosome 11q13 in microdissected *in situ* and invasive human breast cancer. *Cancer Research*, **55**, 467–471.

FURTHER READING

Carmeliet, P. and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature*, **407**, 249–257.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Kohn, E. C. and Liotta, L. A. (1995). Molecular insights into cancer invasion: strategies for prevention and intervention. *Cancer Research*, **55**, 1856.

Nicolson, G. L. (1998). Organ specificity of tumour metastasis: role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer and Metastasis Reviews*, **7**, 173–188.

Chapter 17

Angiogenesis

Ute Modlich and Roy Bicknell
University of Oxford, John Radcliffe Hospital, Oxford, UK

CONTENTS

- What is Angiogenesis?
- 'Tumour Growth is Angiogenesis Dependent'
- Morphology of Blood Vessels
- The Tumour Vasculature is Disorganized
- Mechanisms of Angiogenesis
- Angiogenic Factors
- The Breakdown of the Basement Membrane
- Migration
- Proliferation
- Vessel Formation and Maturation
- The Prognostic Significance of Tumour Angiogenesis

WHAT IS ANGIOGENESIS?

Like all tissues, tumours depend on a continuous supply of oxygen and nutrients for their survival. The ability to recruit a functional blood supply is therefore central to tumorigenesis. Angiogenesis is the growth of new blood vessels from the pre-exisiting vasculature by budding and sprouting of endothelial cells. This is in contrast to vasculogenesis, which is the *de novo* formation of blood vessels from endothelial precursor cells (called angioblasts) (Rissau, 1997). Vasculogenesis occurs mainly during embryogenesis, in particular the development of tissues of endodermal origin, although some ectodermal and mesodermal derived tissues acquire their vessels through angiogenesis, e.g. the kidney and the brain (Beck and D'Amore, 1997).

In an adult, the vasculature is remarkably quiescent and angiogenesis occurs only very rarely. Diffusion of oxygen in tissues is limited to a distance of about 150 μm. Therefore, tissue growth is restricted to a few cubic millimetres if no new vasculature is formed. Under physiological conditions, angiogenesis is involved in the turnover of tissues in the female reproductive system (endometrium, placenta, follicle maturation and corpus luteum formation in the ovaries) and in wound healing. Angiogenesis is, on the other hand, an essential component of many pathologies such as diabetic retinopathy, rheumatoid arthritis, psoriasis and tumour growth.

'TUMOUR GROWTH IS ANGIOGENESIS DEPENDENT'

Around 30 years ago, it was recognized that the growth of solid tumours is angiogenesis dependent (Folkman, 1971, 1990). Thus, tumours remain in a dormant state of a few millimetres in diameter (prevascular phase) unless they are able to recruit their own vascular bed. This does not mean that the tumour cells are unable to proliferate, but an inadequate supply of oxygen and nutrients results in a high rate of apoptosis (cell death). The tumour exists in a balance of proliferation and regression which precludes tumour expansion (tumour dormancy). Only after a tumour has recruited its own blood supply can it expand in size. This involves the production of angiogenic factors secreted into the tissue by the tumour cells and is known as the 'angiogenic switch' (Hanahan and Folkman, 1996). This induces angiogenesis in the adjacent quiescent vasculature, allowing the growth of new vessels into the tumour.

Why and how new vessels are formed is of great interest in terms of anticancer therapy because inhibition of new vessel formation could restrict tumour growth. The tumour is essentially 'starved'. Such strategies could also include the destruction of tumour vessels which are already formed. Increased understanding of the mechanism of angiogenesis could allow subsequent design of therapeutics which interfere with the process. The aim of this chapter is to describe the sequence of events during angiogenesis and to

outline its importance in tumour biology. (See the chapter *Antiangiogenic Therapy*.)

MORPHOLOGY OF BLOOD VESSELS

The entire vasculature is lined by a single layer of specialized cells, the endothelial cells. They form a simple squamous epithelium which rests on a basement membrane and surrounds a lumen in which blood flows (**Figure 1**). The basement membrane contains collagen types I and IV, fibronectin, laminin, entactin and other non-collagen glycoproteins. Endothelial cells, together with their basement membrane, constitute the lamina intima of a blood vessel. The lamina intima is surrounded by pericytes, smooth muscle cells or cardiomyocytes, collectively described as mural cells. Pericytes lie within the basement membrane of the endothelial cells whereas the others surround larger vessels and the endocardium as an additional layer outside the basement membrane (as a part of the lamina media). Pericytes are recruited by endothelial cells during development and play an important role in vessel maturation and stabilization. They appear to suppress the turnover of endothelial cells. During neovascularization pericytes are selectively lost from angiogenic vessels, including the tumour vasculature.

The endothelium in an adult has a surface area of $>1000\,m^2$. In addition to forming a static physiological barrier, it possesses secretory, synthetic, metabolic and immunological functions. The endothelium is highly heterogeneous; its precise nature varies depending upon the function of the vascular bed in different tissues.

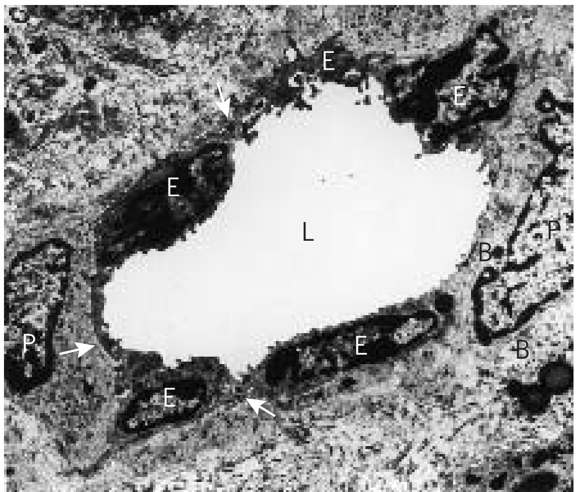


Figure 1 Ultrastructure of a capillary. The vessel lumen (L) is surrounded by five endothelial cells (E). The cells rest on a continuous basement membrame (B). Two pericytes (P) lie within the basement membrane. The endothelial cell contacts (tight junctions) are visible as electron-dense areas of the endothelial cell membranes (arrows).

Histologically there are three main taxonomic classes of endothelia. For example, blood vessels of the brain and the retina are especially tight (continuous endothelium), whereas the endothelium of the sinusoidal vessels in the liver, spleen and bone marrow contains intracellular gaps not covered by a basement membrane and so allows cellular trafficking (discontinuous endothelium); the endothelium in endocrine glands, the kidney and small intestine shows intracellular holes (fenestrated endothelium), which allows extensive exchange of substances.

At the molecular level, the endothelium of different tissues varies in its surface phenotype and protein expression. The distinct differences are induced by the extracellular matrix which adapts the endothelium to its functional needs in a tissue. This observation has led to the concept of ‘vascular addresses.’ It explains the homing of inflammatory cells in specific tissues because they adhere to the vessel wall by attaching to specific endothelial cell surface molecules. The unique vascular addresses of organs are also the reason for tissue-directed metastasis of tumour cells. It has been shown that tumour cells recognize specific adhesion molecules on the endothelium. They adhere there and are subsequently able to extravasate and to invade the tissue. In the tissue stroma they then form micrometastases.

THE TUMOUR VASCULATURE IS DISORGANIZED

The vascular tree of tumour vessels appears ‘chaotic’ and ‘disorganized.’ Tumour vessels are also hyperpermeable (‘leaky’) owing to a discontinuous endothelium and lack of pericytes. Morphologically the vessels appear highly tortuous. The tumour vasculature is formed in two ways: by co-opting pre-existing vessels and by induction of new vessels by angiogenesis. In addition, every tumour induces its own characteristic vascular bed which is different for each tumour type. As a result of persistent growth, the tumour vasculature constantly changes its shape and is dynamic rather than static. In tumours, up to 30% of the vasculature consists of arterio-venous shunts where the blood bypasses the capillaries, precluding exchange of nutrients. Tumour blood vessels lack mural cells as well as appropriate innervation and therefore blood pressure is poorly controlled (only by pre-existing arterioles). The blood flow in the tumour is slow but higher in the tumour periphery than in the centre. Different regions of perfusion can be seen within tumours: necrotic, seminecrotic and well-vascularized regions, in order from the centre of the tumour to the periphery. The blood flow in the periphery of a tumour is higher than that in the surrounding normal tissue (advancing front), whereas in the centre, blood flow can be interrupted transiently resulting in complete stasis; blood flow may then be re-established in the opposite direction.

Extravasation of molecules from the blood occurs through diffusion, convection and partially through transcytosis. Despite the fact that blood flow within tumours is slow, and the vessels are leaky, the delivery of therapeutics into tumours is inefficient. This is due to a high interstitial fluid pressure in the tumour tissue compromising convective movements of therapeutics. In regions of highest pressure in a tumour, the hydrostatic and osmotic pressure in the vessels is opposed by that in the tissue, inhibiting the exchange of substances. This results in hypoxic and hypoglycaemic regions within a tumour. In addition, there are totally anoxic areas. Because the vasculature is dynamic, and perfusion rates in different regions of a tumour change frequently, hypoxic regions can become reoxygenated again. Similarly, normoxic regions can experience sudden oxygen deprivation and become hypoxic. Owing to the poor perfusion of vessels, endothelial cells also can become hypoxic. It has been shown that hypoxia is a potent inducer of many molecules involved in angiogenesis.

These observations highlight the fact that the tumour vasculature is not an extension of 'normal' blood vessels but that it is itself distinct and characteristic. This renders the tumour vasculature an excellent target for antitumour therapy. In addition, targeting the vessels is associated with a number of advantages over current therapies (Augustin, 1998). Targeting its uniqueness should minimize interference with other blood vessels in the body. In addition, endothelial cells are much more accessible than tumour cells to systemically administered therapeutics because of their intimate contact with the blood. Finally, destruction of only a few endothelial cells can have detrimental effects on the tumour. The destruction of a single blood vessel will lead to an amplification of tumour destruction because about 10 layers of tumour cells are dependent on the supply by one blood vessel.

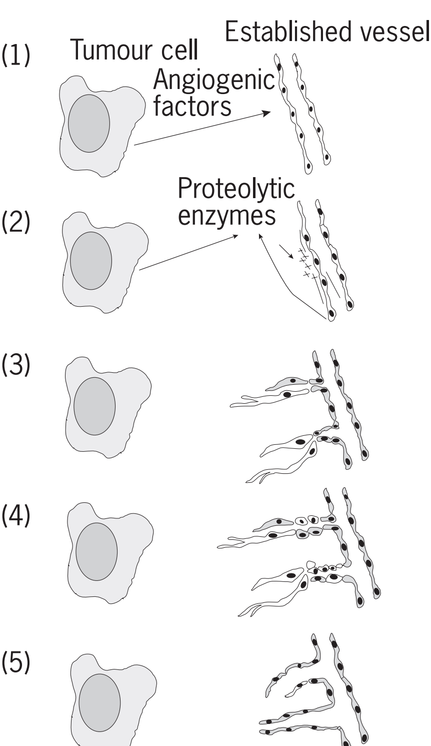


Figure 2 The angiogenic cascade. (1) Angiogenic stimulus; (2) Degradation of the basement membrane and ECM by proteases released from tumour and activated endothelial cells; (3) Migration of EC towards angiogenic stimulus; (4) Endothelial cell proliferation; (5) Tube formation and vessel maturation.

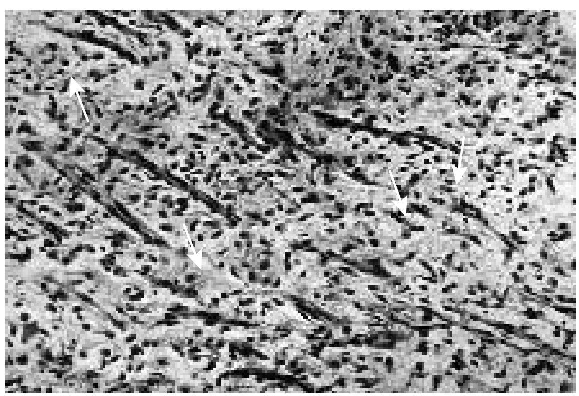


Figure 3 Angiogenic blood vessels in the rat sponge assay (paraffin section, haematoxylin/eosin stained). In this assay inert sponge is implanted subcutaneously. Endothelial cells migrate from the surrounding tissue into the sponge and form new vessels. Some vessels are indicated by arrows.

MECHANISMS OF ANGIOGENESIS

A tumour can only continue to expand in size if it induces a blood supply. When a tumour starts to produce angiogenic factors it activates endothelial cells in the vasculature of the surrounding tissue to initiate angiogenesis (**Figure 2**). The angiogenic stimulus induces the endothelial cells of the 'mother' vessels to change from a quiescent to an activated phenotype (Auerbach and Auerbach, 1994). These endothelial cells produce proteolytic enzymes which break down their basement membrane. This is the prerequisite for endothelial cells to migrate into the surrounding tissue towards the angiogenic stimulus. The migrating endothelial cell changes shape to an elongated phenotype (**Figure 3**). They start expressing typical cell surface molecules which allow the cells to migrate along the extracellular matrix. These endothelial cells also start to proliferate and to form new tubes. Finally, these new tubes anastomose into loops that allow blood flow. The recruitment of pericytes and smooth muscle cells completes angiogenesis through blood vessel maturation. The endothelial cells in the new vessels then revert to a quiescent phenotype.

Table 1 Angiogenic factors

Growth factors
Vascular endothelial growth factor (VEGF)
Placenta growth factor (PIGF)
Basic fibroblast growth factor (bFGF or FGF-2)
Acidic fibroblast growth factor (aFGF or FGF-1)
Transforming growth factor-α (TGF-α)
Transforming growth factor-β TGF-β)
Platelet-derived growth factor (PDGF)
Hepatocyte growth factor (HGF)
Granulocyte colony-stimulating factor (G-CSF)
Cytokines
Tumour necrosis factor-α (TNF-α)
Interleukin 1 (IL-1)
Chemokines
Interleukin 8 (IL-8)
Enzymes
Platelet-derived endothelial cell growth factor (PD-ECGF)
Angiogenin
Prostaglandins
PGE_1
PGE_2

ANGIOGENIC FACTORS

In the last 20 years, many factors that have angiogenic activity have been identified. The first to be identified were the fibroblast growth factors (aFGF or FGF-1 and bFGF or FGF-2), which are now known to be pleiotrophic growth factors. Others include the transforming growth factors (TGF-α and TGF-β), platelet-derived growth factor (PDGF), hepatocyte growth factor (HGF), granulocyte colony-stimulating factor (G-CSF), tumour necrosis factor-alpha (TNF-α), platelet-derived endothelial cell growth factor (PD-ECGF, thymidine phosphorylase), interleukin-8 and prostaglandins (PGE_1, PGE_2) (**Table 1**). All of these have been studied extensively and have been shown to be angiogenic (Folkman and Shing, 1992), but none of these are endothelial cell specific. The only growth factor known to be specific for endothelial cells is vascular endothelial growth factor (VEGF).

Vascular Endothelial Growth Factor (VEGF)

VEGF was the first member to be identified of a growing family of vascular endothelial growth factors and is now referred to as VEGF-A. Other members include VEGF-B to E and placenta growth factor (PIGF-1 and 2). VEGFs are homodimeric proteins and mediate their activity through tyrosine–kinase receptors (VEGF-receptors, VEGFR1-3) which are almost exclusively expressed on endothelial cells. VEGFR-1 (flt-1, fms-like kinase-1) binds VEGF-A, VEGF-B and PIGF with strong affinity whereas VEGFR-2

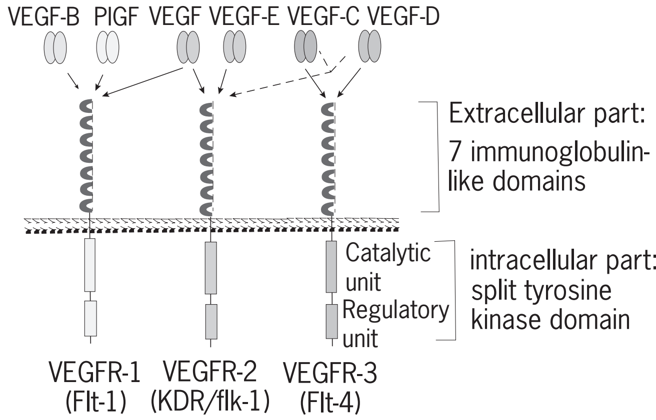


Figure 4 VEGF receptors and their ligands.

(flk-1, foetal liver kinase-1 or the human homologue KDR, kinase-insert domain receptor) binds VEGF-A, VEGF-C and VEGF-D with lower affinity. VEGFR-3 (flt-4) is only expressed on lymphatic endothelium and is the ligand for VEGF-C which induces lymphangiogenesis. VEGF-D has been discovered to be an additional ligand for the VEGFR-3. VEGF-E seems to signal mainly through VEGFR-2 (**Figure 4**).

Of all the VEGFs, VEGF-A has been studied most intensively (for a review, see Neufeld *et al*., 1999). It exists in at least five different splice variants ($VEGF_{121,145,165,189,206}$) encoding isoforms of differing length derived from a single gene. With the exception of $VEGF_{121}$, which is secreted, all VEGFs are heparin binding. Therefore, they accumulate in the extracellular matrix and can be released from there by proteolytic enzymes. There is, in addition, a sixth 110 amino acid isoform of VEGF arising from proteolytic cleavage of $VEGF_{189}$, when bound to the cell surface. Heparan sulfate proteo glycans are known to modulate growth factor signalling of many heparin-binding growth factors from their respective receptors, and the same might apply to VEGF signalling.

VEGF is an endothelial cell mitogen and chemoattractant, it promotes cell migration, inhibits apoptosis and modulates the permeability of the endothelial cell layer (it was first identified as vascular permeability factor (VPF) in 1983). Hence it has major roles as a key regulator in angiogenesis and vasculogenesis. Mice deficient in one allele of VEGF (VEGF +/− mice) show early embryonic death (embryonic day 11–12) due to cardiovascular defects in most sites of early blood vessel formation (embryo and yolk sac). This suggests that a minimal dosage requirement for the growth factor exists because a single allele could not rescue the phenotype. Microinjection of VEGF into quail embryos during development induces uncontrolled and unlimited vascularization at sites that are normally avascular. Together these observations show that the level of VEGF expression is tightly regulated and that small variations can have fatal effects.

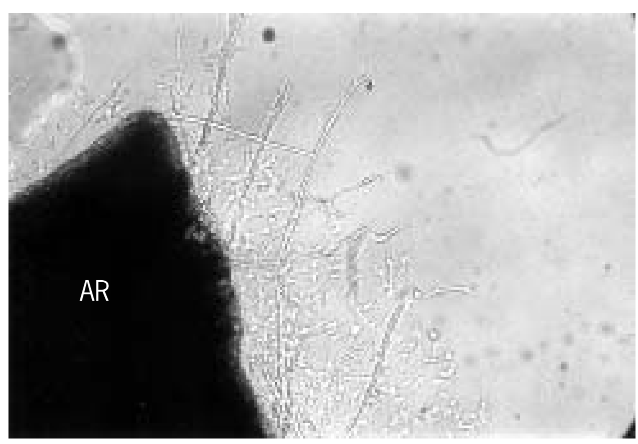


Figure 5 Rat aortic ring angiogenesis assay. A section of a rat aorta is placed in a fibrin gel in cell culture. Endothelial cells from the inner lining of the aortic ring (AR) sprout into the fibrin gel. Some of the newly formed tubes are seen to anastomose.

VEGFR-2-deficient mice (*flk-1 −/−* mice) die *in utero* between embryonic days 8.5 and 9.5. Endothelial cells in these animals fail to differentiate from their angioblastic precursors and the animals die as a result of a total lack of vascular structures. VEGFR-2 is the earliest marker for differentiation towards endothelial cells in development. Similarly, VEGFR-1-deficient mice (*flt-1 −/−*) show abnormal and disorganized vascular structures and die between embryonic days 9.5 and 10.5. These two gene-targeted mice, despite both developing vascular defects and being embryonic lethal, show very different phenotypes. Furthermore the VEGF receptors do not show redundancy in their function.

VEGF has been established as an angiogenic factor *in vitro* in two- and three-dimensional cell culture systems (**Figure 5**). VEGF also proved its angiogenic properties in the rabbit cornea assay and the chick chorioallantoic assay. Tumour cells overexpressing VEGF form faster-growing tumours and contain significantly more blood vessels than tumours formed from untransfected control cells. VEGF also synergizes with bFGF and the angiogenic effects of each growth factor are potentiated.

VEGF expression is regulated by hypoxia and hypoglycaemia, and especially high levels are present within hypoxic regions in tumours. Oxygen tension in tumours is about 50 times lower than in normal tissues and hypoxia could be a major factor driving tumour angiogenesis. The cellular response to hypoxia is mediated by the hypoxia-inducible factor Hif-1. Hif is a dimer comprised of two subunits by Hif-1α and Hif-1β. This dimeric complex recognizes a specific DNA sequence, known as the hypoxia-response element (HRE), which is found within the promoters of hypoxia regulated genes such as the VEGF promoter. The promoter of VEGFR-1 also contains hypoxia response elements whereas the promoter of VEGFR-2 does not. Hence the expression of VEGFR-1 is enhanced by hypoxia directly but not the VEGFR-2. The expression of VEGFR-2 can be induced by binding of VEGF to VEGFR-2 and as a result the VEGFR-2 is indirectly upregulated by hypoxia.

Macrophages in Tumour Angiogenesis

Unlike physiological angiogenesis, tumour growth is usually accompanied by an inflammatory response. Widespread infiltration of inflammatory cells during angiogenesis is a feature of wound healing. Thus, tumours have been described as 'wounds which do not heal.' Tumours recruit inflammatory cells, especially macrophages (tumour-associated macrophages (TAMs)). In some tumours macrophages can represent up to 50% of the cell population and are often the major source of angiogenic molecules in a tumour (see **Figure 6**). Tumour cells express the very potent monotactic chemokine MCP-1 (monocyte chemotactic protein-1) which attracts macrophages into the tumour. TAMs have two important functions within the tumour. First, they confer antitumour effects by their immunological functions and have been implicated as part of the host immune defence mechanism against tumours. Second, those same cells can promote angiogenesis through secretion of cytokines, such as interleukin 1 and TNF-α, chemokines and growth factors (VEGF, bFGF, EGF, PD-ECGF and HGF).

Chemokines act as chemoattractants for haematopoietic cells and confer similar effects on endothelial cells (Moore *et al*., 1998). Chemokines are grouped into two main subfamilies, the CC-chemokines and the CXC-chemokines, based on their structure. In CXC-chemokines the two cysteines near the N-termini are separated by a single amino acid whereas in CC-chemokines these cysteines are adjacent. Endothelial cells express all known receptors for CXC-chemokines (CXCR-1 to CXCR-4 and DARC) but none of the receptors for CC-chemokine (there are to date eight identified receptors). This corresponds well with the observation that CC-chemokines have no direct effects on endothelial cells (members of the CC-chemokine subfamily are MCP-1 and RANTES). The subfamily of CXC-chemokines includes interleukin 8 (IL-8), platelet factor-4, growth-related antigen (GRO-β) and interferon-γ-inducible protein (IP-10). In addition, Il-8 was shown to induce endothelial cell proliferation and to induce angiogenesis *in vivo*. This effect can be potentiated by IL-1 and TNF-α. Interestingly, a number of chemokines from this sub-family have an antiangiogenic function (platelet factor-4, GRO-β, IP-10). These CXC-chemokines have in common that they lack the ELR-motif (glutamine–leucine–arginine) between the first cysteines. The pro- or anti-angiogenic properties of a chemokine might be dependent on this motif.

Macrophages secrete TNF-α, a multifunctional cytokine which has both angiogenic and antiangiogenic properties. These conflicting reports are a result of different actions of TNF-α at high and low doses. Whereas high doses of TNF-α

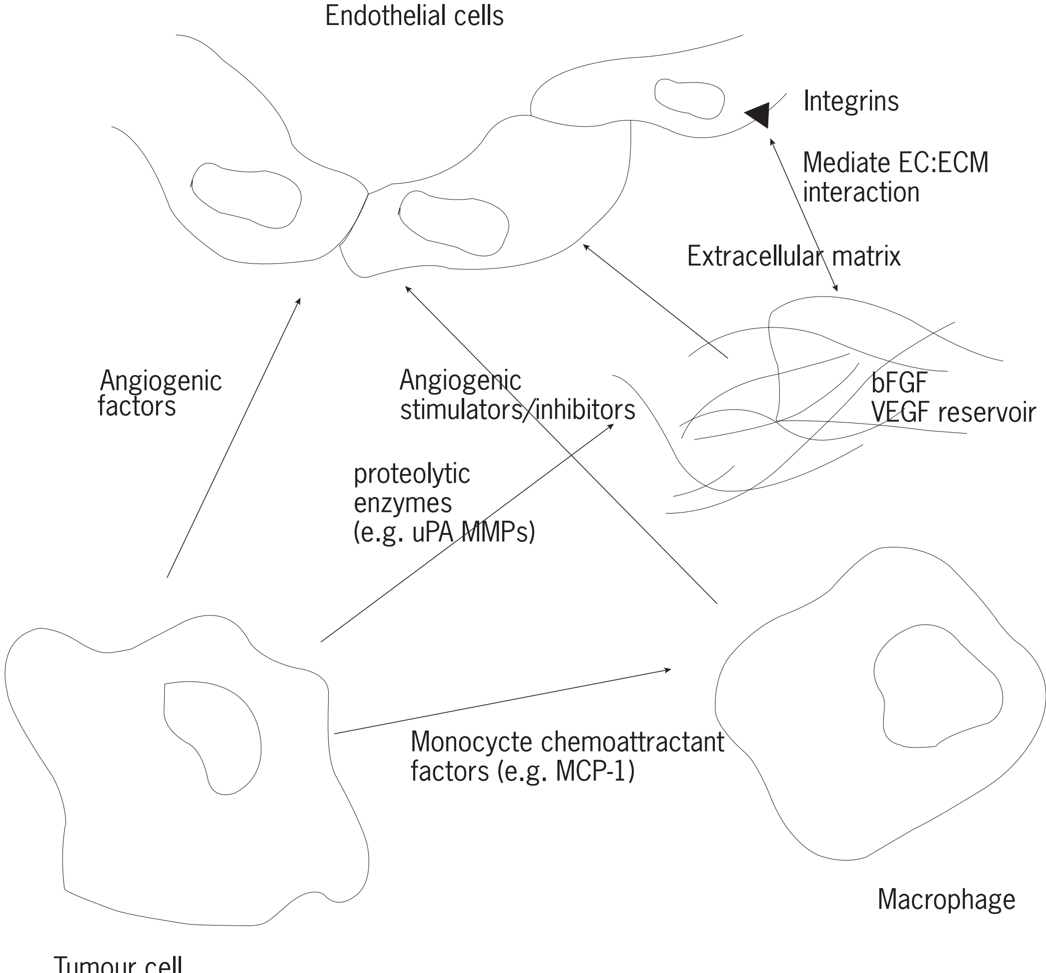


Figure 6 Interaction between tumour cells, endothelial cells and macrophages. Tumour cells produce angiogenic factors, which induce endothelial cells to initiate angiogenesis. Tumour cells also secrete chemoattractants, that recruit macrophages, and proteolytic enzymes, which can release growth factors from the extracellular matrix. Endothelial cells interact with the extracellular matrix by adhesion molecules (e.g. integrins), which mediate endothelial cell migration.

are directly cytotoxic to tumours, it is angiogenic at low doses. Because of its direct cytotoxic effects, TNF-α was studied following systemic administration *in vivo*. Unfortunately, the side effects were severe and may be related to the fact that TNF-α is the major mediator of septic shock. At high doses, TNF-α exerts its antitumour effects mainly through its cytotoxicity on endothelial cells and inhibition of angiogenesis rather than through actions on tumour cells. These findings are consistent with the fact that tumour cells are often insensitive to TNF-α alpha-mediated cytotoxicity.

At low doses, TNF-α induces endothelial cell migration and tube formation *in vitro*, but it also inhibits endothelial cell proliferation. It is chemotactic for leukocytes and induces IL-1 and GM-CSF expression. It also induces the expression of E-selectin, ICAM-1 and VCAM-1 on endothelial cells and can mediate leukocyte recruitment into the tissue stroma.

Fibroblast Growth Factors

aFGF and bFGF are endothelial cell mitogens and stimulate endothelial cell migration. They induce the production of proteases by endothelial cells and stimulate tube formation in three-dimensional cell culture systems. They are angiogenic in all *in vivo* assays and are today often used as a positive control in angiogenesis assays. Because of their heparin-binding abilities, FGFs are bound to the extracellular matrix and are released by proteases involved in the angiogenic cascade. Endothelial cells are not only reactive to FGFs through two FGF receptors but they also produce bFGF themselves. In this way, endothelial cells respond to FGF in both paracrine and autocrine manner. The FGFs interact with almost all molecules involved in angiogenesis. A complete discussion of this growth factor is beyond

the scope of this chapter (for a detailed review, see Christofori, 1997).

THE BREAKDOWN OF THE BASEMENT MEMBRANE

Angiogenesis is an invasive process involving migration of endothelial cells into the surrounding tissue. In the first instance, this requires the degradation of the basement membrane of the pre-existing vasculature, to enable endothelial cells to leave the organized structure of the vessel wall. Several proteolytic enzymes are involved in the degradation of the extracellular matrix (ECM). They include the plasminogen activator system and the matrix metalloproteinases (MMPs).

The Plasminogen Activator System

The plasminogen activator (PA) system includes the urokinase plasminogen activator (uPA) and the tissue-type plasminogen activator (tPA), both of which convert plasminogen into plasmin. The major physiological function of plasmin is to degrade fibrin; however, it also degrades components of the ECM such as laminin and collagen. In addition, it also activates metalloproteinases and elastase. Whereas tPA functions during fibrinolysis, uPA functions mainly in angiogenesis. uPA is secreted as a proenzyme and binds to its receptor on the cell surface, where it becomes activated. By binding to the uPA receptor it also activates downstream signals, resulting in induction of cell migration and invasion. In fact, the presence of the uPA receptor seems to be critical for cell migration because very little migration could be observed in uPA receptor negative cells. Endothelial cells express uPA as well as an uPA inhibitor (PAI). PAI binds to the active uPA and stimulates its internalization into cells following binding to the uPA receptor, resulting in its breakdown. The expression of uPA and its receptor can be induced by FGF-2. This effect is mediated via the FGF receptors. Hence it seems likely that the FGF-2-induced migration in endothelial cells is mediated by uPA. Because of its heparin-binding ability, FGF accumulates in the ECM and is released by proteases including plasmin. Therefore, uPA and FGF increase their function in an autocrine manner because uPA is expected to release FGF-2 and at the same time FGF-2 induces migration through uPA.

The Matrix Metalloproteinases

The MMPs are a family of extracellular endopeptidases which are secreted by a variety of cells including epithelial cells, fibroblasts and inflammatory cells. Endothelial cells express MMP-1, MMP-2, MMP-9 and the membrane-associated MT-1-MMP (for a review, see Stetler-Stevenson, 1999). These endopeptidases are secreted as inactive proenzymes and need to be activated by cleavage of a proteolytic fragment. The substrates of the different endopeptidases include all known components of the ECM but the specific substrate of each enzyme has not yet been identified. Four different endogenous inhibitors of MMPs have been identified; they are the tissue inhibitors of metalloproteinases TIMP-1 to -4. Endothelial cells studied in a two-dimensional culture produce only very low levels of MMPs, whereas expression can be induced in endothelial cells by growing them in three-dimensional collagen gel systems. In addition, changes in cell shape also alter MMP expression (especially mechanical stress) and it has been suggested that integrins, as the mediators of cell–matrix interactions, are involved in the regulation of MMP expression. This has been shown for the $\alpha 2\beta 1$ and the $\alpha v\beta 3$ integrins. The $\alpha v\beta 3$ integrin interacts with MMP-2 through the C-terminal hemopexin-like (PEX) domain of MMP-2. This is interesting because the PEX domain does not contain the integrin binding RGD motif. The PEX domain is also the binding site for TIMP-2. TIMP-2 binding of MMP-2 is necessary for the MT-1-MMP-mediated activation of MMP-2 from the proenzyme. So PEX can inhibit angiogenesis by direct competition to MMP-2 in two different ways. First, PEX inhibits activation of MMP-2 by binding to TIMP-2. Second, PEX binds to the integrin on the cell surface, which would be a prerequisite for MMP-2 activation. Natural breakdown products of MMP-2 are generated and PEX accumulates in tissues where neovascularization occurred. These findings suggest that endogenous PEX acts as a natural inhibitor of MMP-2 function and as a result is an endogenous inhibitor of angiogenesis.

Angiogenesis Inhibitors Encrypted within Larger Proteins

Similarly to PEX, several other endogenous angiogenesis inhibitors have been characterized which are all fragments of larger molecules. These inhibitors are formed by proteolytic breakdown of their parent molecules. The first of these inhibitors was identified as early as 1985 and is a 29-kDa fragment of fibronectin. It has been shown to inhibit endothelial cell proliferation. Fibronectin itself is an abundant molecule which has no such inhibitory function.

More recently discovered endogenous angiogenesis inhibitors are angiostatin and endostatin (for reviews, see Kim Lee Sim, 1998; Oehler and Bicknell, 2000). Angiostatin is a 36-kDa fragment of plasminogen which is cleaved by macrophage-derived MMPs or proteases derived from tumours (e.g. prostate carcinomas). Endostatin is a 20-kDa fragment of collagen XVIII. Systemic administration of both angiostatin and endostatin induces tumour regression by destruction of the tumour vasculature. They also inhibit the formation of new blood vessels in a tumour. They therefore maintain a tumour in a state of

dormancy. Both molecules have been identified in the urine or serum of tumour-bearing mice. In these animals a primary tumour existed but metastases did not occur. Based on these findings, and on the clinical observation that primary tumours suppress the growth of metastasis, it was concluded that inhibitory factors are produced by the primary tumour. The inhibition of metastatic spread was a direct result of angiogenesis inhibition, and angiostatin and endostatin act specifically on endothelial cells. In addition, both molecules are endogenous factors which are generated by proteolytic cleavage of larger molecules. These 'mother' molecules have no inhibitory function but have angiogenesis promoting effects.

N-Terminal fragments of the human prolactin/growth hormone family are antiangiogenic whereas the hormones themselves are angiogenic. They have also been shown to stimulate the expression of PAI. During the maturation of type I collagen, a ~90-kDa N-terminal fragment is cleaved which is homologous to a domain of the antiangiogenic matrix molecule thrombospondin. The fragment has anti-angiogenic effects both *in vitro* and *in vivo*. Restin is a 22-kDa fragment of collagen XV that has been found by homology search with endostatin. It inhibits endothelial cell migration *in vitro* but does not effect proliferation. Vasostatin inhibits bFGF-induced angiogenesis *in vitro* and *in vivo.* It is a 180-kDa fragment of calreticullin. Calreticullin has been shown to exert similar effects.

It seems that endogenous antiangiogenic factors are released under physiological conditions. It might be true that the vasculature is kept in a quiescent state by a balance between angiogenesis-promoting factors and angiogenesis inhibitors and that a deregulation of this balance induces angiogenesis (endogenous angiogenesis inhibitors; see **Table 2**).

Table 2 Naturally occurring angiogenesis inhibitors

Enzyme inhibitors
Tissue metalloproteinase inhibitors (TIMP1-4)
Plasminogen activator inhibitor (uPAI)

Angiogenesis inhibitors encrypted in larger proteins
29-kDa fragment of fibronectin
16-kDa fragment of prolactin
Angiostatin: 36-kDa fragment of plasminogen
Endostatin: 20-kDa fragment of collagen type XVIII
90-kDa N-terminal fragment of collagen type I, homologue to a thrombospondin domain
Cleaved conformation of antithrombin
PEX: hemopexin-like domain of MMP-2
Vasostatin: N-terminal domain of calreticulin
Restin: 22-kDa fragment of collagen XV

Cytokines
Tumour necrosis factor-alpha (TNF-α)
Interferon

Chemokines
Platelet factor-4
Growth-related antigen (GRO-β)
Interferon-γ-inducible protein (IP-10)

MIGRATION

During the formation of new vascular sprouts, endothelial cells migrate into the surrounding tissue following chemoattractant stimuli. Migrating endothelial cells adapt their shape and become elongated with multiple pseudopodia. They cover a larger surface and are therefore in more intimate contact with components of the ECM. Consistent with this, they upregulate the expression of receptors which mediate cell–ECM contacts. Integrins are the major ECM receptors. They are transmembrane heterodimeric cell adhesion molecules which are composed of an α and a β subunit. There are at present 20 members of this family derived from 15 α and eight β chains. Integrins mediate adhesion to all known components of the ECM while one of them can often recognize more than one ECM molecule. Integrins also bind to cell surface molecules of the immunoglobulin superfamily (ICAM-1, ICAM-2 and VCAM-1) and in this situation are involved in cell–cell adhesion processes.

Some integrins recognize ECM molecules through the so-called RGD motif (Arg-Gly-Asp), which is present in fibronectin, vitronectin, proteolysed collagen, fibrinogen, von Willebrand factor, osteopontin and thrombospondin. The $\alpha_5\beta_1$, $\alpha_{\mathrm{IIb}}\beta_3$ and most $\alpha_{\mathrm{v}}\beta$ integrins bind the RGD motif. Some integrins, such as $\alpha_2\beta_1$, bind to an Asp-Gly-Gly-Ala sequence in type I collagen. Both $\alpha_2\beta_1$ and $\alpha_1\beta_1$ integrins are upregulated on migrating endothelial cells following treatment with VEGF and the migration of endothelial cells on a collagen matrix can be inhibited by antibodies against the α chains of these two integrins. An *in vivo* model of angiogenesis was inhibited by combinations of the same antibodies without effects on non-angiogenic vessels.

The most extensively studied integrin involved in angiogenesis is the $\alpha_{\mathrm{v}}\beta_3$ integrin (Eliceiri and Cheresh, 1999). It is upregulated on tumour blood vessels, during wound healing and retinal neovascularization. Angiogenesis is inhibited by a specific monoclonal antibody (LM609) which blocks binding of this integrin and thus disrupts endothelial cell–ECM adhesion. In the chick chorioallantoic membrane assay and in the rabbit cornea eye pocket assay LM609 blocked bFGF-induced angiogenesis whilst VEGF-induced angiogenesis was inhibited by antibodies against the $\alpha_{\mathrm{v}}\beta_5$ integrin. A humanized form of the antibody LM609 (Vitaxin) is in clinical trials.

The RGD peptide itself is able to induce blood vessel regression *in vitro* and *in vivo* through competition with matrix proteins for their integrin binding site. Adhesion is required for endothelial cell survival and these cells undergo apoptosis when adhesion is disrupted. Endothelial

cells of quiescent vessels are less sensitive to antiadhesive treatment than angiogenic and migrating cells. Experiments with endothelial cells grown on RGD-coated beads showed that binding to the ligand (in this case the RGD peptide on the bead) is not sufficient to prevent endothelial cells from undergoing apoptosis. It was suggested that spreading of endothelial cells is necessary for survival. It has also been postulated that the mechanical force generated by cell–ECM interactions mediates gene expression by changes in the cytoskeleton and nuclear morphology.

Integrins interact with the cytoskeleton through their cytoplasmic domains. This interaction is initiated by binding of integrins to the ECM. The mechanical properties of the ECM modulate the strength of this interaction. In fibroblasts the binding of integrins to the cytoskeleton enables the cells to pull themselves forward over a stationary substrate by retrograde movement of the cytoskeleton. This mechanism allows them to migrate. Similarly, integrins are involved in endothelial cell migration.

Integrins also act as signalling receptors. Binding to ECM has been shown to result in elevated intracellular calcium levels, elevated pH, activation of the inositol and DAG pathways and tyrosine phosphorylation. Furthermore, ligand binding to $\alpha_v\beta_3$ is essential for the sustained activation of MAP kinases by angiogenic factors. Tyrosine phosphorylation of VEGFR-2 by VEGF was enhanced by growing endothelial cells on vitronectin, a ligand for $\alpha_v\beta_3$. This was inhibited by antibodies directed against the β_3 integrin subunit. As mentioned earlier, $\alpha_v\beta_3$ integrin binds to MMP-2 and is involved in proteolysis.

PROLIFERATION

Another important step in angiogenesis is the proliferation of endothelial cells. This would provide new vasculature with additional endothelial cells to cover the inner surface of the new vessels. Indeed, many angiogenic factors have mitogenic activity, e.g. bFGF and VEGF.

The turnover rate of endothelial cells in the adult human body is extremely low. Incorporation experiments with radioactively labelled thymidine (H^3TdR) in rabbits *in vivo* estimated a proliferative index for endothelial cells in retinal vessels of about 0.01% (1 h after injection). Following two days of treatment with H^3TdR, about 0.2% of the endothelial cells showed incorporation of thymidine and a turnover time of ~1000 days was calculated. Interestingly, pericytes were discovered to have an even lower proliferation index (0.06%). In contrast, rapidly renewing tissues were shown to have much higher labelling indices, such as the epithelium of the cornea (3%) and the epithelium of the duodenum (14%). In a separate study, endothelial cell proliferation indices of different tissues were compared. Endothelium of the brain and muscle was especially quiescent (0.8% and 0.5% of cells incorporated H^3TdR after 3 days), whereas higher indices were measured in the liver endothelium (4.4% after 3 days).

In contrast to endothelial cells within normal tissues, endothelial cell proliferation indices in tumours is several orders of magnitude higher. Labelling indices in experimental tumours in animals have been reported as up to 32% (anaplastic sarcoma of Wistar rats). Others have estimated proliferation indices between 4.5 and 20%. There was, however, no correlation between the tumour growth rate and the endothelial cell proliferation index. Some studies have reported higher proliferation indices for tumour cells than for endothelial cells (35–11.4%), whereas others have shown the opposite. In general, however, the endothelial cell proliferation is higher in the tumour periphery than in the centre. Artefactual results caused by inadequate perfusion of the tumour centre can be excluded because tumour cells were efficiently labelled.

Endothelial cell proliferation indices in animal tumour models differ significantly from those obtained from human tumours. Endothelial cell proliferation is considerably slower in human tumours than in experimental tumours in animals (2.2–2.8% in breast carcinomas, 5% in gliomas, 0.1–0.6% in prostatic carcinomas). However, these endothelial cell proliferation indices are still significantly higher than in benign human tissues (benign breast tissue, 0.06%; benign hyperplasia in the prostate, 0.023%). These observations argue for active proliferation of endothelial cells during tumour-induced angiogenesis.

The discrepancy between human and experimental tumours becomes obvious if one looks at the rate of growth of experimental tumours. Within 1 week these tumours grow to about 20% of the body weight of the animals, a rate unthinkable in human tumours. Nonetheless, the proliferation indices for endothelial cells in human tumours are 30 times higher than the proliferation indices for normal tissues.

VESSEL FORMATION AND MATURATION

In the final stage of angiogenesis, vessel assembly and maturation result in a vessel that is a stable conduit for blood flow. This is achieved by two processes, namely anastomosis of the developing sprouts and recruitment of pericytes.

Subsequent to the degradation of the basement membrane and outgrowth of vessel sprouts, the sprouts develop into hollow tubes and two sprouts may then join to form a tube through which blood flows. Once formed, the endothelial tubes recruit a layer of pericytes to surround and stabilize the vessel. Pericytes are recruited as local mesenchymal cells, and induced to differentiate into pericytes by the endothelial cells. These cells are large with a prominent nucleus and multiple long processes that embrace the vessel. They also secrete factors, such as

TGF-β, which stabilize the endothelial cells and prevent their proliferation. Furthermore, pericyte density is negatively correlated with the turnover time of the endothelial cells in various tissues, consistent with the stabilizing actions of these cells.

In addition to the pericyte-derived vessel-stabilizing factors, angiopoietin-1 (Ang-1) is a local tissue-derived vessel-stabilizing factor that acts on the receptor Tie-2. Tie-2 is a receptor–tyrosine kinase and is mainly expressed on endothelial cells. The closely related ligand angiopoietin-2 (Ang-2) acts as an antagonist to Ang-1 on this receptor. The related receptor–tyrosine kinase Tie-1 is mainly expressed on endothelium during embryogenesis but downregulated in adulthood. Its ligands have so far not been identified. Targeted disruption of the *Tie-2* gene in mice results in embryonic death between days 9.5 and 10.5 and these animals show heart defects and disorganized vessels on sites where vessels form by vasculogenesis. The endothelial cells appear rounded and have only weak connections with their mural cells. A similar but milder phenotype was developed in *Ang-1* −/− mice which die on embryonic day 12.5. In addition mice overexpressing Ang-2 show very similar phenotypes to *Tie-2* and *Ang-1* −/− mice. These observations show that the angiopoietins and Tie-2 are factors controlling the 'tightness' of vessels and that indeed the angiopoietins have antagonistic effects. *Tie-1* −/− mice die between embryonic day 13.5 and birth as result of the loss of vascular integrity. These animals show 'electron light' endothelial cells because of numerous intra- and intercellular holes which allows blood and plasma extravasation.

During development Ang-1 and Ang-2 are expressed throughout the vasculature, both by the developing endothelium and by its supporting cells, with Ang-2 antagonizing the vessel-stabilizing and maturation function of Tie-2, allowing vasculogenesis and angiogenesis. Ang-1 continues to be expressed into adulthood, but Ang-2 in only expressed in areas that undergo vascular remodelling, such as the female reproductive system. Indeed, expression of Ang-2 appears to be permissive of angiogenesis, suggesting a requirement for 'dematuration' to initiate vessel sprouting (Lauren *et al.*, 1998). Vessels which are destabilized by Ang-2 can respond to VEGF and initiate angiogenesis. Without an angiogenic stimulus blood vessels influenced by Ang-2 will regress (Holash *et al.*, 1999). Only vessels where no pericytes support the endothelial cells can undergo regression whereas the contact of endothelial cells with pericytes prevents regression.

THE PROGNOSTIC SIGNIFICANCE OF TUMOUR ANGIOGENESIS

This chapter has presented evidence that tumour growth is angiogenesis dependent and in the chapter on *Antiangiogenic Therapy* it is shown that inhibition of tumour angiogenesis is currently one of the most exciting avenues of anti-cancer therapy. To conclude this chapter, we will review the prognostic significance of tumour angiogenesis. This encompasses primarily three areas of interest: (1) analysis of the primary tumour microvessel density, (2) expression of angiogenic factors within tumours and (3) analysis of systemic markers of active angiogenesis.

Intratumoural Microvessel Density

In 1991 it was shown that the presence of vascular hotspots, that is, areas of high vascular density, within primary human breast tumours correlated with poor patient survival. This was the first such study linking tumour vascularization, or the end result of angiogenesis, with prognosis. Subsequently, many similar studies have appeared showing a correlation between the presence of vascular hotspots and prognosis in a wide range of different tumour types. Making use of this information in the clinic has, however, proved difficult. To date no simple, quick and reliable way to quantitate the vascular density has appeared. All methods require examination of slides by an experienced pathologist, which is time consuming, expensive and potentially subjective. Nevertheless, proof of the correlation between angiogenesis and prognosis encourages the search for less subjectively quantitated molecular markers of tumour angiogenesis that could substitute for the assessment of vascular density. Such markers will no doubt be found in time.

Systemic Markers of Active Angiogenesis

Angiogenesis is an active process occurring within the tumour, but (excluding women in the menstrual cycle) in the healthy individual not elsewhere in the body, thus metabolic changes arising as a result of the angiogenesis could in principle provide a diagnostic test of tumorigenesis. Potential markers include plasma levels of angiogenic factors and markers of activated endothelium. Amongst angiogenic factors, VEGF has received the most attention as a potential prognostic marker. Thus, studies have suggested that serum VEGF could be a predictor of relapse-free survival in primary human breast cancer. Recently, high urinary VEGF has been shown to correlate with recurrence in bladder cancer, providing an easily accessible marker with which to monitor the disease. Thymidine phosphorylase (platelet-derived endothelial cell growth factor) was shown to be elevated in cancer patients some 15 years before it was recognized to have angiogenic activity and to be elevated in virtually all primary tumours. No doubt more studies will follow.

Other markers of vascular activity are the endothelial leukocyte adhesion molecules VCAM and E-selectin. VCAM and E-selectin mediate leukocyte exit from the

blood through the endothelium. Both molecules are upregulated on endothelium in inflammatory sites. This is also widespread in tumour endothelium where tumours are frequently awash with inflammatory cytokines. Several studies are now examining whether the presence of VCAM or E-selectin in the plasma are useful markers of tumour angiogenesis.

REFERENCES

Auerbach, W. and Auerbach, R. (1994). Angiogenesis inhibition: a review. *Pharmacology and Therapeutics*, **63**, 265–311.

Augustin, H. G. (1998). Antiangiogenic tumour therapy: will it work? *Trends in Pharmacological Science*, **19**, 216–222.

Beck, L. and D'Amore, P. A. (1997). Vascular development: cellular and molecular regulation. *FASEB Journal*, **11**, 365–373.

Christofori, G. (1997). The role of fibroblast growth factors in tumor progression and angiogenesis. In: Bicknell, R., *et al*. (eds), *Tumour Angiogenesis*, 201–237 (Oxford, Oxford University Press).

Eliceiri, B. P. and Cheresh, D. A. (1999). The role of the αv integrins during angiogenesis: insight into potential mechanisms of action and clinical development. *Journal of Clinical Investigation*, **103**, 1227–1230.

Folkman, J. (1971). Tumor angiogenesis: therapeutic implications. *New England Journal of Medicine*, **285**, 1182–1186.

Folkman, J. (1990). What is the evidence that tumor growth is angiogenesis dependent? *Journal of the National Cancer Institute*, **82**, 4–6.

Folkman, J. and Shing, Y. (1992). Angiogenesis. *Journal of Biological Chemistry*, **267**, 10931–10934.

Hanahan, D. and Folkman, J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, **86**, 353–364.

Holash, J., *et al*. (1999). New model of tumor angiogenesis: dynamic balance between vessel regression and growth mediated by angiopoietins and VEGF. *Oncogene*, **18**, 5356–5362.

Kim Lee Sim, B. (1998). Angiostatin and endostatin: endothelial cell-specific endogenous inhibitors of angiogenesis and tumour growth. *Angiogenesis*, **2**, 37–48.

Lauren, J., *et al*. (1998). Is angiopoietin-2 necessary for the initiation of tumor angiogenesis? *American Journal of Pathology*, **153**, 1333–1339.

Moore, B. B., *et al*. (1998). CXC chemokines mechanism of action in regulating tumor angiogenesis. *Angiogenesis*, **2**, 123–134.

Neufeld, G., *et al*. (1999). Vascular endothelial growth factor (VEGF) and its receptors. *FASEB Journal*, **13**, 9–22.

Oehler, M. K. and Bicknell, R. (2000). The promise of anti-angiogenic cancer therapy. *British Journal of Cancer*, **82**, 749–752.

Rissau, W. (1997). Mechanisms of angiogenesis. *Nature*, **386**, 671–674.

Stetler-Stevenson, W. G. (1999). Matrix metalloproteinases in angiogenesis: a moving target for therapeutic intervention. *Journal of Clinical Investigation*, **103**, 1237–1241.

FURTHER READING

Bicknell, R., *et al*. (eds) (1997). *Tumour Angiogenesis*. (Oxford, Oxford University Press).

Chapter 18

Cell Proliferation in Carcinogenesis

Nicholas A. Wright
Imperial Cancer Research Fund and University of London, London, UK

CONTENTS

- Introduction
- How do Tumours Begin in Humans?
- How do Clones Establish Themselves?
- How do Established Tumours Grow?
- Cell Proliferation in Prognosis and Cancer Therapy
- Conclusions

INTRODUCTION

Investigators in the field of cancer research have always regarded cell proliferation, *a fortiori*, as of paramount importance. The reasons for this are not difficult to identify. It is generally accepted that tumours arise as a result of a series of mutations occurring in a cell, often said to be a stem cell. In the colorectal epithelium, for example, several of the mutations required for malignant transformation have been identified; such a series of mutations accumulate in a single cell and its progeny, and this single cell, having acquired properties which endow it with characteristics ensuring its growth and survival, undergoes a series of divisions which eventually result in the development of a neoplasm. A similar series of molecular events is envisioned for the development of other tumours such as the lung, gastric carcinoma and skin. These mutations are thought to confer upon the transformed cell an advantage which enables it to survive – as a mutated clone – and replace the normal cells in the tissue, eventually establishing itself as a neoplasm.

Thus neoplasms are clonally derived, that is, the single cell transformed is the ancestor of all cells which compose the neoplasm. Once established, a mechanism must ensure the growth and propagation of the mutated clone within the epithelium. It is often said that the mechanism lies in an inherent ability of the mutated clone to outgrow its normal counterparts – the 'carcinogenic advantage' is thus a proliferative advantage.

A second reason for the putative importance of cell proliferation in this context is the spread of the mutated clone as it evolves and establishes itself within the host tissue. Third, when a mutant clone has indeed established itself, we observe the phenomenon of tumour growth as the lesion becomes macroscopically evident. During this phase, cell proliferation, in the context of cell birth and cell loss, defines the rate of growth of the tumour.

Finally, when such a lesion is treated, there have been attempts to predict the outcome by the components of cell proliferation.

We shall examine these proposals in turn. Central to our initial proposal is the concept that tumours are indeed clonally derived.

HOW DO TUMOURS BEGIN IN HUMANS?

We often say that tumours arise from a single cell, and therefore form clonal populations. 'If a proliferation is clonal it is a neoplasm' is a statement often heard. Tumour clonality is an important concept in our attempts to understand malignant transformation, and it is worthwhile reminding ourselves of the experimental basis of this proposal. Why do we believe that epithelial tumours, such as the early adenoma in the colonic mucosa, are clonal proliferations? The main methods which have been used for the analysis of clonality in human tumours have been based on X-chromosome inactivation analysis and the detection of somatic mutations. Viral integration, e.g. by Southern blotting in Epstein–Barr virus (EBV)-associated tumours or in hepatitis B- or C-associated liver tumours, are also useful; there is excellent agreement between X-inactivation and EBV integration in nasophayrngeal carcinoma, for example.

In early X-chromosome inactivation studies, the haplotypes of glucose-6-phosphate dehydrogenase (G6PD) were used (Beutler *et al*., 1967; Fialkow, 1976), replaced more recently by methods based on restriction length polymorphisms of X chromosome-linked genes such as glycerophosphate kinase (PGK), the androgen receptor gene (HUMARA), hypoxanthine phosphoribosyltransferase (HPRT), the M27β probe for *DXS285* and p55 and glucose-6-phosphate dehydrogenase. Early on in embryogenesis, genes on one of the two X chromosomes are

randomly inactivated by methylation of cytosine residues within promoter regions; once methylated, such CpG islands are functionally and heritably inactive and it is usually believed that this inactivation is stable, even during malignant change. Thus in approximately half of the cells of the embryo the paternal X chromosome is active, and in the rest it is the X chromosome from the mother. The pattern of fragments produced by DNA digestion with a methylation-sensitive enzyme such as SnaB1 and a further endonuclease corresponding to a restriction fragment length polymorphism, in PGK–BstCX1, for example, can be used to investigate the clonality of any tissue specimen. Informative cases in woman using these markers are reported to vary from 45% with PGK and HUMARA to over 90% with M27β/DXS255.

But can we be certain that the results from such studies will be reliable? Well, the methylation pattern of DNA can be abnormal in malignancy, with both increases and reductions in methylation, and the possibility exists that X-chromosome inactivation may not be valid as an indicator of clonality because of such abnormalities in DNA methylation (Jones, 1996). Moreover, it is possible that X inactivation might be nonrandom, being either constitutive or cell-type specific. Studies in normal haematopoietic and lymphoid tissues have shown skewed X inactivation, possibly favouring the paternal or the maternal X chromosome, which could indicate a nonrandom X-chromosome inactivation pattern. Although there are claims that extremely unbalanced inactivation of the X chromosome is an uncommon phenomenon, skewed inactivation is seen in 23% of women with HPRT and PGK and 22–33% with M27β, in peripheral blood and in bone marrow and skin, indicating tissue specificity, and perhaps related to the number of stem cells in the tissue at the time of X chromosome inactivation. If this number is small, it will result in skewing, with increased probability as the stem cell pool size diminishes (Fialkow, 1973). Moreover, in some embryonic tumours such as retinoblastoma and Wilms' tumour, with LOH on 13q and 11p, respectively, show preferential loss of maternal and paternal alleles. This is also seen in sporadic osteosarcoma. Hence X inactivation analysis is not without its problems, and conclusions drawn from it must be viewed critically.

In mutation analysis, finding the same mutation in key genes, such as k-*ras* or *p53* in multiple tissue sample from the same tumour, or from unconnected tumours, has also been said to indicate a clonal origin, but we should note the possibility that the same mutation is induced in separate precursor cells by a single carcinogen, e.g. aflatoxin causes specific *p53* mutations in hepatocellular carcinomas.

The demonstration of heterogeneity of microsatellite instabilities, e.g. in multiple gastric carcinomas, has also been proposed as a marker for polyclonality – origin from more than one cell – but if genome alterations continue to occur at microsatellite loci with evolution of the tumour, with resulting genetic diversity within the same clone, then microsatellite instability will not be an appropriate molecular marker of clonality. The presence of cytogenetically unrelated abnormal clones demonstrated by karyotypic analysis has also been used as evidence of polyclonality, but the existence of such clones might, however, be due to chromosomal rearrangement in non-neoplastic epithelial or stromal cells: cytogenetically abnormal clones are present in apparently non-neoplastic breast lesions such as fibrocystic disease (from which breast carcinoma probably arises). Mutated clones have been reported even in histologically normal breast epithelium, with several at potential tumour-suppressor gene sites, indicating that genetic abnormalities accumulate before pathological changes can be detected.

We should note that these observations involve the biochemical or molecular examination of homogenized tissues, and in human tissues there have been few opportunities to observe directly the clonal development of very early tumours, and this introduces other problems. In development, a clone is a family of cells which derive from a common progenitor, and these remain more or less contiguous throughout the growth of the embryo. A patch, on the other hand, is defined as a group of cells which share the same genotype, contiguous at the moment of consideration, which, for the current argument, share the same X-chromosome inactivation pattern. Clone size and patch size are not strictly equivalent, since multiple clones of the same genotype could contribute to a single patch; similarly, a single clone could be anatomically separated into different patches.

Figure 1 illustrates this concept as it applies to the clonal origin of tumours. A tumour arising from the centre of a patch will be of clonal origin when assessed by the pattern of X inactivation. The only chance of detecting a polyclonal proliferation would be when such a lesion arises from the margin of a patch boundary (Schmidt and Mead, 1990). This is seen in normal mouse epidermis at patch boundaries, where hair follicles appear polyclonal, but of course are clonal within patches. As X inactivation occurs at about the time of implantation, the pool of stem cells is small, $\sim$15 cells for the skin. Skewing towards one parental allele is therefore very possible, and indeed this is seen in human skin specimens, indicating a large patch size. Polyclonal tumours would be commonest at patch boundaries, the incidence being dependent on the size of the patch, and the incidence of such tumours could be small. And if such rare polyclonal lesions are found, because of a large patch size, it is usually explained away by the lesion consisting of more than one tumour, of clonal origin, that have mixed or collided (Farber, 1990, 1997). However, the rarity of such lesions would indeed be expected if as seems likely, X-linked patches are fairly large.

This is well illustrated by an example: for many years, based in X-chromosome inactivation analysis, it has been believed that atheromatous plaques, occurring in the walls of large- and medium-sized arteries, are monoclonal, and

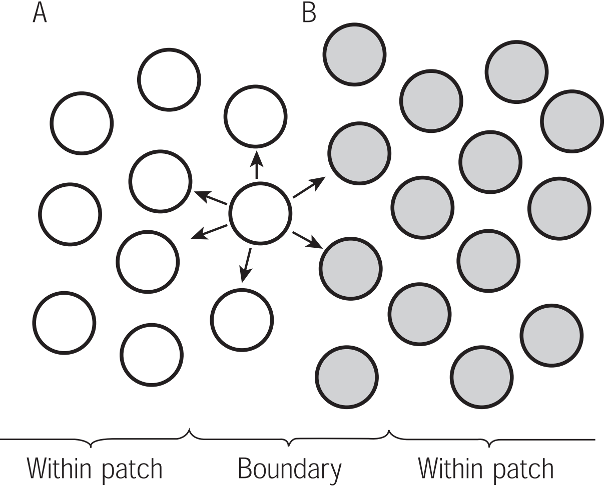


Figure 1 Demonstrating the importance of patch size in the clonal histogenesis of tumours. Different patches are shown, for example, illustrating a different pattern of X-inactivation. In the patches are tissues units, e.g. colonic crypts. A tumour arising within a patch will show a single pattern of X-inactivation of the other marker, and the only chance of detecting a polyclonal tumour is if such a lesion arose from the border of the patch, i.e. between A and B. (From Schmidt and Mead, 1990, *Bioessays*, **12**, 37–40.)

such plaques have even been regarded as some sort of benign neoplasm (Benditt and Benditt, 1973). This clonal proliferation is supported by X-inactivation studies, and was proposed to be due to a somatic mutation, induced perhaps by genotoxic chemicals, or an infection, giving a hit in a single cell, which then develops into an atheromatous plaque (Murry *et al*., 1997). However, by mapping X-chromosome inactivation patterns in human aortic smooth muscle, using the HUMARA method, the patch size in aortic media and intima was found often to exceed 4 mm, indeed a large area in terms of numbers of cells (Chung *et al*., 1998). Because of this large patch size, X-chromosome inactivation analysis cannot distinguish between a monoclonal and a polyclonal origin for atheromatous plaques. This underlines the importance of knowledge of the patch size in such studies.

The study of multiple lesions in the same patient does provide a way of avoiding the problem of patch size, especially if more than one clonal marker is used. Thus the probability that all the tumours examined would have the same X-chromosome inactivation pattern is 0.5^{n-1}, where n equals the tumour number. For example, if allele loss on 9q, an independent event, was also present, then the probability that the same pattern in each tumour is due to chance is $(0.5^{n-1})(0.5^{n-1})$, and so on for each independent marker used.

Similarly, when methods based on the analysis of homogenized tissues give results suggesting polyclonality, this is usually attributed to contamination with underlying stromal cells of different clonal derivation and some results are attributed to normal tissue contamination even when the incidence of polyclonal tumours amounts to 40% of the total. However, it is clear that even the fibroblasts or myofibroblasts very closely applied to the epithelium in the colonic crypt are of different clonal derivation, and there may be invading inflammatory cells, which, if numerous enough, can give rise to disparate results in clonality analyses involving PCR techniques. Normal tissue can be trapped and enclosed by surrounding neoplastic tissue, and again give discordant results. Finally, the problem of poor sampling may give a false impression of monoclonality in a polyclonal tumour.

It is also important to know when, during the life history of the development of the tumour, it is examined. Mature lesions are inappropriate, since tumours of polyclonal origin may become clonal because of clonal evolution (Nowell, 1976, 1986). All clones, except one, are eliminated or reduced to the point of being undetectable, seen in chemically induced mouse fibrosarcomas, initially polyclonal, which evolve to a clonally derived population in time, because of the later selection of a dominant clone (Woodruff *et al*., 1982). This is also seen in colorectal adenomas, where heterogeneity of k-*ras* mutations is observed, which is lost after the lesion has evolved to become an invasive carcinoma (Ajiki *et al*., 1994). As noted above, genetic heterogeneity in a tumour does not necessarily indicate a polyclonal origin, since genetic instability is a major feature of malignant tissues and many new clones may arise during tumour development.

Having critically reviewed the available methods, let us see what they say about how tumours arise in humans.

Non-neoplastic and Preneoplastic Lesions

Many tumours arise from preneoplastic lesions, not themselves neoplastic, which were previously regarded as hyperplasias, which involve changes in many cells and therefore by definition polyclonal. However, such lesions are often clonal proliferations themselves. Extremely relevant is the growing recognition that mutations in important genes such as *p53* are found in tissues such as squamous epithelium preceding any dysplastic change, such as sun-exposed normal epidermis. Clones of keratinocytes with immunoreactive p53 and p53 mutations (exons 5–8) are seen in sun-exposed but otherwise normal skin. This has also been reported in morphologically normal mucosa from individuals with upper aerodigestive tract tumours. Moreover, microsatellite instability has been found in normal mucosa from patients with ulcerative colitis, reflecting the increased risk of malignancy in these patients, since microallelotyping shows no allelic loss in transitional mucosa adjacent to colorectal neoplasms.

However, such losses have been reported in normal breast tissue.

In the gastric mucosa of a single patient, three separate hyperplastic polyps of the fundus have been found which harbour foci of dysplasia showed the same k-*ras* codon 12-point mutation, present in both hyperplastic and dysplastic areas. An explanation is that the progeny of a single transformed cell spreads through the mucosa: either surface spreading and surface implantation has occurred or a single cell could populate a gastric gland, which then spreads through the mucosa by gland fission.

In the female ovary, malignant tumours, endometrioid carcinomas, are thought to arise from foci of endometriosis and the majority of such endometriotic foci are clonal (Jimbo *et al*., 1997). Endometriosis arises either from implantation of shed endometrial cells, or from metaplasia of the pelvic peritoneum: if patch size considerations can be excluded, this means from a single endometrial or mesothelial cell. Endometriotic foci can show aneuploidy and loss of heterozygosity at candidate tumour-suppressor loci in 9p, 11q and 22q. The derivative tumours, the endometrioid carcinomas, are clonally derived, arising from a monoclonal proliferation that itself can show genetic defects.

The preneoplastic lesions from which breast cancers develop – proliferative breast disease and similar lesions – show cytogenetic abnormalities indicating the presence of multiple clones. X-chromosome inactivation studies and detection of microsatellite alterations both show that atypical duct hyperplasia and intraduct papillomas appear clonal proliferations and consequently cytogenetic alterations have already occurred at this stage (Lakhani *et al*., 1995).

Naevocellular naevi, either congenital or acquired, are, in some cases, the precursor lesions of malignant melanomas, are apparently polyclonal lesions on X-chromosome inactivation analysis (HUMARA and PGK combined), but malignant melanomas are clonal. So such naevi are presumably hamartomas – an abnormal proliferation of cells in an organ or tissue where these cell types would normally be found. However, hamartomas in tuberous sclerosis, where multiple cell types are seen, show clonal 9q34 or 16p13.3 LOH and clonal X-inactivation patterns, while pulmonary chondroid hamartomas also contain clonal cytogenetic abnormalities in the chondroid component.

In the human liver opinion is divided concerning the clonality of lesions often regarded as preneoplastic: some maintain that while lesions such as benign adenomatous hyperplasia and focal nodular hyperplasia are polyclonal, but hepatic adenomas and even small (<25 mm) hepatocellular carcinomas are clonal. Others, examining the integration patterns of hepatitis B virus, claim that atypical adenomatous hyperplasisa and focal nodular hyperplasias are clonal. Hepatocellular carcinoma is frequently multifocal and whether these arise from a single clone or independently is controversial, but most authors agree that an independent origin is more likely. Thus, after liver damage, clonal selection occurs during regeneration, leading to the genesis of persistent benign focal proliferations, which may be themselves clonal. This is followed by the development of clonal hepatocellular cancer from one or more such nodules. There is also substantial evidence from hepatitis B and C virus integration that between 0.5 and 43% of regenerative nodules in the resulting cirrhosis are monoclonal (Aihara *et al*., 1994), whereas in other types of cirrhosis, some 54% of regenerative liver nodules are clonal, but that the associated hepatocellular carcinomas are clonal by X-inactivation analysis, and differ from the nodules by 18q loss. This suggests that regenerative liver nodules showing a polyclonal pattern evolve into a clonal population, developing into hepatocellular carcinomas, which are also clonal.

Preinvasive Lesions

The field cancerization hypothesis states that multiple cells form independent tumours, since carcinogenic exposure affects multiple cells in the field (Slaughter *et al*., 1953), and predicts that second primary or synchronous tumours arise from independent genetic events. Thus 11% of individuals with oral cancer had multiple upper aerodigestive tract tumours, and multiple invasive foci are associated with overlying areas of *in situ* squamous carcinoma in these lesions. There is now genetic evidence for such an independent origin: in the upper aerodigestive tract, multiple synchronous squamous tumours appear independent and multicentric in origin. However, the concept of clonal origin and expansion is problematic in organs where several metachronous tumours appear; such a synchronous or second primary tumour may indicate recurrence or indeed lateral spread from a single tumour. Thus, although the field cancerization hypothesis would predict a multicentric, polyclonal origin, with the demonstration of a clonal origin for these tumours, lateral migration from the original clone would be a distinct possibility.

There is conflicting evidence for the nature of field cancerization from the study of tumours of the upper aerodigestive tract and adjacent mucosa. In laryngeal and pharyngeal tumours, multiple samples taken at tumour-distant sites show different and independent mutations in the *p53* gene, favouring a discontinuous, multifocal and polyclonal process, rather than migration of premalignant basal keratinocytes giving a clonal development of multiple primary, secondary or recurrent tumours.

In the stomach, discordant mutation patterns of *APC*, *MCC* and *p53* are found in many cases of multiple gastric carcinomas, again in accord with a multicentric origin, and this is also seen in multiple colorectal tumours, parathyroid adenomas and in separate (Emmert-Buch *et al*., 1995) intraepithelial neoplastic lesions, which show different clonal patterns of allele loss at 8p12-21, suggesting an

independent origin. However, multiple synchronous carcinomas in the bladder and other pelvic organs apparently show a common clonal origin – X-chromosome inactivation and allele loss at 9q and 17p are identical, as are c-*erB2* and *p53* mutations in multiple synchronous urothelial tumours (Lunec *et al*., 1992). Multiple serous adenocarcinomas in the ovaries, peritoneum and endometrium show the same p53 mutation and clonal cytogenetic abnormalities, findings confirmed by X-chromosome inactivation. Multiple sites of occurrence of sporadic ovarian cancer on the ovarian surface and pelvic peritoneum are clonal, as assessed by LOH, p53 mutations and X-chromosome inactivation analysis in the same patient (Jacobs *et al*., 1992).

In multifocal breast carcinomas, an increasingly common finding, karyotypically identical clones are detectable, indicating intramammary spread from a single carcinoma either by focal lymphatic spread or by intraductal spread. Moreover, in phyllodes tumours, widely separated deposits show the same monoclonal stromal component.

If lesions some distance away from each other are clonal, how can this be explained? Of course, there is always the possibility, always very difficult to exclude, that the disease process which causes the tumour to develop has a characteristic genetic fingerprint, therefore seen in all examples of the tumour. The other possibility is that of a common mutated progenitor cell, and the expansion of this mutant clone in some way, at an early stage in tumour development. Examination of *p53* mutations in the nonmalignant but dysplastic tracheobronchial mucosa of individuals who smoke shows the same G:C to T:A transversion in codon 245 at multiple sites in both lungs (Franklin *et al*., 1998). Thus the expression and dispersion of a single mutant progenitor bronchial epithelial cell clone throughout the airways is possible, aided possibly by a proliferative advantage. p53 mutations are early events in upper aerodigestive tract carcinogenesis, prior to the development of invasive lesions, being found in premalignant lesions of the head and neck, lung and oesophagus. p53 mutations apparently do not show an increased incidence with cancer progression, but do show clonal fidelity in a variety of tumours. An early event, prior to p53 mutation, might establish a mutated clone, which migrates laterally, possibly aided by a mutation in a cell cycle control gene, or a cell adhesion gene. In multiple bladder cancer, all tumours lose the same 9q allele as an early event (Chung *et al*., 1995, 1996), possibly encoding for a growth control or adhesion molecule, and cells repopulate the urothelium by lateral migration or mucosal seeding. In most discontinuous foci of CIN3 in the cervix, individual lesions show the same X-chromosome inactivation pattern, suggesting intraepithelial spread. In the skin, it is not uncommon to see migration of morphologically abnormal cells laterally from a lesion such as Bowen's disease.

The multiple deposits of Kaposi sarcoma, a widely disseminated malignancy, appear clonal in any one patient, indicating a clonal origin and wide intravascular dissemination, or an initial vascular hyperplasia, with later clonal evolution. Diffusely infiltrating gliomas are clonal, and multiple discrete meningiomas share clonal neurofibromatosis 2 (*NF2*) mutations, while most individual meningiomas appear clonally derived.

Sporadic, multinodular goitres contain nodules which are regarded as being hyperplastic and therefore polyclonal, but there is substantial X-chromosome inactivation evidence that many of these nodules are clonal populations. The presence of a TSH receptor mutation may be pivotal; most cases showed monoclonality on HUMARA analysis, raising the possibility that during thyroid hyperplasia a cell with a mutation at this locus leads to the initiation of autonomous clonal growth. Multiple nodules in the same patient are mostly clonal, with activation of the same allele, indicating intraglandular spread by follicular budding, although clonal nodules with different X-chromosome inactivation patterns can be seen in the same gland. This could mean a different pathogenetic mechanism for clonal and polyclonal nodules, or indeed evolution of clonal from polyclonal nodules, as we have seen above in the liver. Most follicular, papillary and anaplastic carcinomas are clonal. Parathyroid adenomas are monoclonal, as indicated by X-chromosome inactivation and the identification of clonal abnormalities of the parathyroid hormone gene. In MEN1, where all four parathyroids are enlarged and appear hyperplastic, allelic loss on 11q indicates that proliferation is clonal, although when LOH on 11q is combined with X-chromosome inactivation studies, the parathyroid lesions in MEN1 were shown to be polyclonal, suggesting that multiple neoplastic clones grow, coalesce and replace the parathyroid gland. Sporadic parathyroid hyperplasias, which are either a primary phenomenon or secondary to such conditions as chronic renal failure, were previously regarded as polyclonal proliferations, but X inactivation and allelic loss of 11q shows that 38% of primary parathyroid hyperplasias and 64% of hyperplasias secondary to renal failure harbour clonal proliferations. Again we can interpret this as clonal neoplastic evolution from a pre-existing polyclonal hyperplasia. Diffuse parathyroid hyperplasias in uraemia are polyclonal, but the individual nodules in nodular hyperplasia are clonal, indeed showing different clonal patterns of X inactivation in the same gland, once more favouring the hypothesis that monoclonal proliferations evolve from hyperplasias. These findings favour the view that hyperplasia begins in endocrine glands as a polyclonal process, but then becomes a clonal hyperplasia, and the borderline between this phase and that of a benign clonal neoplasm is difficult to delineate.

Invasive carcinomas of the cervix, which are clonal in origin, arise in areas of cervical intraepithelial neoplasia (CIN). Severe dysplasia or CIN3 is also a clonal

proliferation, although lesser degrees of dysplasia (CIN2) appear more commonly polyclonal in X-inactivation studies. In this respect, some vulvar hyperplasia, considered to be preneoplastic lesions in this tissue, appear clonal: the derivative VIN and invasive carcinomas are also clonal proliferations. In nasopharyngeal carcinoma (NPC), combined X-inactivation studies (PGK), X-linked RPLPs, and EBV integration show that carcinomas are mainly clonal, but hyperplastic epithelia, and early atypical hyperplastic epithelia, are polyclonal, and clonality emerges only at the moderate/severe dysplastic stage.

A relationship between intestinal metaplasia and carcinoma of the stomach and lower oesophagus has long been suspected: the non-dysplastic metaplasia adjacent to carcinomas is clonal on X-inactivation analysis (HUMARA), and also shows LOH for *APC*, changes also seen in the dysplastic and neoplastic tissues of Barrett oesophagus. Although it is not yet clear if the *APC* change is clonal, microsatellite analysis shows allelic imbalance not only on 5q but at multiple other sites, in both premalignant and malignant Barrett epithelium, supporting the concept of clonal expansion from metaplasia through dysplasia to carcinoma (Zhuang *et al*., 1996).

In most, but not all, of the above discussion the commonality has been the need to homogenize the tissue, albeit, in some cases, after microdissection. Evidence from studies where direct observation is possible gives disparate results. In a patient with familial adenomatous polyposis, who was also a sex chromosome chimaera (XO/XY), the colon contained hundreds of adenomas, ranging in size from monocryptal adenomas to microadenomas 2.5 mm in diameter; no larger adenomatous polyps were seen (Novelli *et al*., 1996). If an adenoma was of clonal origin, all dysplastic crypts within it would be expected to be entirely XO or entirely XY. Localization of the Y chromosome in tissue sections showed that monocryptal adenomas were entirely XO or XY, with no mixed pattern. However, many adenomas (76%) were polyclonal. Isoenzyme studies of G6PD in black females have shown that colonic adenomas from patients with Gardner's syndrome were polyclonal (Hsu *et al*., 1983), while studies using X-linked RFLPs show that both spontaneous and familial adenomas are clonal (Fearon *et al*., 1987). The minimum size of the adenomas in this might explain the discrepancy, through monoclonal conversion as size increases. Polyclonal derivation of adenomas has also been found in chimaeric mice – made between a *Min* mouse (which has an *APC* mutation) and a *Min*/ROSA mouse (Merrit *et al*., 1997), and in *Apc*$+/^{\mathrm{min}}/+/-$ (De Wind *et al*., 1998) mouse chimaeras. Possible mechanisms of polyclonality in these lesions include 'field' effects causing adenomas to cluster (nonrandom collision), a passive process involving fusion of two or more *APC*-negative clones early in tumour development, but the high frequency of mixed adenomas found in both these studies is inconsistent with a random appearance of *APC*-negative clones in the mucosa, and suggests that some regions of the intestine have an increased potential for initiation. Multiple clones may be required for early adenoma growth, or perhaps early adenomas may induce adenomatous growth in surrounding crypts, especially in FAP since all cells already have a single *APC* mutation and perhaps some derangement of APC function. These latter scenarios imply active cooperation between clones. Indeed, conversion of normal crypts to adenomatous crypts apparently occurs at the margins of FAP adenomas.

What can we conclude? In many tissues the bulk of the evidence indicates that preneoplastic changes occur which are hyperplastic, but polyclonal, indicating increased rates of cell production among many cells. After some time, a genetic event occurs in which a clone of cells appears, which expands and establishes itself. This clone has a 'carcinogenic advantage,' which allows such expansion. Further growth of the clone is associated with genetic evolution and the appearance of mutations, which lead to the development of an early clonal neoplasm. Further clonal evolution develops the invasive phenotype. However, we must concede that further studies, with methods which allow the direct examination of clonality, might provide different conclusions, and we cannot disregard the concept that the earliest lesions are polyclonal, needing the cooperation of several clones, and that monoclonality is itself the result of further clonal evolution.

HOW DO CLONES ESTABLISH THEMSELVES?

The simplest way in which we can conceive of a single mutated cell establishing itself among a population of normal cells is if the mutation selects for an advantage that involves factors which modulate cell proliferation – the ability to divide faster – or the ability to survive longer. While conceptually these are significant advantages to being able to survive longer, or indefinitely, in the genesis of a neoplasm, it is, in practice, very difficult to demonstrate. While some early neoplastic lesions do show abnormal expression of genes involved in apoptosis, this is a far cry from demonstrating a aggregation of cells brought about by reduced cell death. Moreover, simple modelling approaches have suggested that mechanisms involving apoptosis are unlikely to be involved at this stage.

Is there any evidence that very early neoplastic lesions show an increased cell production rate? Using the colon as a paradigm, the earliest lesions detectable, the so-called 'aberrant crypt foci,' show increased labelling indices with a number of markers of cell proliferation, whereas in established adenomas, the rate of cell production also increases as assessed by these markers. The importance of even a minor increase in cell proliferation

can be assessed by examination of how a mutated stem cell – of which there are several in each intestinal crypt – competes with its colleagues for ascendancy, and in some cases succeeds, populating the crypt with its progeny. Intestinal crypts, in both animals and humans, are clonal populations, ultimately derived from a single cell. The evidence from this comes from the analysis of chimaeric mice and mice heterozygous for a defective G6PD gene carried on the X chromosome and randomly expressed after Lyonization. In these chimaeras, crypts are always composed wholly of cells from one component mouse strain (Ponder *et al*., 1985), and with similar findings in the G6PD heterozygote crypts are indeed clonal populations and the stem cell repertoire includes all intestinal crypt and gastric gland lineages. **Figure 2 (see colour plate section)** shows the results of an experiment in which mice showing uniform staining for G6PD are given a single dose of mutagen: in the weeks that follow, crypts appear which are composed of cells with a different, mutated phenotype. There is an induction of a rapid, but transient, increase in crypts which show a partial, or segmented, mutated phenotype (**Figure 2a**) (Park *et al*., 1995). Later, there is an increase in crypts showing a wholly mutated phenotype. The emergence of the partially mutated crypts, and their replacement by wholly mutated crypts, can be explained by a mutation at the G6PD locus in a single stem cell from which all lineages are derived. The partially mutated crypts are crypts in the process of being colonized by progeny from the mutated stem cell, and this crypt will ultimately develop into a wholly mutated crypt.

Approximately 9% of the Caucasian population secrete sialic acid lacking in *O*-acetyl substituents, readily recognized by histochemical staining. This is explained by genetic variability in the expression of the enzyme *O*-acetyl transferase (OAT) and this 9% of the population is homozygous for inactive *OAT* genes, *OAT−/OAT−*. Some 42% of the population are heterozygous, *OAT−/OAT+*, and loss of this gene converts the genotype to *OAT−/−OAT−*. This is indeed seen in about 42% of the population, and again is explained by a mutation or loss of the gene by nondisjunction in a single crypt stem cell and the colonization of this crypt by the clonal progeny of the mutated stem cell.

If we again imagine that a stem cell carries a mutation, not in its *G6PD* or its *OAT* gene, but in a cell cycle gene, increasing its cell production rate, then the monoclonal conversion process will not be stochastic, with equal competition between stem cells, but the mutant stem cell would have an advantage, and would succeed in colonizing the crypt. **Figure 2c** shows a further method in which mutated clones can be propagated – the mutated clone has expanded by crypt fission, where crypts divide to make two daughter crypts, and thus spread through the mucosa. If a gene controlling crypt fission became mutated, then of course the crypt would spread further through the mucosa. Indeed, in conditions such as ulcerative colitis, large areas of the mucosa become occupied by the same aneuploid stem line, conceptually, at any rate, the result of increased crypt fission.

In squamous epithelium we have seen above that tiny p53-positive clones, containing mutated p53, occur early in the carcinogenic process. These clones may owe their existence to prolonged survival, given the presence of a mutated *p53* gene. However, there is substantial evidence that rates of cell production, in preneoplastic lesions such as actinic keratosis, are elevated. In fact, it is possible to predict the shape of preinvasive skin lesions such as actinic keratosis, Bowen disease and lentigo maligna, by assuming different values for the carcinogenic advantage, in terms of increased cell production. Similar observations of elevated rates of cell production have been recorded in other situations where there is a defined morphological sequence of carcinogenesis – in the liver, bladder, oesophagus and lung, for example.

We have seen above that, in many tissues, clonal evolution occurs on a background of hyperplasia, implicitly involving increased cell production. It is further evident that the establishment of the mutated clone is, in many tissues, associated with an increase in cell production rate. It is easy to conceive now such increases in cell proliferation, in cells such as stem cells, can lead to the establishment and development of a mutated clone (**Figure 2**). However attractive the concept seems, we cannot say definitively whether such increased cell proliferation of stem cells is involved. For example, in **Figure 2**, increased longevity, or changes in the adhesion molecule status of the progeny, could give much the same result. An ability to study cell proliferation, and then expression of cell cycle-related genes, in stem cells during carcinogenesis, would be a very useful, if experimentally difficult, prospect.

HOW DO ESTABLISHED TUMOURS GROW?

The basic concepts of cell proliferation in tumour growth have been extensively and well reviewed.

Tumour Growth Curves

There has been extensive analysis of tumour growth curves – obtained by plotting the volume or weight of the tumour against time – in the past, in the (forlorn) expectation that basic truths about the nature of tumour growth might be discovered. Although this has not been the case, some interesting facts have emerged. Many tumours, usually transplantable, growing in experimental animals, initially grow in an exponential manner, that is, after an initial lag phase, there is logarithmic growth (**Figure 3**). However, as growth proceeds, and the

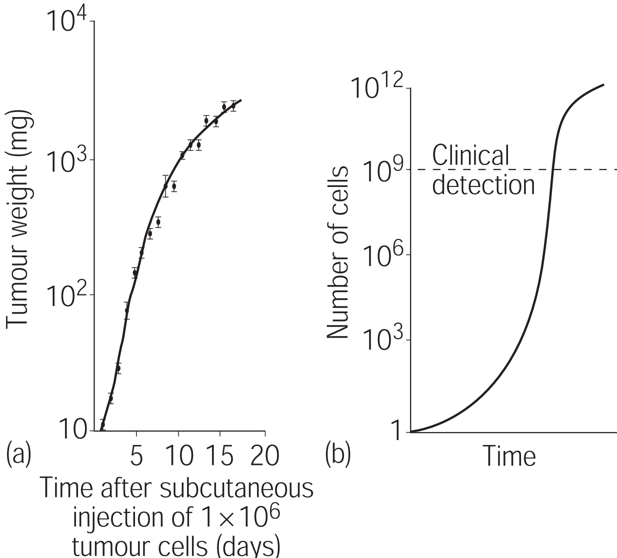


Figure 3 (a) A typical growth curve for a rapidly growing mouse tumour, with the curve fitted by eye; (b) idealized growth curve for a human tumour, showing the initial lag phase, a period of exponential growth and a retardation in the rate of increase of tumour cell number at large size. Note the long latent period before clinical detection.

tumour increases in volume, growth progressively slows. This is the so-called Gompertzian growth curve, where growth retardation is itself exponential. Although such tumours rarely obey such a growth curve in its entirety, it is usually a good approximation. Note that, at the asymptote, growth is barely discernible. Most experimental tumours in rodents grow very rapidly, particularly the transplantable ones, since they have been selected for a rapid growth rate.

In humans, tumours grow much more slowly; most of the doubling in cell number has occurred before the tumour becomes clinically evident. In circumstances where it has been possible to measure the rates of tumour growth, e.g. by serial observation of pulmonary primary or metastatic lesions, a number of tumours do indeed appear to grow exponentially. However, in other cases, growth is extremely variable, with periods of quiescence followed by rapid growth. Some human tumours grow extremely slowly, e.g. colorectal carcinoma has a reported doubling time (the time taken for the tumour to double in size) of over 600 days. Others, such as childhood or embryonal tumours, show doubling times of 10–20 days. There is some evidence that tumour growth rates vary with age; for example, breast carcinomas in women below the age of 50 years show a median doubling time of 80 days, whereas women over the age of 50 years show a doubling time of 140 days.

Kinetic Parameters During Tumour Growth

Why does tumour growth slow down? Well, the net rate of growth is defined by the relationship between the rate of cell production – the birth rate – and the rate of cell death or other modes of loss, the cell loss rate. In the earlier phases of growth, as seen in the early exponential part of the curve in **Figure 3**, most cells are growing exponentially; thus every cell which divides contributes two cells to the population, with effectively no cell loss. What defines the rate at which the exponential portion of the curve ascends?: two parameters – the cell cycle time (the interval between two divisions of the same cell) and the growth fraction, or proliferating population (the proportion of cells in the tumour which are dividing). There is no evidence in solid experimental tumours, at any rate, that the cell cycle time changes at all during tumour growth, but as growth advances, there are substantial studies showing that the growth fraction declines with time. In experiments where the growth fraction has been measured in different parts of the tumour, it is the areas furthest from the blood supply that show the largest reductions in growth fraction – the centre of the tumour in most lesions, or when the tumour is 'corded,' or shows cords of surviving tumour cells surrounding blood vessels – in the distal portions of the cord.

While the decline in the growth fraction is certainly responsible for a proportion of the decline in the rate of tumour growth, by far the largest factor is the considerable increase in the cell loss rate. This is due, of course, to the tumour outgrowing its blood supply and to large numbers of cells in the lesion dying. In fact, at the asymptote in the curve, where the curve levels off, the rate of cell loss almost equals the rate of cell birth, with little resulting in the way of net growth. It is sometimes useful to define a cell loss factor, ϕ, which is the ratio of the cell loss rate to the birth rate. At the asymptote in **Figure 3**, this will be close to 1.

Where data are available, experience culled from animal tumours is also applicable to humans. However, measurements are usually confined to the later phases of tumour growth, since most lesions present, biologically speaking, very late. Thus many carcinomas appear to be low-growth fraction, high-cell loss lesions. In clinically detected colorectal carcinomas, for example, the cell cycle time is 48 h, the growth fraction is 14% and the cell loss factor 98%, i.e. for every 100 cells which are born, two survive. At least this gives us some insight into why such tumours are difficult to treat with modalities such as radiation and cytotoxic chemotherapy. Of the surviving cells in the tumour, some 86% are not in the cell cycle, mainly because they are hypoxic. Such cells would be resistant to irradiation, and even many fractions might not recruit many cells into cycle. Similarly, 5-fluorouracil, commonly used in these lesions, is, in the main, an S-phase poison. With so few cells in cycle, a brisk response would

not be expected, unless the growth fraction is increased by debulking and redistribution.

CELL PROLIFERATION IN PROGNOSIS AND CANCER THERAPY

There has been a great deal of effort and energy expended in attempts to relate measurements of cell proliferation to both prognosis and the prediction of tumour response, and to define subgroups requiring additional or adjuvant therapy. It is probably true to say that such attempts have failed.

It is now some 80 years since measurements of mitotic activity were first routinely made in human tumours, and some 40 years since tritiated thymidine, the first real S-phase-specific marker, became available. Since then, measurements using techniques such as flow cytometry, bromodeoxyuridine labelling and the several antibodies against the Ki67 antigen, have followed. to say nothing of numerous attempts to clone human tumour cells *in vitro* and predict their levels of chemosensitivity against a panel of agents, a method recently resurrected. There have also been attempts to plan chemotherapeutic regimens on the basis of measurements of cell proliferation, e.g. the so-called 'synchronization therapy' for head and neck cancers and acute leukaemias, where doses of cytotoxics are given when the tumour cells are 'synchronized' in a cell cycle phase most sensitive to the agent in question.

It is perhaps salient to ask which of these measurements are generally and routinely used in cancer treatment outside of research protocols. Precious few, is the sobering answer. In tumour diagnosis, pathologists routinely do a formal mitotic count (rather than merely observe the presence of mitoses) in smooth muscle cell tumours of the uterus, some central nervous system tumours and stromal tumours of the gastrointestinal tract. In the grading of tumours, and therefore in the prediction of prognosis, mitotic counts are routinely carried out in a number of circumstances, as in the assessment of grade of invasive ductal carcinoma of the breast as a component of the Nottingham Prognostic Index or sometimes in ovarian carcinomas, either as a raw count as a component of a morphometric index. And have more sophisticated measurements reached routine status? Some advocate the use of Ki67 in the differential diagnosis of difficult melanocytic lesions, but there has been little use of such methods elsewhere. Perhaps the advent of new markers, such as antibodies against the minichromosome maintenance proteins, said to be more sensitive than Ki67, may make a difference.

CONCLUSIONS

In human cancer, it is clear that cell proliferation is involved in the origin and evolution of mutated clones, although the details at present elude us. Cell proliferation studies give us insight into how tumours grow, and in some instances why treatment is so difficult. Measurements of cell proliferation are at present of little routine value in diagnosis and prognosis. It is clear that a great deal of work needs to be done if we are to understand fully the role of cell proliferation in tumour development.

REFERENCES

Aihara, T., *et al*. (1994). Clonal analysis of regenerative nodules in hepatitis C virus-induced liver cirrhosis. *Gastroenterology*, **107**, 1805–1811.

Ajiki, T., *et al*. (1994). K-*ras* gene mutation related to histological atypias in human colorectal adenomas. *Biotechnical Histochemistry*, **70**, 90–94.

Benditt, E. P. and Benditt, J. M. (1973). Evidence for a monoclonal origin of human atherosclerotic plaques. *Proceedings of the National Academy of Sciences of the USA*, **70**, 1753–1756.

Beutler, E., *et al*. (1967). Value of genetic variants of glucose-6-phosphate dehydrogenase in tracing the origin of malignant tumours. *New England Journal of Medicine*, **276**, 389–391.

Chung, G. T. Y., *et al*. (1995). Sequential molecular changes in lung cancer development. *Oncogene*, **11**, 2591–2598.

Chung, G. T. Y., *et al*. (1996). Clonal evolution of lung tumours. *Cancer Research*, **56**, 1609–1614.

Chung, I.-M., *et al*. (1998). Clonal architecture of normal and atherosclerotic aorta. Implications for atherogenesis and vascular development. *American Journal of Pathology*, **152**, 913–923.

De Wind, N., *et al*. (1998). Mouse models for hereditary non-polyposis colorectal cancer. *Cancer Research*, **58**, 248–255.

Emmert-Buch, M. R., *et al*. (1995). Allelic loss on chromosome 8p12–21 in microdissected prostatic intraepithelial neoplasia. *Cancer Research*, **55**, 2959–2962.

Farber, E. (1990). Clonal adaptation during carcinogenesis. *Biochemical Pharmacology*, **39**, 1837–1846.

Farber, E. (1997). Monoclonal and polyclonal development of digestive tract tumors in chimeric mice. *Japanese Journal of Cancer Research*, **88**, 663.

Fearon, E. R., *et al*. (1987). Clonal analysis of human colorectal tumours. *Science*, **238**, 193–197.

Fialkow, P. (1973). Primordial cell pool size and lineage relationships of five human cell types. *Annals of Human Genetics*, **37**, 39–48.

Fialkow, P. J. (1976). Clonal origin of human tumours. *Biochimica Biophysica Acta*, **458**, 283–321.

Franklin, W. A., *et al*. (1997). Widely dispersed p53 mutation in respiratory epithelium. *Journal of Clinical Investigation*, **100**, 2133–2137.

Hsu, S. H., *et al*. (1983). Multiclonal origin of polyps in Gardner syndrome. *Science*, **251**, 951–953.

Jacobs, I. J., *et al*. (1992). Clonal origin of epithelial ovarian carcinoma; analysis by loss of heterozygosity, p53 mutation,

and X-chromosome inactivation. *Journal of the National Cancer Institute*, **84**, 1793–1798.

Jimbo, H., *et al*. (1997). Evidence for monoclonal expansion of epithelial cells in ovarian endometrial cysts. *American Journal of Pathology*, **150**, 1173–1178.

Jones, P. A. (1996). DNA methylation errors and cancer. *Cancer Research*, **56**, 2463–2467.

Lakhani, S. R., *et al*. (1995). Atypical ductal hyperplasia of the breast: clonal proliferation with loss of heterozygosity on chromosomes 16q and 17p. *Journal of Clinical Pathology*, **48**, 611–615.

Lunec, J., Challen, C., Wright, C., Mellon, K. and Neal, D. (1992). *c-erB2* amplification and identical p53 mutations in comcomitant transitional carcinomas of renal pelvis and urinary bladder. *Lancet*, **339**, 439–440.

Merrit, A. J., *et al*. (1997). Polyclonal structure of intestinal adenomas in ApcMin/+ mice with concomitant loss of Apc+ from all tumor lineages. *Proceedings of the National Academy of Sciences of the USA*, **94**, 13927–13931.

Murry, C. E., *et al*. (1997). Monoclonality of smooth muscle cells in human atherosclerosis. *American Journal of Pathology*, **151**, 697–706.

Novelli, M. R., *et al*. (1996). Polyclonal origin of colonic adenomas in an XO/XY patient with FAP. *Science*, **272**, 1187–1190.

Nowell, P. C. (1976). The clonal evolution of tumour cell populations. *Science*, **194**, 23–28.

Nowell, P. C. (1986). Mechanisms of tumor progression. *Cancer Research*, **46**, 2203–2207.

Park, H.-S., *et al*. (1995). Crypt fission in the small intestine and colon: a mechansism for the emergence of glucose-6-phosphate locus-mutated crypts after mutagen treatment. *American Journal of Pathology*, **147**, 1416–1427.

Ponder, B. A. J., *et al*. (1985). Derivation of mouse intestinal crypts from single progenitor cells. *Nature*, **313**, 689–691.

Schmidt, G. H. and Mead, R. (1990). On the clonal origin of tumours – lessons from studies of intestinal epithelium. *Bioessays*, **12**, 37–40.

Slaughter, D. P., *et al*. (1953). 'Field cancerisation' in oral stratified squamous epithelium. *Cancer*, **6**, 963–968.

Woodruff, M. F. A., *et al*. (1982). Clonal interaction in tumours. *Nature*, **299**, 822–824.

Zhuang, Z., *et al*. (1996). Barrett's esophagus: metaplastic cells with loss of hetrozygosity at the *APC* gene locus are clonal precursors to invasive adenocarcinoma. *Cancer Research*, **56**, 1961–1964.

FURTHER READING

Garcia, S., *et al*. (1999). Field cancerization, stem cells and clonality: the spread of mutated clones in epithelial sheets. *Journal of Pathology*, **187**, 61–81.

Garcia, S., *et al*. (2001). The clonal origin and clonal evolution of human epithelial tumours. *International Journal of Experimental Pathology*, **81**, 117–143.

Hall, P. A., *et al*. (eds) (1992). *Clinical Aspects of Cell Proliferation* (Springer, Heidelberg).

Kinzler, K. W. and Vogelstein, B. (1997). Gatekeepers and caretakers. *Nature*, **386**, 761–763.

Pretlow, T. P., *et al*. (1994). Aberrant crypt foci and colon tumours in F344 rats have similar increases in proliferative activity. *International Journal of Cancer*, **56**, 599–602.

Quinn, C. and Wright, N. A. (1990). The clinical assessment of proliferation and growth in human tumours; evaluation of methods and applications as diagnostic variables. *Journal of Pathology*, **160**, 93–102.

Shibata, D. (1998). The dynamics of early intestinal tumour proliferation: to be or not to be. In Precancer: Biology, Importance and Possible Prevention. *Cancer Surveys*, **32**, 181–200.

Steele, G. G. (1977). *Growth Kinetics of Tumours* (Clarendon Press, Oxford).

Wright, N. A. and Alison, M. R. (1984). *The Biology of Epithelial Cell Populations*, Vols 1 and 2 (Oxford University Press, Oxford).

Wright, N. A. (2001). Stem cell repertoire in the intestine: the origins of self-renewal, cell lineage and cancer. *International Journal of Experimental Pathology*, **81**, 89–116.

Chapter 19

Identifying Cancer Causes Through Epidemiology

Hans-Olov Adami
Karolinska Institutet, Stockholm, Sweden

Pagona Lagiou
University of Athens Medical School, Goudi, Athens, Greece

Dimitrios Trichopoulos
Harvard School of Public Health, Boston, MA, USA

CONTENTS

- Introduction
- The 'Natural History' of the Epidemiological Identification of a Cancer Cause
- Descriptive Cancer Epidemiology
- Formulation of Aetiological Hypotheses
- Epidemiological Studies Evaluating Cancer Causation
- Chance, Confounding and Bias
- Inference of Cancer Causation in Epidemiology
- Conclusion

INTRODUCTION

Cancer control is an important objective of contemporary medicine. It may be achieved through primary prevention (e.g., reduction of exposure to carcinogenic agents or increase in resistance towards them by immunization or other procedures), down-staging (earlier recognition of clinical symptoms or signs), screening for preclinical diagnosis of asymptomatic cancer or preneoplastic conditions, curative and palliative treatment and rehabilitation schemes. Although these approaches are complementary, there are strong theoretical arguments and overwhelming empirical evidence suggesting that primary prevention represents the most promising strategy for effective cancer control. Implementation of primary prevention requires identification of carcinogenic agents and of the conditions that favour the exposure of susceptible individuals to these agents. An agent is considered to be carcinogenic when a change in the frequency or intensity of exposure to this agent is accompanied by a predictable change in the frequency of occurrence of cancer of particular type(s) at a later time.

The issue of causal inference has generated intense debates among both philosophers and scientists (Rothman, 1988). In a simplified scheme that reflects, to a considerable extent, the ideas of Sir Karl Popper (Buck, 1975), causal inference follows a pattern of interconnected cycles. Each cycle includes an examination of the existing data in the light of the prevailing views and questions, formulation of an aetiological hypothesis that adequately addresses these questions and evaluation of the compatibility of the hypothesis with new sets of data specifically generated or assembled for this purpose. Compatibility cannot be equated to proof, but a hypothesis gains credibility when it repeatedly resists refute.

In 1969, the World Health Organisation (WHO), through its International Agency for Research on Cancer (IARC), initiated a programme to evaluate carcinogenic risks to humans and to produce monographs on individual agents, groups or mixtures (IARC, 1987). For the evaluation, evidence for carcinogenicity in humans and experimental animals, and also other relevant data in experimental systems and humans, are taken into account. However, in the final overall evaluation, an agent is judged to be 'conclusively' carcinogenic in humans when there is sufficient evidence of carcinogenicity in humans derived from relevant and valid epidemiological studies. Evidence for carcinogenicity in experimental animals and *in vitro* experimental systems is given less weight, because of well-known problems related to species specificity. Series of clinical case reports are also considered, since they can be thought of as incomplete epidemiological processes, with 'control' information based implicitly on background clinical experience and information.

THE 'NATURAL HISTORY' OF THE EPIDEMIOLOGICAL IDENTIFICATION OF A CANCER CAUSE

The formulation of aetiological hypotheses is usually based on the examination of existing data. These data may represent the results of studies in experimental animals, e.g. the occurrence of papillary carcinoma in the bladder of mice after exposure to tobacco tar encouraged investigators to examine whether an association between tobacco smoking and bladder cancer also existed in humans. In other instances, the data may refer to 'unusual' or 'interesting' cases reported in the clinical literature, e.g. the hypothesis linking inorganic trivalent arsenic compounds to skin cancer (IARC, 1980a) and phenacetin-containing analgesics to renal pelvic carcinoma (IARC, 1980b) have been based, to a large extent, on clinical observations and pathophysiological considerations. There have also been situations where hypotheses were developed and subsequently tested on the basis of biological and theoretical arguments. For example, it was hypothesized that passive smoking may cause lung cancer, because sidestream smoke is not qualitatively different from mainstream smoke, and there is no threshold in the dose-dependent relation between active smoking and risk of lung cancer (Trichopoulos, 1994).

In most instances, however, aetiological hypotheses are developed on the basis of statistical associations between the cancer under consideration on the one hand, and various personal characteristics of the affected individuals, in addition to the time and place occurrence pattern of their disease, on the other. By revealing who, when and where are affected by a particular cancer, one has already gone a long way towards discovering why the particular cancer has occurred. Such observations, collectively considered under the term 'descriptive epidemiology,' represent either the products of routinely recorded information or the byproducts of analytical epidemiological studies designed to address other specific aetiological hypotheses.

An aetiological relation presupposes the existence of a statistical association, but for diseases such as cancer, which are defined according to criteria at the histological, cytological or subcellular level, a cause does not have to be, and usually is not, either necessary or sufficient. For example, not all hepatitis B virus carriers develop hepatocellular carcinoma, and this cancer can develop without the presence of, or even exposure to, the hepatitis B virus (Stuver, 1998). Furthermore, the existence of a statistical association between a particular agent and a particular form of cancer does not necessarily imply the existence of an underlying causal relation. It is possible, in fact common, that the association reflects coexistence of the particular agent with another agent (the confounding factor) which represents the real cause of the particular cancer. Thus, lung cancer patients may report excessive use of alcoholic beverages, simply because in several cultures tobacco smoking and alcohol drinking tend to be positively correlated (**Figure 1**). Even in the absence of confounding, a statistical association is not an infallible indication of a causal relation. Non-smoking lung cancer patients, for example, may report higher alcohol intakes compared with healthy individuals, because they provide more truthful and accurate histories of habits for which there is real or perceived social disapproval (information bias).

Analytical epidemiological studies are designed to explore whether an association between a particular agent or characteristic and a particular cancer actually exists, and what is its real strength after eliminating, as far as possible, all recognizable effects of confounding and bias. The most commonly used measures of strength of the association between a particular agent or characteristic and a particular cancer is the relative risk; this generic term covers the rate ratio, the risk ratio and the odds ratio. The relative risk indicates how many times higher (or lower) is the frequency of occurrence of the particular cancer among individuals exposed to the agent (or possessing the characteristic), compared with individuals not exposed to the agent (or not possessing the characteristic).

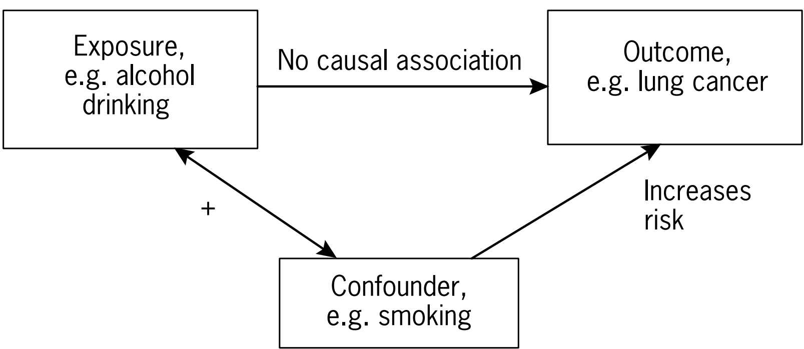


Figure 1 Confounding of the association between consumption of alcoholic beverages and lung cancer by tobacco smoking.

Analytical epidemiological studies have been traditionally designated as cohort or case control. In cohort studies, exposed and nonexposed individuals are followed over time and the frequency of occurrence of the cancer under investigation in the two (or more, if several exposure levels can be ascertained) groups is calculated, allowing the direct estimation of the incidence rate ratio, which is a variant of the relative risk (**Figure 2**). In case-control

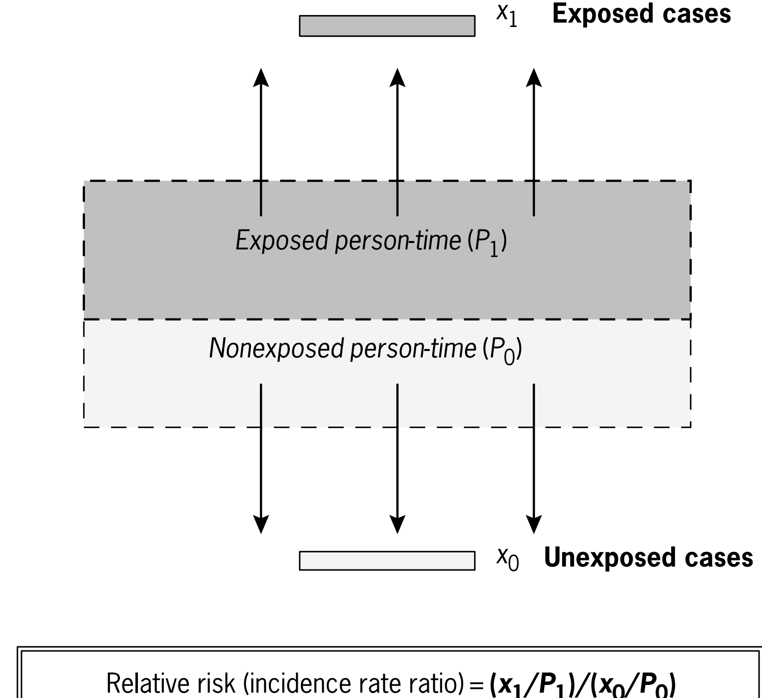


Figure 2 A cohort study. (Adapted from Walker, 1991.)

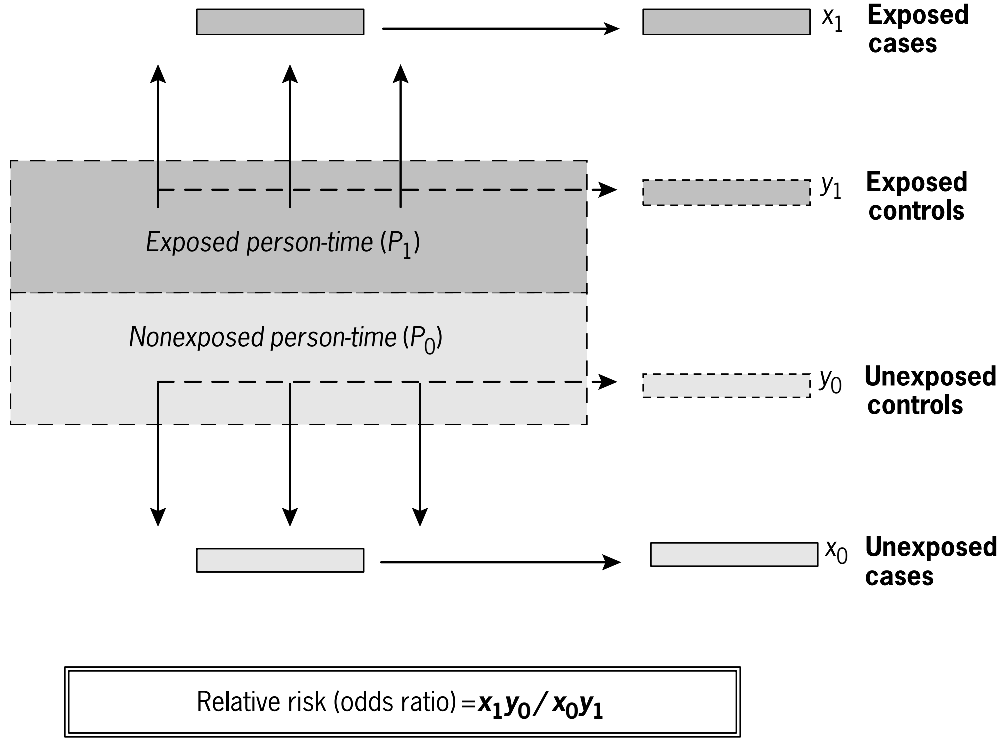


Figure 3 A case-control study. (Adapted from Walker, 1991.)

studies, patients with a recent diagnosis of the cancer under investigation are compared with individuals free of this cancer. These individuals (controls) should be representative of the population which gave rise to the cases with respect to the antecedent frequency of exposure to the agent under study. It can be shown that this design permits the calculation of the odds ratio, which is another variant of the relative risk. The actual frequency of occurrence of the cancer under consideration among exposed and nonexposed individuals usually cannot, however, be calculated in a case-control study (**Figure 3**).

In both cohort and case-control studies, measurable confounding factors and identifiable biases can be controlled for in the analysis, either by stratification or by multivariate modelling (Breslow and Day, 1980, 1987). Contemporary authors have indicated that a sharp distinction between cohort and case-control studies is artificial and unjustifiable (Miettinen, 1985; Rothman and Greenland, 1998). However, most epidemiological studies are still reported under these headings and their distinction may be useful for readers who are not methodologically orientated.

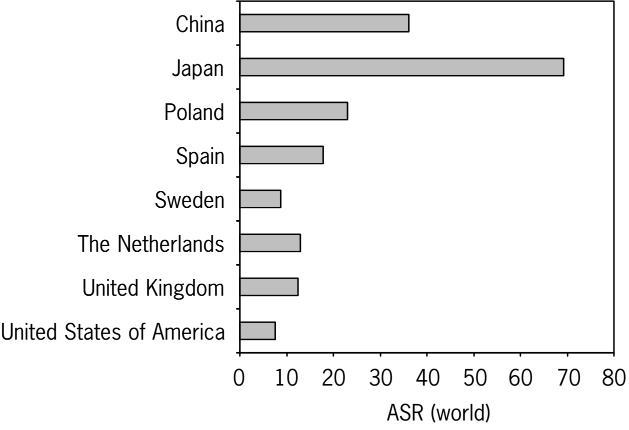


Figure 4 Age-standardized incidence rates of stomach cancer among males. (Based on data from Ferlay *et al.*, GLOBOCAN 2000: Cancer Incidence, Mortality and Prevalence Worldwide, Version 1.0. IARC CancerBase No. 5. IARC Press, Lyon, 2001.)

DESCRIPTIVE CANCER EPIDEMIOLOGY

Descriptive cancer epidemiology examines the distribution of individual cancer types by a number of characteristics. These characteristics may refer to individuals themselves (personal characteristics, including age, gender, occupation, education, marital status, etc.), to the place of disease initiation or occurrence (place characteristics, including the country and region of residence, the urban or rural nature of the area, the altitude, the latitude, etc.) or the time pattern of occurrence (time characteristcs, including long-term trends, time elapsing between exposure to a certain agent and the appearance of a particular cancer, etc.). Certain characteristics are inherently multidimensional, e.g. migration refers to individuals, in addition to the country of origin and the time of migration.

In most instances, descriptive epidemiological associations can only generate aetiological hypotheses, because observed associations are frequently confounded by many factors and are also subject to a number of potential biases. For example, the high incidence of stomach cancer in Japan (**Figure 4**) can be attributed to a number of intercorrelated genetic, nutritional and socio-cultural characteristics of the Japanese people. However, on rare occasions, descriptive epidemiological data can be revealing by themselves, particularly when an otherwise rare tumour occurs frequently among persons with a certain defining characteristic, usually their occupation or medical treatment for an earlier condition. For example, the high incidence of cancer of the nasal cavity among workers in nickel refineries strongly suggests that the agents or the processes involved in nickel refining are carcinogenic.

Personal Characteristics

Among the many personal characteristics, those most frequently explored in descriptive epidemiology are age, gender, religion, marital status, occupation and socioeconomic class. These characteristics are rarely themselves aetiologically relevant, but they are commonly correlated with exposure to carcinogenic factors, and for this reason they are frequently considered risk indicators or risk factors. Furthermore, several of these characteristics are readily available because they are routinely recorded for administrative reasons.

Because the incidence of most cancers increases rapidly with age, exceptions to this pattern may have aetiological significance. Acute leukaemia, for example, shows both an early and a late peak, which led investigators to consider intrauterine carcinogenesis. The subsequent discovery that intrauterine exposure to ionizing radiation may increase the risk of childhood leukaemia provided empirical support to this hypothesis. Similarly, the bimodal incidence pattern of Hodgkin disease in many countries (**Figure 5**) led to the prevailing hypothesis that the early peak is due to an infectious agent. Finally, the slowing of the rate of increase of the incidence of breast cancer after menopause has underlined the importance of ovarian oestrogens in the aetiology of this disease.

Many cancers are more common among men than among women. This has been attributed, in many instances, to the higher past exposure of men to tobacco smoking and alcohol drinking. However, some intriguing gender differences remain unexplained. The incidence of hepatitis B-positive hepatocellular carcinoma is three times higher among men, even though the prevalence of chronic hepatitis B virus infection is only two times higher among them.

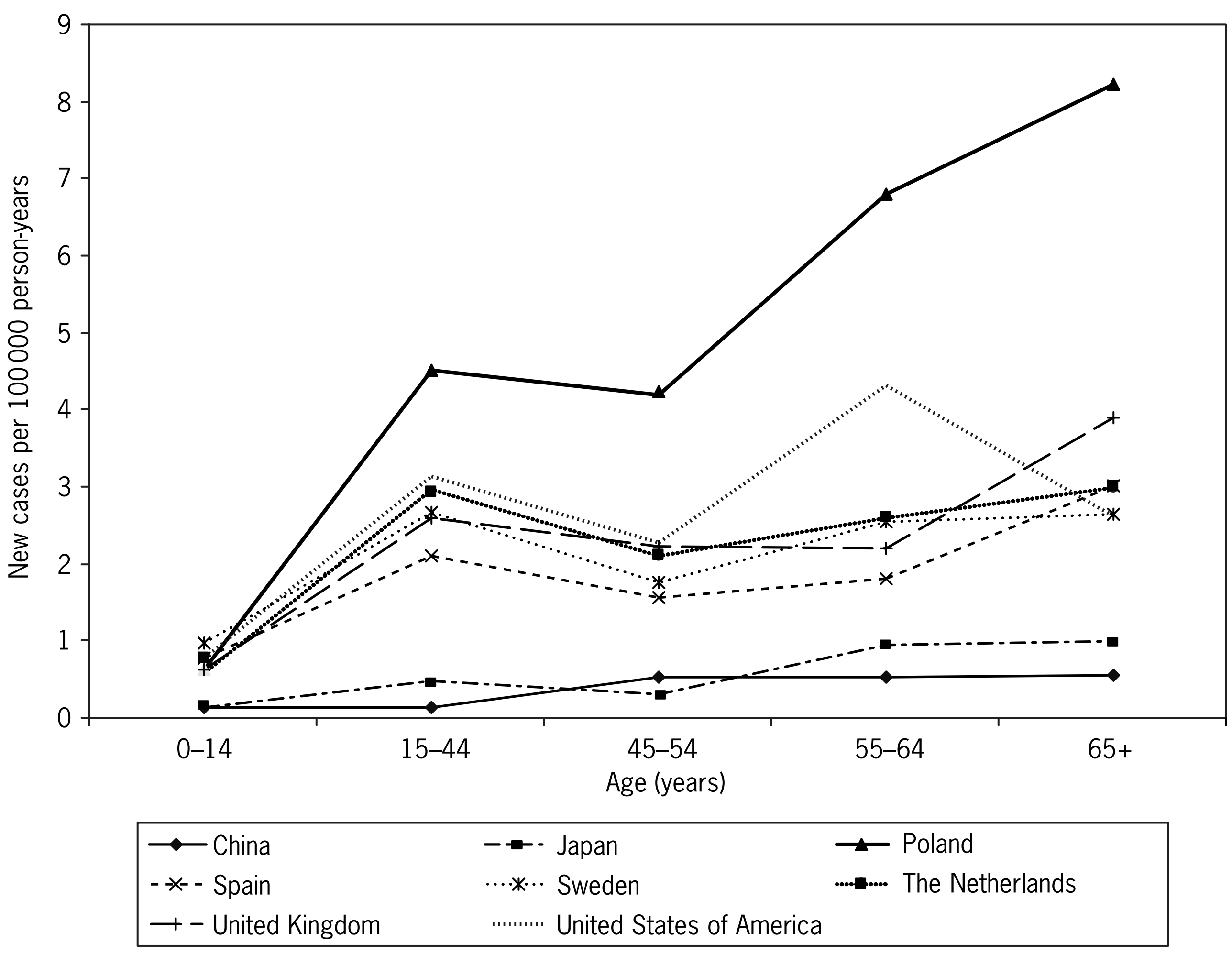


Figure 5 Age-specific incidence rates of Hodgkin disease among males. (Based on data from Ferlay *et al.*, GLOBOCAN 2000: Cancer Incidence, Mortality and Prevalence Worldwide, Version 1.0. IARC Cancer Base No. 5. IARC Press, Lyon, 2001.)

In contrast, the incidence of gallbladder cancer (and also cholelithiasis) is considerably more common among women.

Religion, ethnic origin and race are frequently interrelated. The realization that environmental factors, including lifestyles, are more important than hereditary factors in explaining cancer variation among population groups (Lichtenstein *et al.*, 2000) has suggested that those groupings should be viewed in the socio-cultural, rather than genetic, context. Several hypotheses concerning cancer causation have been generated by the unusual lifestyles and cancer occurrence patterns among Mormons and Seventh Day Adventists in the United States. Also, the low incidence of cervical cancer among nuns has led to the hypothesis that sexual activity is an important determinant of this cancer.

Marital status has been an important parameter for the formulation of aetiological hypotheses concerning cancer occurrence in the female reproductive organs. Marital status associations, most likely reflecting underlying causal links of nulliparity with breast cancer, sexual activity with cervical cancer and low parity with ovarian cancer, are stronger in socio-cultural settings in which marital status is strongly correlated with sexual activity and reproduction.

Several forms of cancer, notably cancers of the stomach and uterine cervix, are more frequent in the lower socio-economic classes, and these observations have led to the development of aetiological hypotheses implicating, respectively, particular dietary patterns (salty foods, inadequate intake of vegetables and fruits), or inadequate hygienic conditions and infrequent utilization of barrier contraceptive methods. Furthermore, occupational mortality statistics have been extremely valuable in pointing out occupations that increase the risk of particular types of cancer.

Place and Time Characteristics

Cancer mortality statistics by site, age and gender are published, on an annual basis, by the WHO for most countries of the world. Furthermore, the IARC publishes, every 5 years, Cancer Incidence in the Five Continents, containing data – by site, age and gender – generated by a worldwide network of cancer registries (Parkin *et al*., 1992). Correlations between the incidence of, or mortality from, particular cancers in various countries on the one hand, and per capita consumption in these countries of various food groups or nutrients on the other, generated hypotheses about the nutritional aetiology of several cancers (Armstrong and Doll, 1975).

National cancer rates have also been used extensively in studies of migrant populations, to distinguish the relative importance of heredity and environment (including lifestyle) in the generation of the large international variability of cancer occurrence at most sites. In migrant populations, e.g. Japanese migrating to the United States, the incidence of large bowel cancer approaches that of the host country within a few decades, whereas for stomach cancer a whole lifetime is needed (Haenszel and Kurihara, 1968) and for breast cancer several generations may be required for incidence assimilation (Ziegler *et al*., 1993). Lastly, large-scale geographic patterns have been used to generate hypotheses linking malaria endemicity with the occurrence of Burkitt lymphoma (Burkitt, 1970) and sunlight exposure with the incidence of skin cancer (Scotto *et al*., 1982).

Long-term temporal trends of lung and stomach cancer in developed countries (**Figure 6**) have led to hypotheses implicating tobacco smoking and food preservation modalities in their aetiology. Specifically, the rapid rise of lung cancer mortality among men during the first half of the twentieth century, and the much slower rise of the corresponding mortality among women, indicated that a novel strong carcinogen became widespread among men but much less so among women. Few agents, besides tobacco, would meet these criteria, and many analytical epidemiological studies have demonstrated that smoking was indeed the responsible factor. Whilst no comparable clear explanation exists for the remarkable decline in stomach cancer mortality, many investigators speculate that the large-scale

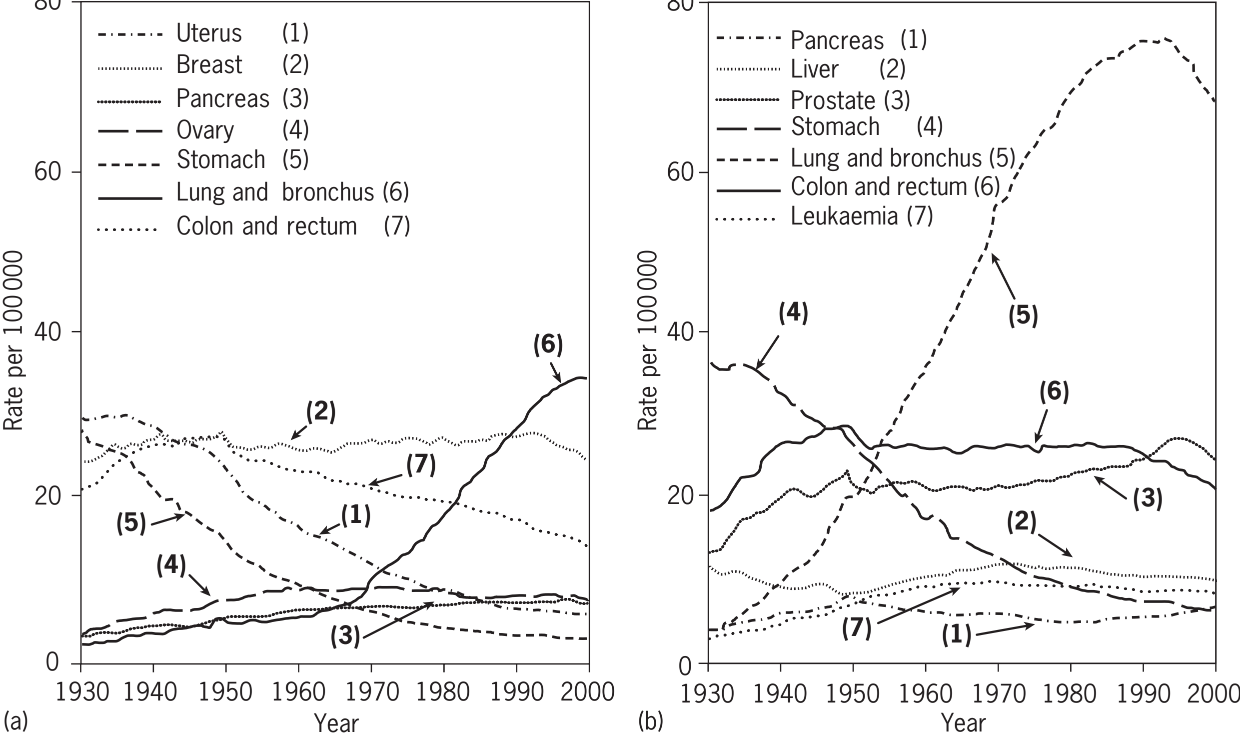


Figure 6 Age-adjusted (to the 1970 US standard population) cancer death rates by site among (a) Females and (b) males in the USA from 1930 to 1996, expressed per 100 000 person-years. (a) Uterus cancer death rates are for uterine cervix and uterine corpus combined. Note: Owing to changes in ICD coding, numerator information has changed over time. Rates for cancers of the uterus, ovary, lung and bronchus, and colon and rectum are affected by these coding changes. (b) Note: Owing to changes in ICD coding, numerator information has changed over time. Rates for cancers of the liver, lung and bronchus, and colon and rectum are affected by these coding changes. (From Greenlee *et al*., 2000, *Cancer Journal for Clinicians*, **50**, 7–33.)

introduction of refrigeration for food preservation may have been responsible.

Clustering in both time and place, a characteristic of diseases caused by infectious agents, has been investigated using an ingenious procedure developed by Knox (1963) for cases of acute childhood leukaemia. Notwithstanding inconclusive results, the hypothesis postulating involvement of infectious agent(s) is still widely entertained.

FORMULATION OF AETIOLOGICAL HYPOTHESES

The reasoning leading to the generation of aetiological hypotheses follows, explicitly or implicitly, a set of rules which focus on difference, agreement, concomitant variation and analogy (MacMahon and Trichopoulos, 1996).

According to the rule of difference, when both a particular set of factors and the incidence of a particular cancer differ between two populations, one or more of these factors will likely contribute to the occurrence of the disease. When the difference in incidence is large, in relative terms, and only one or two factors form the differentiating set, then the hypothesis is more likely to be correct, and vice versa. Thus, the relatively high incidence of Kaposi sarcoma among homosexual men in the United States in the early 1980s indicated that sexual preference was involved in the occurrence of the disease, since few other, if any, characteristics distinguish homosexual from heterosexual men. Also, the lower incidence of cervical cancer among nuns points to sexual activity as an important causal factor, since nuns differ from other women with respect to only a few characteristics and practices. By contrast, the usually higher incidence of lung cancer in large cities cannot be attributed confidently to air pollution, since city dwellers differ from rural inhabitants with respect to a number of exposures, including those related to occupation, diet, smoking, environmental tobacco smoke and radon.

The rule of agreement focuses on common factors identifiable in different settings characterized by a high incidence of particular cancers. Thus, exposure to ionizing radiation was the common factor characterizing patients with ankylosing spondylitis therapeutically irradiated and children diagnostically X-rayed *in utero*, when both groups were found to have increased incidence of leukaemia. Also, multiplicity of sexual partners is the likely common factor characterizing women with squamous carcinoma of the uterine cervix and male homosexuals with anal cancer.

Concomitant variation represents a quantitative expression of the first two rules, particularly the first. The concomitant variation of per capita fat consumption and the incidence of, or mortality from, colorectal cancer, in international correlations (Armstrong and Doll, 1975) was important for the hypothesis linking fat consumption to colorectal cancer – a hypothesis supported by the results of many, although not all, analytical epidemiological studies. Correlations based on countries, regions or other population groups (as distinct from correlations based on individuals) are termed population, group or ecological correlations. From the epidemiological point of view, they may be considered half-way between descriptive and analytical studies, and their aetiological importance is, as a rule, limited.

Inference by analogy represents the fourth and arguably the softest rule. The paradigm of poliomyelitis, in which early infection is harmless, whereas late infection causes the paralytic form of the disease, has been invoked by several authors in attempts to explore the aetiology of childhood leukaemia.

The credibility of an aetiological hypothesis developed on the basis of descriptive epidemiological data increases substantially if the suspected agent causes cancer in experimental animals, preferably in more than one species and in the same organ as in humans. In the absence of a satisfactory experimental model, demonstration of *in vitro* mutagenicity provides supportive evidence, as does the identification of a likely pathogenetic mechanism. Nevertheless, the ultimate criterion of the validity of an aetiological hypothesis – whether derived from descriptive epidemiological findings or formulated on the basis of experimental data, clinical information or theoretical considerations, or any combination of the above – is in most cases epidemiological. Specifically, an aetiological hypothesis is refuted when it generates predictions that appear incompatible with the collective evidence of analytical epidemiological studies. In contrast, it remains provisionally valid when it generates predictions compatible with the accumulated empirical evidence from human studies.

EPIDEMIOLOGICAL STUDIES EVALUATING CANCER CAUSATION

In theory, the best empirical evidence regarding causation should come from randomized trials in humans. Studies in humans do not face the pitfalls inherent in the reasoning by analogy from animal to human data, whereas randomization allows control for distorting influences by both known and unsuspected confounding factors; in addition, double-blind designs minimize the potential for several types of bias. Indeed, therapeutic clinical trials are also aetiological studies exploring the causation of a better clinical outcome by a certain process, the study treatment. However, experimental studies of cancer causation cannot be undertaken in humans for ethical reasons, except when there is evidence that a particular factor may actually protect from cancer (in which instance the absence of the factor can be thought of as the carcinogenic agent). Even under these conditions, experimental studies in humans

are still impractical. Among cancer patients, the outcome under investigation (death, metastasis or recurrence) is a relatively frequent effect, and the corresponding study size can be manageable. Among healthy individuals, the frequency of occurrence of any particular cancer is low and the corresponding study size must be very large, making the follow-up and compliance problems exceedingly difficult. Nevertheless, a few such studies have been undertaken, either by investing large resources (e.g. the IARC study targetting the prevention of hepatocellular carcinoma in the Gambia through active immunization against hepatitis B virus) or by reducing the required sample size by focusing on preneoplastic conditions that identify high risk individuals.

Epidemiological studies specifically designed to address a particular aetiological hypothesis are usually called analytical. The objective of analytical studies is to document causation between exposures and a certain disease. In analytical investigations, measurements and categorical assignments apply to individuals, whereas this is not necessary in descriptive epidemiological studies. Thus, in order to examine in an analytical study whether chronic carriers of hepatitis B virus (HBV) are more likely to develop liver cancer than noncarriers, it is necessary to classify the individuals under study according to their HBV carrier state and the development or not of liver cancer during a specified time period (Hennekens and Buring, 1987; MacMahon and Trichopoulos, 1996; Rothman and Greenland, 1998).

Ecological studies in epidemiology, as opposed to individual-based studies, occupy an intermediate position between descriptive and analytical investigations, in that they share many characteristics with descriptive studies, but may also serve aetiological objectives. In ecological studies, the exposure and the disease under investigation are ascertained not for individuals but for groups or even whole populations (Morgenstern, 1982). Thus the prevalence of HBV in several populations could be correlated with the incidence of liver cancer in these populations, even though no information could be obtained as to whether any particular individual in these populations was or was not an HBV carrier and has or has not developed liver cancer.

When an exposure is fairly common (e.g. smoking, sunlight, poverty, even prevalence of HBV carriers), ecological studies should be able to reveal the effects of these exposures. Thus, following the increase in tobacco consumption, the incidence of lung cancer increased sharply over time; skin melanoma is more common in geographic latitudes with more sunshine exposure; stomach cancer is generally more common in low-income social strata; and the incidence of primary liver cancer is higher in populations with higher prevalence of HBV (Tomatis, 1990). As a corollary, the inability of ecological studies to demonstrate an association between a widespread exposure that has rapidly increased over time (e.g. extremely low-frequency magnetic fields) and the incidence of a disease allegedly caused by these fields (e.g. childhood leukaemia) challenges the causal nature of the positive association reported from some analytical investigations.

In analytical epidemiological studies, there are several ways through which an association, or lack thereof, is assessed, but the most common measure is the relative risk. A value equal to 1 implies that the exposure under study does not affect the incidence of the disease under consideration. In contrast, values <1 and >1 indicate, respectively, an inverse or a positive association (MacMahon and Trichopoulos, 1996; Rothman and Greenland, 1998). When the relative risk is >2 and the associated unbiased, unconfounded, precise and causal, an exposed-case patient is more likely than not to have developed the disease because of the exposure. When the relative risk is >1 but <2 the exposed patient is more likely than not to have developed the disease for reasons other than the exposure. This is because a relative risk of, say, 1.5 has a baseline component equal to 1 that characterizes the unexposed, and a component equal to 0.5 that applies only to the exposed (MacMahon and Trichopoulos, 1996). For instance, if the risk of a nonsmoking 55-year-old man developing lung cancer in the next 10 years is 1%, and that of a same age and gender smoker is 5% (relative risk 5), only 4% in the smoker's risk (i.e. 4/5 of the total 5%) can be attributed to his smoking.

Cohort Studies

In cohort studies (**Figure 2**), individuals are classified according to their exposure and are followed over time for ascertainment of the frequency of disease occurrence in the various exposure-defined categories. In each category the frequency of occurrence is calculated either as risk (proportion of those who developed the disease under study among all individuals in this category) or as incidence rate (number of those who developed the disease divided by the sum of the time each of the individuals in this category has been under observation). Defining characteristics of cohort studies are that they are exposure based and are patently or conceptually longitudinal.

In cohort studies, the groups to be studied are selected on the basis of exposure. In some studies the groups are chosen on the basis of a particular exposure (special exposure cohorts), whereas in others, groups offering special resources for follow-up are initially chosen and the individuals are subsequently allocated according to exposure status (general population cohorts). The first approach may be necessary when rare exposures need to be studied, while the second approach is appropriate when the exposure under consideration is fairly common in general populations (e.g. smoking or major dietary components).

The cohort study groups are observed over a period of time to determine the frequency of occurrence of disease among them. The distinction between retrospective and

prospective cohort studies depends on whether or not the cases of disease have occurred in the cohort at the time the study began. In a retrospective cohort study, all the relevant exposures and health outcomes have already occurred when the study is initiated. In a prospective cohort study, the relevant causes may or may not have occurred at the time the study began, but the cases of disease have not yet occurred and, following selection of the study cohort, the investigator must wait for the disease to appear among cohort members.

From a methodological point of view, there are two types of cohort studies: closed or fixed cohorts and open or dynamic cohorts. Closed cohorts are frequent in occupational epidemiology and study of outbreaks, whereas open cohorts dominate cancer epidemiology and are the basis for most case-control studies. The key distinction between open and closed cohorts is how membership in the cohort is determined. In a closed cohort, membership is determined by a membership-defining event which occurs at a point in time. Studies based on follow-up of closed cohorts may be analysed using either cumulative incidence (risk) measures or by counting person-time and calculating incidence rate measures. Analyses based on cumulative incidence measures are useful only under certain conditions (that is, no loss to follow-up; no competing risks; exposure status unchanged throughout follow-up; study subjects followed for the same period of time). Whether or not these assumptions are met, it is always valid to conduct analyses based on person-time, using incidence rate measures in the setting of a closed cohort.

Open cohorts are composed of individuals who contribute person-time to the cohort, while meeting criteria for a membership-defining state. Once individuals can no longer be characterized by this state, they cease to contribute person-time to the corresponding cohort. In studies based on open cohorts it is not possible to measure cumulative incidence (risk) directly. Instead, analyses are based on person-time using incidence rate measures. Thus, if among 5000 non-smoking men followed for an average period of 10 years ($P_0 = 50\,000$ person-years) $x_0 = 25$ were diagnosed with lung cancer, whereas among 10 000 smoking men followed for an average period of 8 years ($P_1 = 80\,000$ person-years) $x_1 = 600$ were diagnosed with lung cancer, the incidence rate among nonexposed would be 50 per 10^5 person-years, that among exposed 750 per 10^5 person-years and the incidence rate ratio would be $(600/80\,000)/(50/50\,000) = 15$.

Case-control Studies

In case-control studies (**Figure 3**), patients recently diagnosed with the disease under consideration form the case series, and their exposure to the factor under investigation is ascertained through questionnaires, interviews, examination of records, undertaking of laboratory tests in biological samples and other means. The pattern of exposure to the study factor in the population that generated the case series is then evaluated through a properly selected control series. If only two categories of exposure are relevant (exposed and unexposed), the relative risk can be estimated by multiplying exposed cases with unexposed controls and dividing this product by the product of unexposed cases and exposed controls (the so-called odds ratio, that adequately estimates rate ratio and risk ratio). Thus, if among 200 male patients diagnosed with lung cancer (cases), $x_1 = 150$ were smokers and $x_0 = 50$ nonsmokers, whereas among 300 men of similar age as the cases but without lung cancer (controls), $y_1 = 50$ were smokers and $y_0 = 250$ were nonsmokers, the odds ratio x_1y_0/x_0y_1, which is a good approximation to the relative risk, is $(150 \times 250)/(50 \times 50) = 15$.

Two features of case-control studies make them susceptible to bias: the ascertainment of exposure after the occurrence of disease (information bias), and suboptimal processes for control selection (selection bias) (Hennekens and Buring, 1987; MacMahon and Trichopoulos, 1996; Rothman and Greenland, 1998).

Some case-control designs, in particular the population-based ones, are methodologically superior to other case-control variants. A case-control study is called population-based when controls are chosen from the clearly defined population from which all cases have arisen – in other words, had one of the controls developed the disease under study it would have definitely been included among the cases. Cohort studies, however, generally produce more credible results than those emerging from case-control investigations, because, if properly designed and implemented, they are, as a rule, free from information and selection bias.

Frequently, case-control studies are matched, in the sense that controls are chosen so as to match particular cases with respect to gender, age, race or any other factor that is probably related to the disease under study but not intended to be analysed in the particular study. Matching is not strictly necessary, nor does it confirm the validity of results, but it improves statistical efficiency, i.e. statistical power, or the ability to substantiate a true association (Rothman and Greenland, 1998). If matching has been used in the enrollment of cases and controls, the statistical analysis should accommodate the matching process, through either a matched analysis (e.g. conditional modelling) or unmatched analysis with explicit control of the matching factors (proper application of unconditional modelling).

The advantages and disadvantages of cohort and case-control studies in exploring the causes of cancer in humans are described in several textbooks on epidemiology. Briefly, cohort (or prospective or follow-up) studies trace in a more natural way the time sequence of causal phenomena. However, the case-control approach is, in essence, an extension of a practice familiar to physicians, that of case history taking. Confounding is just as common

in cohort as in case-control studies, but information bias concerning exposure to suspected agent(s) is more common in case-control investigations, since the existence of a serious disease, such as cancer, can affect recollection and reporting, as well as several directly ascertainable biochemical and immunological variables. By contrast, case-control studies are usually more powerful, in the statistical sense, since the frequency of exposure to many suspected agents is more common than the frequency of occurrence of most cancers during a not-too-long period of follow-up.

CHANCE, CONFOUNDING AND BIAS

Three issues need to be resolved before an epidemiological association could be considered true and therefore deserve interpretation in causal terms: chance, confounding and bias.

Probabilistic processes always have a built-in uncertainty, but we can reduce the chance-related uncertainty by using progressively larger numbers in a study, and we can assess its possible influence by utilizing statistical procedures that generate what has become known as the *P* value. This value indicates how likely one would have been to observe an association as extreme as, or more extreme than, that found between a particular exposure and a certain disease, if there were in fact no association. The true meaning of the *P* value, however, is poorly understood and the concept itself is widely misused.

P values are traditionally expressed as numerical fractions of unity. For example, a *P* value of 0.1 for a particular positive association indicates that there is a 10% chance that such an association or more extreme (or a symmetrically opposite one, i.e. an inverse association) would appear by chance, even if there were in reality no association at all. In essence, the *P* value is interpretable as such when only one comparison or one test is performed. When multiple comparisons or multiple tests are carried out, the set of the respective *P* values lose their collective interpretability. Procedures for adjusting *P* values according to the number of comparisons undertaken or tests performed have been proposed, but they are not universally accepted (Rothman, 1990) and are rarely utilized. However, other things being equal, the real significance of a certain *P* value is weaker when the number of tests performed is larger. *P* values can be more confidently interpreted as suggestive of real phenomena or genuine associations when there is independent evidence in support of the process that generated the *P* value.

A *P* value of 0.05 or smaller is traditionally treated in medical research as evidence that an observed association may not have arisen by chance. However, small *P* values, including values considerably smaller than 0.05, do not guarantee that an association (or difference) is genuine, let alone causal. Even when the *P* value is very small and was generated from a randomized trial, it could still be dismissed when the relevant result makes no sense (Miettinen, 1985). As a corollary, our daily lives are full of highly unlikely events and coincidences and, at the extremes, thousands of people have become wealthy from lotteries and many more have died in strange accidents, even though the probabilities of the respective events are extremely small. The lesson is simple: highly unlikely events that would have been associated with extremely small *P* values do happen by chance. In addition to small *P* values, science requires judgement relying on sound substantive knowledge in order to discard chance.

The *P* value itself does not convey information about the strength of the respective association: a weak association may be statistically highly significant (very small *P*) when the study is large, and a strong association may be statistically nonsignificant (large *P*) when the study is small (Rothman and Geenland, 1998). In order to integrate information about the strength of an association (as reflected in the relative risk effect measure) and its statistical significance, the concept of confidence interval has been developed. The confidence interval describes the range of values that the true relative risk is likely to take with, for example, 95% confidence (95% confidence interval) on the basis of the data through which the association was evaluated. The confidence interval is intimately linked to the *P* value, so that for a statistically significant ($P < 0.05$) association, the 95% confidence interval of the measure of this association (e.g. the relative risk) does not include ('rejects') 95 times out of 100 the value of this measure which would have indicated no association (the no association value, also called the 'null value,' is 1 for relative risk).

Random variation *per se* in epidemiological studies is not an insurmountable problem. Larger studies and eventually combined analyses, through systematic statistical evaluations of results of several independent investigations, can effectively address genuine chance-related concerns. Such combined analyses have been termed meta-analyses and pooled analyses. There is no real distinction between the two terms, although meta-analysis has been more frequently used when published data are combined, whereas in pooled analysis primary data from different studies may be made available to the investigator who undertakes the task to combine them. Meta-analyses and pooled analyses have been widely and effectively used for randomized clinical trials and intervention studies, in which confounding and bias are of limited concern (Sacks *et al*., 1987). For observational epidemiological studies, however, the role of meta-analysis is not universally accepted (Shapiro, 1994; Feinstein, 1995), because no statistical summarization can effectively address problems generated by residual confounding, unidentified bias and the different ways investigators choose to present their results.

Confounding is the phenomenon generated when another factor that causes the disease under study, or is

otherwise strongly related with it, is also independently related to the exposure under investigation (**Figure 1**). Thus, if one wishes to examine whether hepatitis C virus (HCV) causes liver cancer, hepatitis B virus (HBV) would be a likely confounder because HBV causes liver cancer and carriers of HBV are more likely to also be carriers of HCV (because these two viruses are largely transmitted by the same route). Hence, if the confounding influence of hepatitis B is not accounted for in the design (by limiting the study to hepatitits B-negative subjects) or in analyses of the data, then the strength of the association between hepatitis C and liver cancer would be overestimated.

A more trivial example is the strong association between carrying matches (or a lighter) on the one hand and developing lung cancer on the other. Neither matches nor lighters cause lung cancer and their association to the disease is due entirely to confounding by cigarette smoking. The confounding factor, cigarette smoking, is the true cause of lung cancer and the dependence of cigarette lighting on matches or lighters generates the confounded entirely spurious association of the latter two factors with the disease.

There are several ways to deal with confounding, some simple, some more complicated. They all assume that two conditions are satisfied: (1) that all the confounders have been identified or at least suspected and (2) that the identified or suspected confounders can be adequately measured (Greenland, 1980). When the study is fairly large, it is always possible to evaluate in the analysis all suspected confounders, but the ability to conceptualize and measure all of them accurately frequently goes beyond the control of the investigator. The result is what has been termed residual confounding (Greenland, 1980; MacMahon and Trichopoulos, 1996; Rothman and Greenland, 1998).

Compounding the problems of epidemiological studies is that the data are never of optimal quality. Data collection relies on the memory of and accurate reporting by study participants, laboratory procedures or existing records. None of these sources is perfect, and frequently the errors are unequal between the compared groups and generate biased results. A reasonable concern is that cases, or their relatives, are inclined to link the disease under study to particular exposures for conscious or subconscious reasons. Cases may also try harder than controls or their relatives do to recall or identify past exposures (information bias).

A well thought-out protocol, standardized procedures and built-in quality control measures can reduce bias and allow quantification of its potential impact. However, complete assurance that bias has been eliminated can never be achieved. In addition, the reliance of case-control studies on a control series that simultaneously has to meet criteria of compliance, comparability to the case series, statistical efficiency and general practicality makes them susceptible to selection bias of unpredictable direction and magnitude (Wacholder *et al*., 1992a–c). Hospital controls, neighbourhood controls and controls enrolled through searches of telephone lists have their own problems, and these have been extensively discussed (MacMahon and Trichopoulos, 1996).

When results of an observational epidemiological study designed to address a specific hypothesis are striking, the study is large and there is no evidence of overt confounding or major biases, it is legitimate to attempt aetiological inferences. The rationale is that powerful confounding presupposes strong risk factors that are unlikely to be missed, and that major biases can be traced to gross and easily identifiable protocol violations. Interpretation becomes particularly problematic when a weak association turns out to be statistically significant perhaps in a large but imperfect data set. Although that association could reflect a weak but genuine causal association, it might more likely be the result of residual confounding, subtle unidentifiable bias or chance following a multiple testing process.

Repeated demonstration of an association of similar direction and magnitude in several studies, undertaken by different investigators in different population groups, increases confidence in a genuine causal basis but cannot conclusively establish this. Nor do meta-analyses prove causality; these techniques essentially address the issue of chance and provide no guarantee that a particular bias, unrecognized confounding or selective reporting have not operated in the constituent studies. It is at this stage that biology confronts epidemiology and the ability to reconcile the two perspectives should be the guiding principle in interpreting epidemiological results.

INFERENCE OF CANCER CAUSATION IN EPIDEMIOLOGY

Criteria for inferring causation from epidemiological investigations have been proposed, over the years, by several authors (MacMahon *et al*., 1960; US Department of Health, Education and Welfare, 1964; Hill, 1965; IARC, 1987; Evans, 1993). The philosophy of causation in epidemiology and medicine has also been examined in various essays (Rothman, 1988). In spite of differences in emphasis, a similar set of principles have been invoked by most authors. Sir Austin Bradford Hill (Hill, 1965) advocated the nine widely used criteria listed in **Table 1** to distinguish causal from noncausal associations. These criteria, although sensible and useful, do not separately address the inherently different issues that are posed by the results of a single study, the results of several studies and the likelihood of causation in a certain individual (Cole, 1997). In reality, the perceived likelihood of an association between a particular exposure and a certain disease being causal moves forward or backward in a continuous

Table 1 The Hill criteria for inferring causation (Hill, 1965)

Strength	A strong association is more likely to be causal. The measure of strength of an association is relative risk and not statistical significance
Consistency	An association is more likely to be causal when it is observed repeatedly and in different populations
Specificity	When an exposure is associated with a specific outcome only (e.g. a cancer site or even better a particular histological type of this cancer), then it is more likely to be causal. There are exceptions, however (e.g. smoking causing bronchitis, lung cancer, pancreatic cancer)
Temporality	A cause should not only precede the outcome (disease), but also be compatible with its latency period (in noninfectious diseases) or its incubation period (infectious disease)
Biological gradient	This criterion refers to the presence of a dose–response relationship. If the frequency of the outcome increases when an exposure is more intense or lasts longer, then it is more likely that the association is causal
Plausibility	An association is more likely to be causal when it is biologically plausible
Coherence	This criterion implies that a cause and effect interpretation of an association should not conflict with what is known of the natural history and biology of the disease
Experimental evidence	If experimental evidence exists, then the association is more likely to be causal. Such evidence, however, is seldom available in human populations
Analogy	The existence of an analogy (e.g. if a drug causes birth defects, then another drug could also have the same effect) could enhance the credibility that an association is causal

spectrum as research results accumulate. The evidence for causality is declared as sufficient when a particular threshold has been reached (MacMahon and Trichopoulos, 1996; Cole, 1997).

Criteria for causality can be invoked, explicitly or implicitly, in evaluating the results of a single epidemiological study, although, in this instance, a firm conclusion is all but impossible (single study level, or level I, according to the classification introduced by Cole and used here (Cole, 1997)). Criteria for causality are more frequently used for the assessment of evidence accumulated from several epidemiological studies and other biomedical investigations. At this stage, the intellectual process is inductive, moving away from the specifics towards generalization (several studies level, or level II).

Regulatory agencies and policy makers may recommend standards, set limits or authorize action even when the scientific evidence is weak, particularly at levels surrounding the proposed standards or limits. These procedures serve public health objectives by introducing a wide safety margin, but should never be confused with the establishment of causation based on scientific considerations alone. Finally, when causation has been established at level II, then, and only then, can the cause of the disease in a particular individual be considered (specific person level, or level III). At this level, the intellectual process is deductive and deterministic, moving from the general concept of disease causation to the examination of what has caused disease in a particular individual.

Whereas causality can be conclusively established between a particular exposure as an entity and a particular disease as an entity, it is not possible to conclusively establish such a link between an individual exposure and a particular disease of a certain individual, e.g. smoking and development of lung cancer. It is possible, however, to infer deductively that the specific individual's illness more likely than not was caused by the specified exposure. For this conclusion all the following criteria must be met (Cole, 1997):

- The exposure under consideration, as an entity, must be an established cause of the disease under consideration, as an entity (level II).
- The relevant exposure of the particular individual must have properties comparable (in terms of intensity, duration, associated latency, etc.) to those of the exposure that, as an entity, has been shown to cause the disease under consideration, again as an entity.
- The disease of the specified person must be identical with or within the symptomatological spectrum of the disease which, as an entity, has been aetiologically linked to the exposure.
- The patient must not have been exposed to another established or likely cause of this disease. If the patient has been exposed to both the factor under consideration (e.g. smoking) and to another causal factor (e.g. asbestos), individual attribution becomes a function of several relative risks, all versus the completely unexposed: relative risk of those who only had the exposure under consideration, relative risk of those who had only been exposed to the other causal factor(s) and relative risk of those who have had a combination of these exposures.
- The scientifically defensible, properly adjusted, relative risk of the disease among individuals who have had the exposure under consideration at the given magnitude should be >2. Only then (as previously discussed) is the diseased individual more likely than not to have developed the disease because of the specified exposure (US Department of Health, Education and Welfare, 1964; MacMahon and Trichopoulos, 1996). If, for

example, the relative risk is 2.5, the component 1 of the background risk is less than the component 1.5 of the excess risk and the individual is more likely than not to have developed the disease because of the exposure.

CONCLUSION

Modern epidemiology has become a rich and powerful toolbox for the study of biological phenomena in humans. Because manipulation of exposures, many of which may be harmful, is usually neither feasible nor ethical, epidemiologists have to base their inferences on experiments that human subjects undertake intentionally, naturally or even unconsciously. The study of risk of lung cancer among smokers compared with nonsmokers is one classical example of a 'natural' experiment. Because human life is characterized by a myriad of complex, often interrelated behaviours and exposures – ranging from genetic traits and features of the intrauterine environment to growth rate, physical activity, sexual practices, use of tobacco, alcohol and pharmaceutical compounds, dietary intake, exposure to infections, environmental pollutants and occupational hazards, etc. – epidemiologic investigation is difficult and challenging. Given this complexity, it is not surprising that from time to time epidemiological studies generate results that appear confusing, biologically absurd or even contradictory. However, it is reassuring that a wealth of new knowledge has been generated by epidemiological studies over the last few decades, and that this knowledge now lays the scientific ground for primary prevention of many major cancers and other chronic diseases among humans globally.

Given the multitude of problems with which epidemiological studies have to deal, the variability of the conditions under which they are undertaken and the abundance of potential sources of confounding and bias, it is remarkable that their results are as consistent as they are. However, it should be stressed that, in epidemiology, more than in any other field of biomedical research, it is the collective evidence that matters, rather than the results of a particular study, however large and well done the latter may be. Only the quality-adjusted collective evidence can address the issues of hidden bias and confounding that may find their ways into any particular study, and only the collective evidence can provide an approximation to generalizability.

REFERENCES

Armstrong, B. and Doll, R. (1975). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International Journal of Cancer*, **15**, 617–631.

Breslow, N. E. and Day, N. E. (1980). *Statistical Methods in Cancer Research. Vol. I. The Analysis of Case-control Studies.* IARC Scientific Publication No. **32**, 5–338 (International Agency for Research on Cancer, Lyon).

Breslow, N. E. and Day, N. E. (1987). *Statistical Methods in Cancer Research. Vol. II. The Design and Analysis of Cohort Studies.* IARC Scientific Publication No. **82**, 1–406 (International Agency for Research on Cancer, Lyon).

Buck, C. (1975). Popper's philosophy for epidemiologists. *International Journal of Epidemiology*, **4**, 159–168.

Burkitt, D. P. (1970). Geographical distribution. In: Burkitt, E. P. and Wright, D. H. (eds), *Burkitt's Lymphoma*. 186–197 (Livingstone, Edinburgh).

Cole, P. (1997). Causality in epidemiology, health policy and law. *Environmental Law Reporter*, **27**, 10279–10285.

Evans, A. (1993). *Causation and Disease: A Chronological Journey* (Plenum, New York).

Feinstein, A. R. (1995). Meta-analysis: statistical alchemy for the 21st century. *Journal of Clinical Epidemiology*, **48**, 71–79.

Greenland, S. (1980). The effect of misclassification in the presence of covariates. *American Journal of Epidemiology*, **112**, 564–569.

Haenszel, W. and Kurihara, M. (1968). Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *Journal of the National Cancer Institute*, **40**, 43–68.

Hennekens, C. H. and Buring, J. E. (1987). *Epidemiology in Medicine* (Little, Brown, Boston).

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, **58**, 295–300.

IARC (1980a). *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Vol. 23. Some Metals and Metallic Compounds*. 39–142 (International Agency for Research on Cancer, Lyon).

IARC (1980b) *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans.* Vol. 24, Some Pharmaceutical Drugs. 135–62. (International Agency for Research on Cancer, Lyon).

IARC (1987). *IARC Monographs on Carcinogenic Risks to Humans. Overall Evaluations of Carcinogenicity: an Updating of IARC Monographs Volumes 1 to 42*, Supplement 7. 440 (International Agency for Research on Cancer, Lyon).

Knox, G. (1963). Detection of low intensity epidemicity. Application to cleft lip and palate. *British Journal of Preventive and Social Medicine*, **17**, 121–127.

Lichtenstein, P., *et al*. (2000). Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine*, **343**, 78–85.

MacMahon, B. and Trichopoulos, D. (1996). *Epidemiology: Principles and Methods* (Little, Brown, Boston).

MacMahon, B. *et al*. (1960). *Epidemiologic Methods* (Little, Brown, Boston).

Miettinen, O. S. (1985). *Theoretical Epidemiology: Principles of Occurrence Research in Medicine.* (Wiley, New York).

Morgenstern, H. (1982). Uses of ecologic analysis in epidemiologic research. *American Journal of Public Health*, **72**, 1336–1344.

Parkin, D. M., *et al*. (eds) (1992). *Cancer Incidence in Five Continents*, Vol. VI. IARC Scientific Publication No. 120 (International Agency for Research on Cancer, Lyon).

Rothman, K. J. (ed.) (1988). *Causal Inference* (Epidemiology Resources, Chestnut Hill, MA).

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, **1**, 43–56.

Rothman, K. J. and Greenland, S. (1998). *Modern Epidemiology*, 2nd edn (Lippincott-Raven, Philadelphia, PA).

Sacks, H. S., *et al*. (1987). Meta-analysis of randomized controlled trials. *New England Journal of Medicine*, **316**, 450–455.

Scotto, J., *et al*. (1982). Solar radiation. In: Schottenfeld, D. and Fraumeni, J. F. (eds), *Cancer Epidemiology and Prevention*. 254–276 (W. B. Saunders, Philadelphia, PA).

Shapiro, S. (1994). Meta-analysis/Shmeta-analysis. *American Journal of Epidemiology*, **140**, 771–778.

Stuver, S. O. (1998). Towards global control of liver cancer? *Seminars in Cancer Biology*, **8**, 299–306.

Tomatis, L. (ed.) (1990). *Cancer: Causes, Occurrence and Control*. IARC Scientific Publication No. 100 (International Agency for Research on Cancer, Lyon).

Trichopoulos, D. (1994). Risk of lung cancer from passive smoking. *Principles and Practice of Oncology: PPO Updates*, **8**, 1–8.

US Department of Health, Education and Welfare (1964). *Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service*. Publication 1103 (US Government Printing Office, Washington, DC).

Wacholder, S., *et al*. (1992a). Selection of controls in case-control studies: I. Principles. *American Journal of Epidemiology*, **135**, 1019–1028.

Wacholder, S., *et al*. (1992b). Selection of controls in case-control studies: II. Types of controls. *American Journal of Epidemiology*, **135**, 1029–1041.

Wacholder, S., *et al*. (1992c). Selection of controls in case-control studies: III. Design options. *American Journal of Epidemiology*, **135**, 1042–1050.

Ziegler, R. G., *et al*. (1993). Migration patterns and breast cancer risk in Asian-American women. *Journal of the National Cancer Institute*, **85**, 1819–1827.

FURTHER READING

Authors' Work

Adami, H.-O. and Trichopoulos, D. (eds) (1998). Progress and enigmas in cancer epidemiology. *Seminars in Cancer Biology*, **8**, 215–314.

Adami, H.-O. and Trichopoulos, D. (1999). Epidemiology, medicine and public health. *International Journal of Epidemiology*, **28**, s1005–s1008.

MacMahon, B. and Trichopoulos, D. (1996). *Epidemiology: Principles and Methods*, 2nd edn (Little, Brown, Boston).

Saracci, R. and Trichopoulos, D. (1996). Aetiological leads. In: Peckham, M., *et al*. (eds), *Oxford Textbook of Oncology*. 167–174 (Oxford University Press, Oxford).

Trichopoulos, D. and Adami, H.-O. (2001). Concepts in cancer epidemiology and etiology. In: Adami, H.-O., *et al*. (eds), *Cancer Epidemiology* (Oxford University Press, Oxford).

Trichopoulos, D., *et al*. (1997). Epidemiology of cancer. In: DeVita, V. T., Jr, *et al*. (eds), *Cancer: Principles and Practice of Oncology*, 5th edn. 231–257 (Lippincott-Raven, Philadelphia, PA).

Other Publications

Breslow, N. E. and Day, N. E. (1980). *Statistical Methods in Cancer Research. Vol. I. The Analysis of Case-control Studies*. IARC Scientific Publication No. **32**, 5–338 (International Agency for Research on Cancer, Lyon).

Breslow, N. E. and Day, N. E. (1987). *Statistical Methods in Cancer Research. Vol. II. The Design and Analysis of Cohort Studies*. IARC Scientific Publication No. **82**, 1–406 (International Agency for Research on Cancer, Lyon).

Hennekens, C. H. and Buring, J. E. (1987). *Epidemiology in Medicine* (Little, Brown, Boston).

Rothman, K. J. and Greenland, S. (1998). *Modern Epidemiology*, 2nd edn (Lippincott-Raven, Philadelphia, PA).

Schottenfeld, D. and Fraumeni, J. F. (eds) (1996). *Cancer Epidemiology and Prevention*, 2nd edn (Oxford University Press, New York).

dos Santos Silva, I. (1999). *Cancer Epidemiology: Principles and Methods* (International Agency for Research on Cancer, Lyon).

Szklo, M. and Javier Nieto, F. (1999). *Epidemiology: Beyond the Basics* (Aspen, Gaithersburg, MD).

Tomatis, L. (ed.) (1990). *Cancer: Causes, Occurrence and Control*. IARC Scientific Publication No. 100 (International Agency for Research on Cancer, Lyon).

Walker, A. M. (1991). *Observation and Inference: An Introduction to the Methods of Epidemiology* (Epidemiology Resources, Boston).

Chapter 20

Mechanisms of Chemical Carcinogenesis

Nigel J. Gooderham and Paul L. Carmichael
Imperial College School of Medicine, London, UK

CONTENTS

- Introduction
- Chemicals can Cause Cancer
- Metabolism of Chemical Carcinogens
- Chemicals can Damage DNA
- Chemicals can Induce Mutations
- Examples of Chemical Carcinogens and their Metabolism
- The Biology of Chemical Carcinogenesis
- Critical Gene Targets
- Conclusion

INTRODUCTION

The concept that chemicals can cause cancer (chemical carcinogenesis) has been accepted for some time. As long ago as 1775, the English physician Sir Percival Pott noted the incidence of scrotal cancer in chimney sweeps and perceptively suggested that the disease was related to their occupation. He further suggested that it was soot that was the cause of their disease. Since this observation, the number of chemicals strongly associated with the development of cancer has substantially increased. Other notable historical examples include the development of skin tumours associated with oils and bladder cancers due to exposure to dyes and pigments. A more detailed account of the history of chemicals and cancer can be found in Lawley (1994).

The list of chemicals that can induce cancer is extensive; they can show high specificity for the organ in which the tumour is induced and in the molecular mechanisms through which they operate. Early observations of chemical carcinogenicity were often made using crude mixtures such as coal tars and subsequent studies have shown that these mixtures comprise a complex range of chemical entities. Chemical carcinogens include organics and inorganics, fibres and particulates, and biologically active materials such as hormones. The organic chemicals probably comprise the largest group.

CHEMICALS CAN CAUSE CANCER

Many chemicals are direct-acting carcinogens yet many more require metabolic activation in order to exert their carcinogenicity. In most cases it is chemical reactivity that dictates this carcinogenicity and therefore factors and environments, which influence this reactivity, are important in the nature and site of the carcinogenic action.

Organic Chemical Carcinogens

The diversity of organic chemicals that have been shown to cause cancer is considerable. Notable examples are shown in **Figure 1**. They range from low molecular mass simple halogenated hydrocarbons to very complex multiheterocyclic molecules. All possess, or have the potential to possess, key functionalities that are intimately involved in their carcinogenic action. Some of the first chemicals found to be carcinogens were the polycyclic aromatic hydrocarbons (PAHs). These chemicals are major components of coal tars and soots and application to the skin of rodents showed them to be powerful carcinogens. Many of these PAHs have been identified (**Figure 2**) and include benzo[*a*]pyrene, dibenz[*a*,*c*]anthracene, 3-methylcholanthrene, 7,12-dimethylbenz[*a*]anthracene and chrysene. All are multi-ring aromatic chemicals composed of carbon and hydrogen. Substituted PAHs in which a nitro or amino or azo function is incorporated into the structure are also carcinogenic; examples include 4-dimethylaminoazobenzene, 2-naphthylamine, benzidine and 1-nitropyrene (**Figures 1** and **2**). However, aromaticity is not an obligatory feature for chemicals that cause cancer and a group of lower molecular mass chemicals, the nitrosoamines and nitrosoamides, are equally potent carcinogens.

Other examples of cancer-causing organic chemicals include aflatoxins, which are generated by the mould *Aspergillus flavus*. Indeed, aflatoxin B_1 is one of the most potent hepatocarcinogens known, capable of inducing tumours in rodents, fish, birds and primates, including

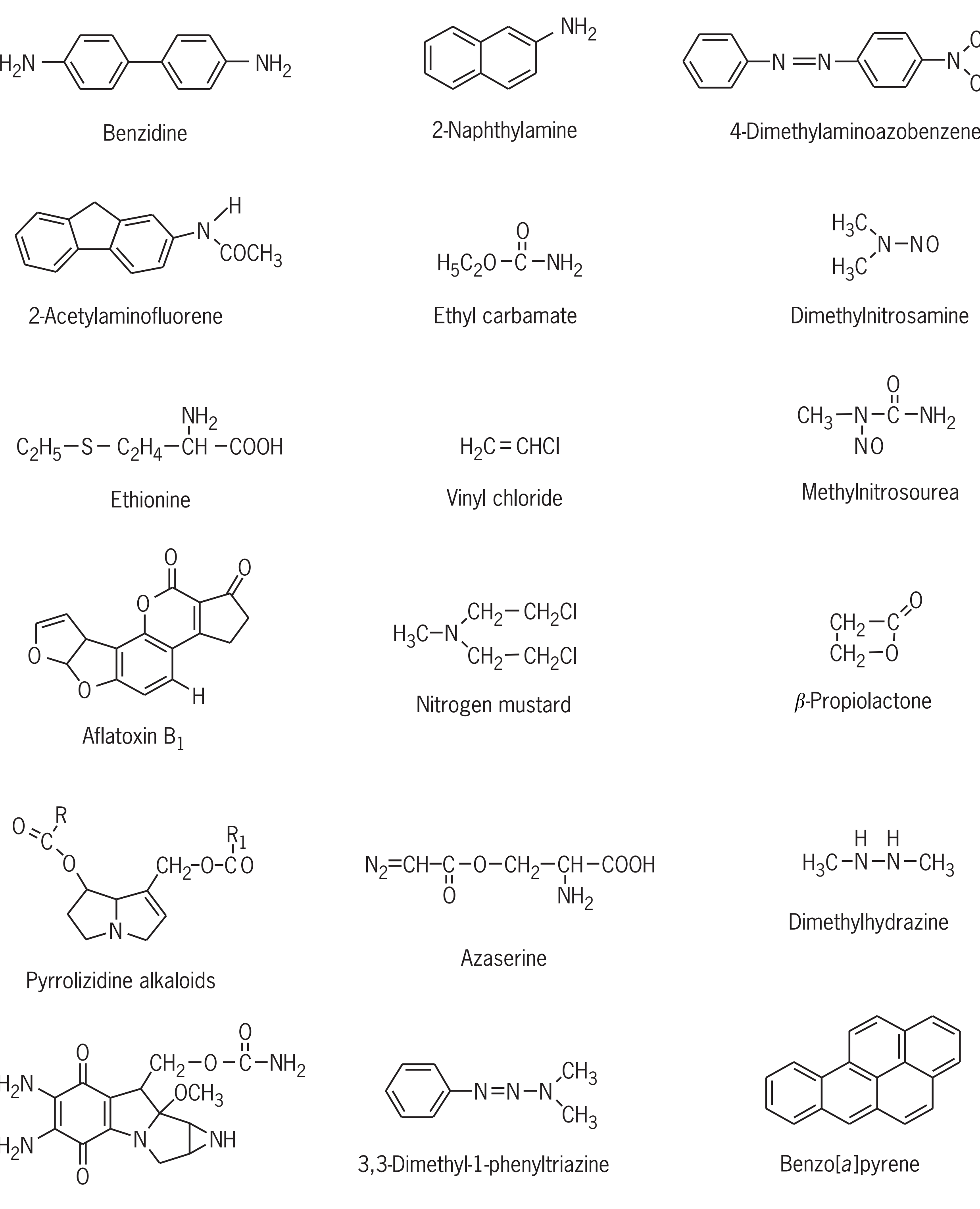


Figure 1 Examples of chemicals that can cause cancer.

humans. The mould is a contaminant of many food crops (e.g. grains and peanuts), especially when stored under warm, humid conditions, and is a significant hazard in many parts of the world, especially Africa and Asia (Wogan *et al*., 1992). However, the aflatoxins are not the only naturally produced organic carcinogens that present a hazard to humans; other notorious examples include hydrazine derivatives in certain mushrooms, cycasin from

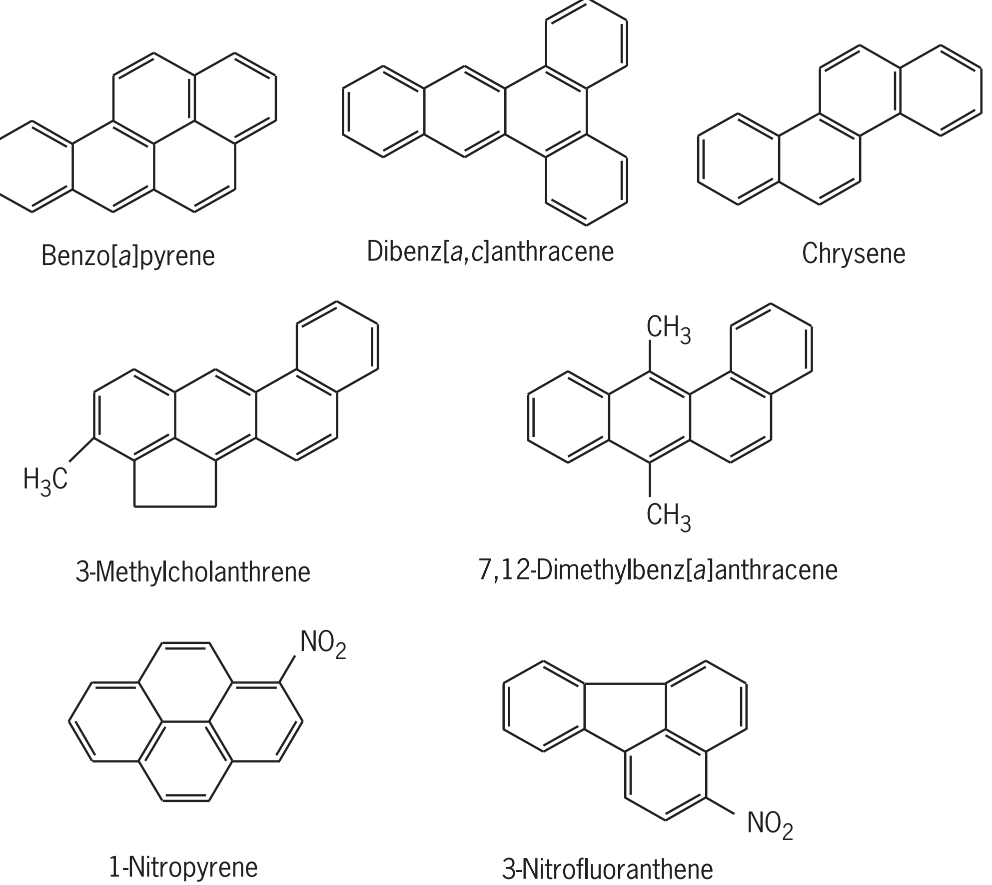


Figure 2 Examples of carcinogenic polycyclic aromatic hydrocarbons.

cycad nuts, pyrrolizidine alkaloids and ptaquiloside in various plants (see the chapter *Dietary Genotoxins and Cancer*).

Although most of the aforementioned chemicals require metabolic activation in order to exert their carcinogenicity, there are many examples of chemical carcinogens whose structure incorporates inherently reactive functional groups. Such chemicals are direct-acting carcinogens (see **Figure 3**). In each case, the reactivity of the key functional group enables the chemical to damage directly DNA, proteins and other cellular macromolecules.

Inorganic Chemical Carcinogens

Although the carcinogenic potential of many elements has not been adequately evaluated, several are known to be carcinogenic in laboratory animals and good epidemiological data support their potential as human carcinogens. Notable examples include compounds of cadmium, chromium and nickel that have all been shown to cause cancer in animals and are established human carcinogens with documented industrial exposure. In contrast, in the case of arsenic, there is significant evidence for it being a human carcinogen despite negative animal data. Other elements that are clearly carcinogenic in animals and therefore are suspect human carcinogens include beryllium, cobalt, iron, lead, titanium and zinc (Sky-Peck, 1986).

Inert Chemical Carcinogens

Some chemicals are carcinogenic even in the absence of chemical reactivity; their physical presence in tissues can be enough to form a tumour. Implantation of certain plastics and fibres into animals can induce sarcomas, usually at the site of implantation. It is the physical size and nature of the material that appear to be important for the development of cancer and not the chemical composition (Brand *et al.*, 1975). Although rodents are susceptible to these agents, other species are resistant, e.g. the guinea pig, and interestingly, implantation of plastic and other inert materials into humans, in the form of prostheses, rarely generates a sarcoma. Of more importance to humans are the fibres that are known to cause cancer; asbestos is notorious in this respect and human exposure results in mesothelioma and bronchiogenic carcinoma. The development of asbestos-induced neoplastic disease is related to the crystal structure and dimensions of the fibres rather than the chemical composition of the material (Lippmann, 1993). Fibres that are about 5 μm in length with a diameter of <0.5 μm induce mesothelioma, whereas

$ClCH_2-O-CH_2Cl$
Bis(chloromethyl) ether

H_2N-NH_2
Hydrazine

$BrCH_2-CH_2Br$
Ethylene dibromide

$(H_3CO)_2SO_2$
Dimethyl sulfate

$H_3CO(H_3C)SO_2$
Methyl methanesulfonate

$H_3C-N(NO)-CONH_2$
N-Methylnitrosourea

$H_3C-N(NO)-C(=NH)NHNO_2$
N-Methyl-*N*′-nitro-*N*-nitrosoguanidine

Epichlorohydrin

Ethylene oxide

Ethylenimine

$CH_3-N(CH_2CH_2Cl)_2$
Nitrogen mustard

$S(CH_2CH_2Cl)_2$
Sulfur mustard

Cyclophosphamide

Benzyl chloride

β-Propionolactone

Figure 3 Examples of direct-acting chemical carcinogens.

fibres that are ≥10 μm in length induce carcinoma. Since not all asbestos fibres conform to these dimensions, different types and sources of asbestos vary in their carcinogenic potency.

Hormonal Chemical Carcinogens

The first link between hormones and carcinogenesis can probably be ascribed to the Italian physician Ramazzini, who in the eighteenth century observed an increased incidence of breast cancer among nuns. It is now well established that never having children is associated with an increased breast cancer risk and, over a century after Ramazzini's observations, George Thomas Beatson (1896) pointed out that a relationship existed between breast cancer and the ovary, the major site for the production of oestrogen. Indeed, the case for endogenous oestrogens in cancer promotion is well established but increasingly concern has arisen regarding external phyto- and xeno-oestrogens in our environment and the role they may play in carcinogenesis. Furthermore, the administration of chemicals that alter the synthesis or secretion of hormones can lead to neoplastic disease. For example, chemical modulation of thyroid and pituitary growth hormone can lead to neoplasms and changes in human growth hormone, transforming or insulin-like growth factors and testosterone (**Figure 4**) are all associated with carcinogenicity under circumstances where their normal function is interfered with. In each case there is an interruption to the normal hormonal relationship experienced by the target organ. An increasing number of synthetic compounds that possess steroid hormone and antisteroid hormone activity

17β-Oestradiol

Testosterone

Diethylstilboestrol

Tamoxifen

Ethinyloestradiol

Figure 4 Examples of hormonal carcinogens.

have been found to be chemical carcinogens (**Figure 4**). One of particular interest has been the highly effective anti-breast cancer agent tamoxifen, which also possesses carcinogenic activity in the uterus in a small proportion of women using the drug (Carmichael, 1998).

METABOLISM OF CHEMICAL CARCINOGENS

Many chemicals require metabolic activation in order to exert their carcinogenic potential. The pioneering studies of Elizabeth and James Miller showed that metabolic activation of azo dyes led to their covalent binding to cellular macromolecules. They went on to show with the model carcinogen 2-acetylaminofluorene that hydroxylation of the amide nitrogen generated a metabolite that was more carcinogenic than the parent molecule. Subsequently it was found that these primary products of metabolism, although activated, could be further metabolized to even more reactive derivatives (Miller, J. A., 1970 and Miller, E. C., 1978). It was the Millers who understood that these products were potent electrophiles and comprehensively described their rapid covalent interactions with cellular macromolecules (Miller and Miller, 1981). This led to their proposal that chemical carcinogens that require such metabolic conversion in order to exert their carcinogenic effect should be called procarcinogens and that their highly reactive electrophilic metabolites were ultimate carcinogens. This further led to the concept of proximal carcinogens (e.g. *N*-hydroxy-2-acetylaminofluorene), which were intermediates between the parental procarcinogen and the ultimate carcinogenic metabolite. Although this concept has now been with us for more than two decades, the structure of many ultimate carcinogens is still not thoroughly understood and in many cases may comprise a number of different metabolites of the same parent compound.

There are a number of these metabolic pathways that together are part of a more extensive defence system, the overall role of which is ideally to process and detoxify noxious chemicals. Enzyme-catalysed and diverse in nature, these reactions have been defined and split into what are called phase I and phase II metabolism (Williams, 1971). Phase I can be separated into oxidation, reduction and hydrolytic reactions and phase II comprises a series of conjugation reactions in which a polar endogenous group is added to the xenobiotic chemical. The overall effect of this biochemistry is to convert xenobiotics, which are often lipophilic molecules, into more polar water-soluble and therefore more readily excreted products. Generally, phase I reactions unmask or introduce a functional group into the molecule and phase II metabolism conjugates the derivative with a polar water-soluble endogenous molecule, that is often acidic in nature. However, it is these same pathways of detoxification metabolism that can inadvertently bioactivate chemical carcinogens. For a more detailed description of phase I and phase II metabolism reactions and associated enzymology, the reader is directed to Jakoby *et al*. (1982), Parkinson (1996) and Gonzalez (1989).

The majority of procarcinogens are activated by mechanisms involving two-electron-mediated metabolic reactions primarily catalysed by the mixed function oxidase enzyme systems, often involving cytochrome P-450 enzymes. However, a number of one-electron reactions are known to be capable of activating xenobiotics in co-oxidation processes. For example, PAHs have been found to be bioactivated during the synthesis of prostaglandins from arachidonic acid. A key enzyme in this process is prostaglandin H synthetase, which catalyses the oxygenation of arachidonic acid to the endoperoxide prostaglandin G_2 and also has peroxidase activity, whereby it reduces the hydroperoxide prostaglandin G_2 to the alcohol prostaglandin H_2. In these reactions the peroxidase activity of the enzyme yields a free radical product that can donate electrons to xenobiotics (Eling *et al*., 1990). Other enzyme systems which can participate in these one-electron activation reactions include constitutive peroxidases such as myeloperoxidase and lactoperoxidase, both of which are capable of activating xenobiotics. Although these co-oxidation pathways are not as quantitatively important as the mixed-function oxidase activities, their presence in tissues that lack mixed function oxidase activity can be an important contributor to xenobiotic activation.

CHEMICALS CAN DAMAGE DNA

As discussed above, the metabolism of chemical agents to reactive species is a common feature of carcinogenicity. Once bioactivated (often via proximal carcinogens or intermediate chemicals formed on the way to the creation of the ultimate carcinogen), for most classical chemical carcinogens, some form of DNA damage is the norm. Because of this DNA-damaging activity, such agents are known as genetic or genotoxic carcinogens. However, this is not the case for all chemical carcinogens and some agents bring about carcinogenicity through no direct alteration or damage to the DNA. These agents can be classified as a separate group known as epigenetic carcinogens and their effects are commonly mediated through other changes involving growth factor expression or complex effects on signal transduction mechanisms (see the chapter *Non-genotoxic Causes of Cancer*). Some of the common ways in which chemical agents may be genotoxic are summarized in **Figure 5**.

The nature of carcinogen damage to DNA is dependent upon the chemical agent and its metabolism, but can often include simple changes to the DNA such as the hydroxylation of the bases. Products of such damage

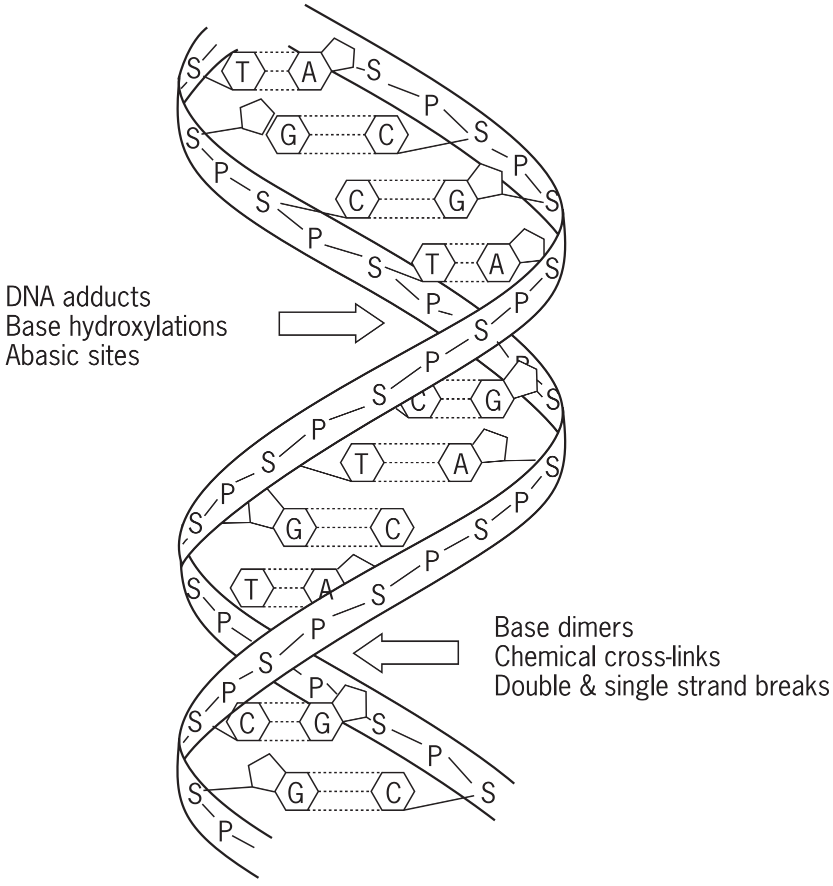


Figure 5 Carcinogen damage to DNA.

include 8-hydroxydeoxyguanosine (formed by the hydroxylation of guanine bases), which is considered to be a mutagenic lesion in DNA, and believed to be formed by attack of the DNA by highly reactive free radicals. Free radicals are a common product of a number of chemical carcinogens, which either carry an unpaired (and so-called radical) electron on the molecule, or are formed from oxygen (oxygen free radicals) via metabolism. Hydroxylations and other more extensive radical damage may result in the loss of bases and the creation of abasic sites, either apurinic or apyrimidinic. Free radical attack of DNA is also responsible for the formation of genotoxic strand breaks in the DNA. These can be formed at either one strand or may span both sides of the sugar–phosphate backbone resulting in a double-strand break.

One of the most important ways, however, in which genotoxic carcinogens may bring about DNA damage is through a chemical binding directly to the DNA. For some small molecules this binding may be considered an alkylation, although some longer chain molecules are capable of forming cross-links across the bases. For many chemical carcinogens the product of their bioactivation (via the creation of an electron- and DNA-seeking electrophile) is the formation of what are known as 'DNA adducts.' These covalent modifications of the DNA, normally at the bases, are generally considered to be one of the primary initiating events (see below) in chemical carcinogenesis and they are formed by many of the agents mentioned above, including PAHs, aflatoxins and aromatic amines. With chronic exposure to such genotoxic agents, DNA adducts may reach steady-state levels in target tissues. During cell replication the DNA adducts can result directly in mutations in genes that control cell growth and thus lead to neoplasia. The levels of these DNA adducts appear to be dose related and are generally predictive of tumour incidence across species. Thus, the accurate estimation and identification of human carcinogen-induced DNA adducts is one of the most important predictive tools or biomarker for the assessment of human cancer risk (discussed in the chapter The Formation of DNA Adducts).

The position and nature of carcinogen-induced DNA adducts dictate the type of mutation that can result (Dogliotti, 1996). For example, small alkylating agents will adduct to guanine at the N7 position owing to the highly nucleophilic nature of the site. In contrast, other more bulky aromatic amine agents will attack the purine ring preferentially, such as the C8 position of guanine and others such as diol epoxides of PAHs bind to the N2 and N6 positions. The adduct products of these reactions are converted to mutations when the cell attempts to repair the damage or replicate itself. These include point mutations, frameshifts involving loss or gain of either a single or multiple bases, chromosomal aberrations, anueploidy or polyploidy. Once introduced, the mutation becomes fixed within the DNA sequence and therefore heritable. The type of mutation that a chemical induces is dependent upon the way in which the chemical interacts with the DNA. The site of chemical attack is important but so is the influence of the bulk of the chemical and the way in which it can influence the structure of the DNA. For example, the amide 2-acetylaminofluorene differs from the amine 2-aminofluorene only by the presence of a carbonyl group. Both chemicals are metabolically activated via their *N*-hydroxy derivatives, yet the bulk of the amide–DNA adduct intercalates within the DNA, distorting the helix whereas the bulk of the amine–DNA adduct remains on the helix exterior. The consequence of these different arrangements is that the amide adduct results in frameshift mutations whereas the amine causes transversions.

The Fate of Carcinogen DNA Damage

Regardless of the specific nature of the carcinogen damage to DNA, it is important to note that aberrations at the primary sequence may not necessarily result in cancer. Indeed, there are at least three possible fates for such carcinogen damage. These are summarized in **Figure 6**.

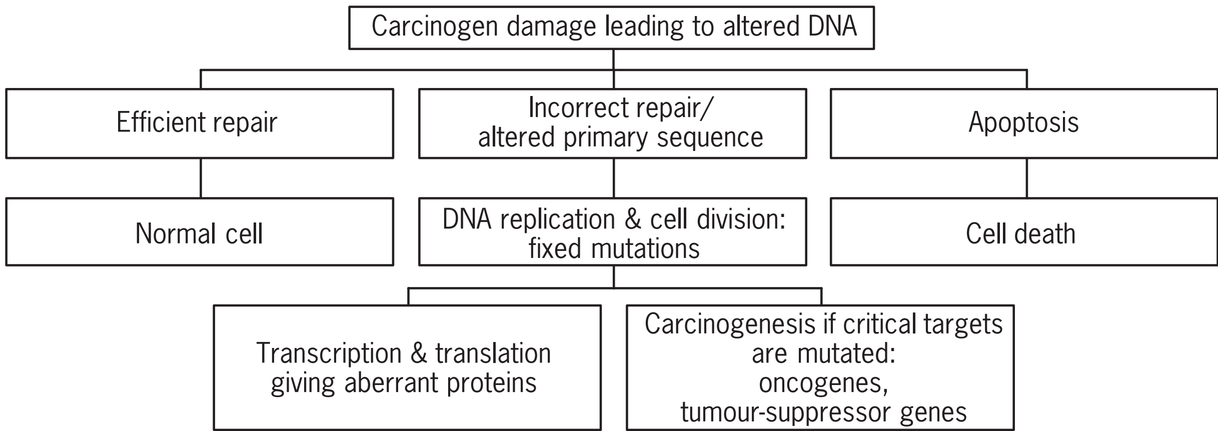


Figure 6 Possible fates for carcinogen-damaged DNA.

In some circumstances, the damage to the genetic machinery by a potential chemical carcinogen may be so extensive or be recognized as so crucial that cells initiate the process of apoptosis or programmed cell death. This behaviour and the involvement of such important proteins as p53 are discussed in detail elsewhere in this book, but through this activity, affected cells may effectively commit suicide and thus prevent the formation of potentially cancerous clones. However, if the carcinogen-altered DNA remains unchanged or if repair of the DNA is incorrect, then following DNA replication and cell division, daughter cells may contain fixed mutations. Transcription and translation of mutated genes in daughter cells will give rise to incorrect or inappropriate proteins being produced but carcinogenesis will not occur unless critical gene targets in the DNA are mutated (see later).

It follows, therefore, that repair of DNA-carcinogen damage in cells should ideally be fast, efficient and highly accurate. Indeed, mammalian cells invest a great deal of effort in order to achieve this and there are many genes and enzymes involved in maintaining the fidelity of the DNA sequence. These include enzymes that can bring about the reversal of DNA damage or repair the DNA through either base excision or nucleotide excision.

The Repair of Chemical-damaged DNA

Clearly the most straightforward way for a cell to repair a gene lesion is to remove it directly and thus regenerate the normal base at its correct position. One example of such direct removal involves the photodamage to DNA caused by ultraviolet (UV) radiation. UV radiation creates a number of photoproducts in DNA including the formation of mutagenic pyrimidine dimers between neighbouring thymine bases. Of these, the cyclobutane–pyrimidine photodimer can be repaired by a photoreactivating enzyme, although a 6–4-photoproduct cannot be repaired. The enzyme operates by binding to the photodimer and splits it back to the original bases, although the enzyme requires light energy and hence cannot work in the dark. Another enzyme involved in the direct removal of DNA damage is alkyltransferase, which can remove certain alkyl groups from altered guanine bases, although if the levels of alkylation are high enough then the enzyme activity can become saturated.

There are several pathways for excision of altered bases, often with the concomitant removal of neighbouring bases. Some systems recognize any lesion that causes a significant distortion of the DNA double helix and creates endonucleolytic cuts, several bases away and on either side of the lesion. The simple loss of a purine or pyrimidine may also initiate endonuclease cutting. The single-stranded DNA containing the damaged or missing base can then be removed and the short gap is filled in by DNA polymerases. Through this form of excision repair, bulky DNA adducts formed from the binding of PAHs and aflatoxins can be removed, as can certain UV-damaged bases. For some forms of damage (e.g. alkylation), DNA glycosylases remove just the altered base, leaving an apurinic or apyrimidinic (AP) site. The AP site is subsequently excised by endonucleases as above. Some repair pathways are capable of recognizing errors even after DNA replication has taken place. One of these systems, termed the mismatch repair system, can detect mismatches that occur during DNA replication (Friedberg, 1985). (See also chapter on *Genomic Instability and DNA Repair*.)

CHEMICALS CAN INDUCE MUTATIONS

The interaction of a chemical with DNA is not, in itself, a mutagenic event. The interaction of a xenobiotic with DNA can lead to mutagenesis due to attempts by the cell to repair the damaged DNA or during replication of the damaged DNA. Either way, a mistake may be made in which an inappropriate base is inserted or is lost (or a series of bases are inserted or are lost) from the region of DNA, with the consequence that a mutation is acquired, which is heritable. If the mutation occurs in a crucial piece of DNA, i.e. within a structural sequence that codes for an important gene controlling cell growth or a sequence that regulates the expression of such a gene, then the consequences may be disastrous. However, despite the human genome (some three billion bases) coding for approximately 28 000 genes, the majority of the DNA in a cell does not code for crucial information and mutations are usually not significant. However, the threat to the cell by DNA-damaging influences is substantial and it is not surprising that numerous proteins are involved in the detection and repair of damaged DNA.

Since mutation is evidence that a chemical is capable of inducing heritable genetic change, and since tumours often contain characteristic mutational patterns or mutational spectra in critical genes, the notion arises that the carcinogenic process is heavily dependent upon the acquisition of mutation in key genes (Harris, 1991). Furthermore, the observation that acquisition of specific types of mutation in such key genes is consistently associated with the formation of tumours, strengthens the argument that these changes are a driver of tumorigenesis. Although it is often difficult to detect and evaluate cancer-causing mutations in primary target cells and tissues, it is relatively straightforward to detect and identify mutations in surrogate reporter genes, and from information such as mutation frequency, nature and sequence context, the potential for a particular chemical to participate in the cancer process can be evaluated. This realization has driven the development of assays designed to evaluate mutagenic potential (and therefore carcinogenic potential). Currently systems are available to assess not only crude mutagenic activity, but also the likelihood of specific transitions, transversions (point mutations) or

deletions or insertions (frameshift mutations) or clastogenicity or aneuploidy. Since each of these types of DNA damage are detectable in tumours, the ability of a xenobiotic to induce such damage and therefore the relevance of the chemical to the neoplastic process can be assessed. For a more detailed description, see the chapter *Short-Term Testing for Genotoxicity*.

In vivo chemical-induced mutation is often subtle and insidious and it is usually difficult to ascribe particular mutations with exposure to a specific chemical, because evidence of the chemical is removed (loss of the DNA–chemical adduct) by introduction of the mutation. However, different classes of reactive chemicals appear to have a preference for individual DNA bases. For example, activated aromatic amines and PAHs appear to target guanine bases preferentially, probably owing to the fact that guanine is the most basic of the four nucleic acid bases. Some activated xenobiotics tend to induce transversions whereas others cause transitions or deletions. This has led to the concept of mutational fingerprints in which a specific activated xenobiotic is thought to induce a particular type of mutagenic response with a frequency which is greater than that expected by chance alone (Thilly, 1990; Aguilar *et al*., 1993; Yadollahi-Farsani *et al*., 1996). The concept of mutational fingerprints raises the prospect of identifying chemical causality in the occurrence of a particular type of mutational event, despite the fact that direct evidence for involvement (presence of the chemical) is not available.

The cooked food-borne carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) offers an example of a chemical whose mutagenic effects are characteristic and predictable (see **Table 1**). Numerous *in vitro* studies have shown this chemical to be an effective bacterial and mammalian cell mutagen (Gooderham *et al*., 1997). DNA sequence analysis of PhIP-induced mutation indicates that activation of this promutagen to damaging electrophilic species results in predominantly guanine-based mutation events. Several studies have shown the spectrum of these mutations to comprise mainly G to T transversions, with a few guanine transitions, and a significant percentage of frameshift mutations involving loss of G–C base pairs but not A–T base pairs (**Table 1**). Interestingly, PhIP-induced mutation appears to involve little or no mutation at A–T base pairs. It has also been commonly observed that PhIP-induced guanine mutation occurs primarily on the non-transcribed DNA strand, implying that damage induced on the transcribed strand is preferentially and successfully repaired. Additionally, detailed examination of the sequence context of PhIP induced mutation, shows that there appears to be a preferred motif (5′ GGGA 3′) within which the mutations are detected. It has been suggested that collectively, these observations describe a PhIP 'mutation fingerprint' and that the appearance of such a combination of mutational changes could be evidence for involvement of the chemical in their causation (Gooderham *et al*., 1996). In the case of PhIP, the nature and preferences of its DNA reactive species (PhIP nitrenium ion) tend to support the type of mutagenic outcome associated with the chemical. PhIP activated by cytochrome P-450-1 family enzymes is oxidised to the *N*-hydroxy derivative. Subsequent esterification of the *N*-hydroxy metabolite generates the acetoxy or sulfoxy derivatives, which spontaneously decompose, forming the highly reactive nitrenium ion (see **Figure 7**). The nitrenium ion shows preference for attacking guanine bases leading to the deoxyguanine-C8-PhIP (dG-C8-PhIP) and dG-N2-PhIP products. The inherent basicity of the purine molecule is a likely driver of the reaction and therefore it is no surprise that monotonous runs of guanine present as highly susceptible targets.

Table 1 PhIP induces a mutation fingerprint *in vitro* and *in vivo*

Mutation type	**Percentage of mutations**					
	hprt[a]	*hprt*[b]	*dhfr*[c]	*lacZ*[d]	*lacI*[e]	*lacI*[f]
Base substitutions	83	89	80	66	71	71
At GC pairs						
GC→TA	63	56	65	33	49	43
GC→AT	15	27	5	25	10	10
GC→CG	5	6	5	5	9	15
At AT pairs	0	0	5	3	3	3
Single bp frameshifts	13	9	5	22	26	22
(GC bp)	13	9	5	20	26	21
(AT bp)	0	0	0	2	0	1
Other mutations	4	2	15	12	3	8

[a] *hprt* gene in V79 Chinese hamster fibroblast cells (Yadollahi-Farsani *et al*., 1996)
[b] *hprt* gene in TK6 human lymphoblastoid cells (Morganthaler and Holzhauser, 1995).
[c] *dhfr* gene in CHO Chinese hamster ovary cells (Carothers *et al*., 1994).
[d] *lacI* gene in Muta™mouse mice large intestine (Lynch *et al*., 1998).
[e] *lacZ* gene in Big Blue mice large intestine (Okonogi *et al*., 1997).
[f] *lacZ* gene in Big Blue rat mammary gland (Okochi *et al*., 1999).

As discussed, the mechanisms whereby mutations arise and their chemical nature can vary considerably. In some instances, adducted DNA can base pair with its correct complementary base, apparently unaffected by the adduct whereas in other cases base pairing becomes degenerate and inappropriate bases are used during repair or replication. If the adduct is not correctly repaired, then replication can result in a mutation being fixed. In other instances, the presence of the adduct physically blocks DNA synthesis, effectively terminating the process at that point. Since such blocks would be lethal, the cell uses bypass mechanisms to overcome the block and adenine is frequently used to pair with the damaged (adducted) base. If adenine is not the original complementary base, then a mutation has been introduced into the sequence.

The ability of chemicals to induce mutation is an important aspect of their involvement in carcinogenesis. This is particularly true for the initiation process but it also

Figure 7 The metabolic activation of 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP). P-450, cytochrome P-450; NAT, *N*-acetyltransferase; ST, sulfur transferase.

contributes to promotion and progression stages during which cells continue to accumulate genetic damage at key gene sites (see later).

EXAMPLES OF CHEMICAL CARCINOGENS AND THEIR METABOLISM

2-Acetylaminofluorene

2-Acetylaminofluorene (2-AAF) is a potent mutagen and carcinogen, which induces tumours in a number of species in the liver, bladder and kidney. As previously mentioned, metabolism of this compound is the key to its carcinogenicity. 2-AAF is a substrate for cytochrome P-450 enzyme, of which there is a superfamily in mammals (Gonzalez, 1989; Porter and Coon, 1991). The product of this metabolism (**Figure 8**) is *N*-hydroxy-2-acetylaminofluorene (*N*-OH-2-AAF), which is a more potent carcinogen that the parent molecule. The formation of *N*-hydroxy metabolites has been found for many aromatic amine, amide, nitro and nitroso compounds, many of which exert toxicity through this type of derivative. For nitro and nitroso compounds, the *N*-hydroxy metabolite is formed by reduction rather than oxidation.

In the case of 2-AAF, the *N*-hydroxy metabolite is not the ultimate carcinogen but a proximal carcinogen, and this compound undergoes several enzymic and non-enzymic rearrangements. The compound can be *O*-acetylated by cytosolic *N*-acetyltransferase enzyme to yield the *N*-acetyl-*N*-acetoxy derivative, which can spontaneously rearrange to form the arylamidonium ion and a carbonium ion and interact directly with DNA to produce DNA adducts. In addition to esterification by acetylation, the *N*-OH-2-AAF can be *O*-sulfated by cytosolic sulfur transferase enzyme giving rise to the *N*-acetyl-*N*-sulfoxy product. This is again unstable and spontaneously generates the arylamidonium ion and carbonium ion, which can adduct to DNA. In addition, the cytosolic *N*,*O*-arylhydroxamic acid acyltransferase enzyme catalyses the transfer of the acetyl group from the N atom of the *N*-OH-2-AAF to the O atom of the *N*-OH group to produce *N*-acetoxy-2-aminofluorene (*N*-OH-2-AF). This reactive metabolite spontaneously decomposes to form a nitrenium ion, which will also react with DNA. However, the product of this latter reaction is the deacetylated aminofluorene adduct (see **Figure 8**). The interconversion of amide and amine metabolites of 2-AAF can further occur via the microsomal enzyme deacetylase, producing the *N*-hydroxy metabolite of the amine derivative. Subsequent esterification of the aryl hydroxylamine by sulfur transferase yields the reactive sulfate ester, which also spontaneously decomposes to form the reactive nitrenium ion. The reactive nitrenium, carbonium and arylamidonium ion metabolites of 2-AAF react with the nucleophilic groups in DNA, proteins and endogenous thiols such as glutathione (Miller and Miller, 1981). These interactions can be demonstrated *in vitro* and *in vivo*. Other metabolites such as the *N*,*O*-glucuronide, although not directly activated products, can be important in the carcinogenic process due to the fact that they are capable of degradation to proximal

Figure 8 The metabolic activation of 2-acetylaminofluorene. P-450, cytochrome P-450; DeAc, deacetylase; AHAT, *N,O*-arylhydroxamic acid acyltransferase; NAT, *N*-acetyltransferase; ST, sulfur transferase.

N-hydroxy metabolites. Thus, the *N,O*-glucuronide of *N*-OH-2-AAF is less reactive than the acetoxy and sulfate esters, yet it may be involved in the formation of bladder tumours. The mechanism for this is thought to involve degradation of the glucuronide in the bladder due to the acidic pH of urine giving rise to the *N*-hydroxy proximal carcinogen, which is then a substrate for esterification.

N-Hydroxylation of 2-AAF is not the only route of metabolism. Ring hydroxylation, catalysed by cytochrome P-450 enzymes, can also occur which generates metabolites that are not carcinogenic *per se*. These ring-hydroxylated products can be further metabolized to glucuronidated products that are readily excreted.

Benzidine

As can be seen from the example of 2-AAF, the reactivity and carcinogenicity of aromatic amides can involve their conversion to aromatic amines. Another example of this class of chemical carcinogen is benzidine. Benzidine (**Figure 9**) is a carcinogenic bifunctional aromatic amine. It can undergo several routes of metabolism, but with regard to its carcinogenicity, *N*-hydroxylation and *N*-esterification are important. The amine function at both ends of the molecule is subjected to *N*-acetylation to the corresponding amide and they can also be *N*-hydroxylated by cytochrome P-450 enzymes. The resulting aryl

Figure 9 The metabolic activation of benzidine. P-450, cytochrome P-450; NAT, *N-acetyltransferase.*

Figure 10 The metabolic activation of aflatoxin B_1. P-450, cytochrome P-450.

hydroxamic acid is unstable and rearranges to form electrophilic nitrenium ion derivatives which will rapidly interact with cellular nucleophiles such as DNA, forming DNA adducts (Searle, 1984).

Aflatoxin B_1

Aflatoxin B_1 is one of a family of mycotoxin contaminants of food crops such as grain and groundnuts. Produced by *Aspergillus flavus*, especially in hot and humid conditions, there are four main types of aflatoxin, B_1, B_2, G_1 and G_2. Aflatoxin B_1 is not only the most toxic, but is also the most carcinogenic. Contamination of food with aflatoxins is a significant problem in parts of Africa and Asia where conditions are particularly favourable for the growth of the *Aspergillus* organism. Indeed, epidemiological studies show that in areas of the world where contamination of food supplies with aflatoxins is high, there is an associated incidence of human hepatocellular carcinoma. Although contamination of grain crops with aflatoxins is not restricted to these parts of the world, in the developed countries food surveillance programmes backed up by tight legislative control regulating maximum permissible levels of aflatoxins in food crops control human exposure to these potent chemical carcinogens.

The basis of the carcinogenicity of aflatoxin B_1 centres on the carbon–carbon double bond (**Figure 10**) in the terminal furan ring of the molecule (Wogan *et al*., 1971). The chemical is a substrate for cytochrome P-450 enzymes that oxidize the carbon–carbon double bond. The epoxide product is an electrophile that rapidly reacts with cellular nucleophiles such as DNA and the resulting DNA–aflatoxin B_1 adduct is powerfully promutagenic. This and the fact that it is highly hepatotoxic makes aflatoxin B_1 one of the most powerful carcinogens known.

Benzo[*a*]pyrene

The PAHs are a large group of environmentally important chemical carcinogens. Benzo[*a*]pyrene is a prominent example of these compounds, being ubiquitous in our environment since it is present in the smoke and fumes from many diverse sources, including cigarettes, exhaust fumes and the burning of many different organic materials including wood and fossil fuels. It is one of the most extensively studied PAHs. It is a substrate for the mixed-function oxidases, being extensively converted to a variety

Bay region
K-region
Benzo[*a*]pyrene
P-450
Benzo[*a*]pyrene-7,8-oxide
EH
Benzo[*a*]pyrene-7,8-dihydrodiol
P-450
DNA adducts
Benzo[*a*]pyrene-7,8-dihydrodiol-9,10-oxide

Figure 11 The metabolic activation of benzo[*a*]pyrene. P-450, cytochrome P-450; EH, epoxide hydrolase.

of different oxidized metabolites including epoxides, phenols, diols, dihydrodiols and their conjugated products, particularly with glutathione, glucuronic acid and sulfate. The genetic toxicity of the PAHs is based around the formation of their epoxides (Sims *et al*., 1974). For example, with benzo[*a*]pyrene, cytochrome P-450-1 family enzymes can generate a series of epoxides around the different rings of the molecule, some of which are known to be more carcinogenic than others. Electronically, the most reactive portion of the benzo[*a*]pyrene molecule is the so-called 'K region' (see **Figure 11**), yet it is epoxides of the 'bay region,' which are thought to be the most tumorigenic. The formation of the ultimate carcinogen of benzo[*a*]pyrene involves cytochrome P-450-mediated epoxidation at the 7,8-position of the molecule. This epoxide is hydrolysed by the enzyme epoxide hydrolase to form the 7,8-dihydrodiol. The dihydrodiol subsequently undergoes cytochrome P-450-mediated oxidation, forming the 7,8-dihydrodiol-9,10-epoxide, which is thought to be the ultimate carcinogen. However, diastereoisomers of these metabolites are formed metabolically and the (+)-benzo[*a*]pyrene, (7*R*,8*S*)-dihydrodiol-(9*S*,10*R*)-epoxide formed by cytochrome P-450 1A1 and epoxide hydrolase generates a species that is more mutagenic than other isomers. Interestingly, although the K region 4,5-epoxide is highly mutagenic, the bay region 7,8-dihydrodiol-9,10-epoxide is carcinogenic whereas the 4,5-epoxide and the 9,10-epoxide are not, indicating the requirement for the specific dihydrodiol formation. Other metabolites of benzo[*a*]pyrene are known to be cytotoxic and mutagenic, e.g. the 3,6-quinone derivative, but have poorer carcinogenic potential (Conney, 1982).

Dimethylnitrosamine

The nitrosamines are another extensively studied family of chemical carcinogens. Dimethylnitrosamine, representative of this family, is hepatotoxic, mutagenic and carcinogenic, causing kidney tumours with acute exposure and liver tumours after chronic exposure (Magee *et al*., 1976). Additionally, tumours of the stomach, oesophagus and central nervous system are found. Again, cytochrome P450 enzymes are central to the metabolic activation of dimethylnitrosamine, involving *N*-demethylation to *N*-methylnitrosamine. This metabolite rearranges to form methyldiazohydroxide, then methyldiazonium ion and ultimately methyl carbonium ion. It is the methyl carbonium ion that is the DNA-damaging species since it is a highly reactive alkylating agent (see **Figure 12**). The degree of DNA methylation that occurs after exposure to dimethylnitrosamine correlates with susceptibility to tumour formation. The DNA sites that appear to be most sensitive to alkylation are the N^7 position of guanine and to a lesser extent the O^6 position of guanine, but it is the latter site that appears to correlate best with mutagenicity and carcinogenicity. Alkylation at guanine O^6 leads to guanine–thymidine mispairing, causing a GC to AT transition. Paradoxically, acute doses of dimethylnitrosamine cause greater DNA methylation in the liver than the kidney, yet tumours are preferentially found in the kidney. Clearly, although the liver has more metabolic activation capacity, it also has better protective mechanisms. This balance between activation and protective mechanisms is also likely to be a significant factor in the susceptibility

$$H_3C(H_3C)N-N=O \xrightarrow{P\text{-}450} HOH_2C(H_3C)N-N=O \longrightarrow [H_3C-N=N-OH] + HCOH$$

Dimethylnitrosamine → Hydroxymethyl+ nitrosamine

$\overset{+}{CH_3} + N_2 + H_2O$ (Methylcarbonium ion) → DNA adducts

Figure 12 The metabolic activation of dimethylnitrosamine. P-450, cytochrome P-450.

$H_2C=CHCl$ (Vinyl chloride) $\xrightarrow{P\text{-}450}$ $H_2C-CHCl$ (epoxide, O bridge) (Vinyl chloride epoxide) → $O=C(H)-CH_2Cl$ (Chloroacetaldehyde); → DNA adducts

Figure 13 The metabolic activation of vinyl chloride. P-450, cytochrome P-450.

of other organs and tissues to dimethylnitrosamine-induced carcinogenesis.

Vinyl chloride

Vinyl chloride is a simple halogenated allyl compound that is extensively used in the plastics industry, being the starting point for a number of polymer syntheses, particularly the manufacture of poly(vinyl chloride). A gas at room temperature and therefore usually stored as a liquified gas under pressure, the use of the material in an industrial setting is substantial and numerous workers have suffered well-documented accidental occupational exposure, for example being overcome by vinyl chloride solvent narcosis during the cleaning of reaction vessels. Acute exposure to the material is associated with an unusual form of liver tumour, known as haemangiosarcoma, a tumour of reticuloendothelial cells, and not hepatocytes, giving rise to tumours of the hepatic vasculature. This very rare type of liver cancer has only been observed in workers who have been exposed to vinyl chloride.

For tumorigenicity, vinyl chloride requires metabolic activation by cytochrome P-450 enzymes (Bolt, 1988). Oxidation to the epoxide intermediate (**Figure 13**) is the first step resulting in the subsequent formation of chloroacetaldehyde. Although a number of other metabolites are known to be generated, the tumorigenicity of the compound is likely to be dependent upon the epoxide and chloroacetaldehyde metabolites, both of which react with cellular nucleophiles. Subsequently glutathione is depleted and the excess reactive metabolites can then react with nucleic acids and protein. The available evidence suggests that the epoxide is the predominant species binding to DNA where it reacts with deoxyguanosine at the N^7 position and with RNA to give $1,N^6$-ethenoadenosine and $3,N^4$-ethenocytidine, whereas the chloroacetaldehyde probably binds to protein.

THE BIOLOGY OF CHEMICAL CARCINOGENESIS

The involvement of chemicals in damaging DNA is only one part of their potential role in carcinogenesis. In progressing to the neoplastic state, cells must undergo fundamental changes in their biology and many of these changes can be driven by chemical intervention of essential cellular activity. Current perception of the neoplastic process has been shaped by fundamental studies performed in the 1940s using PAHs with the mouse skin carcinogenesis model (Berenblum and Shubik, 1947). From such studies emerged the concept of the multistep nature of carcinogenesis and the defining of these steps as initiation, promotion and subsequently progression (see **Table 2**).

Initiation

A key feature of the initiation process (**Table 2**) is the requirement for cell replication (Pitot and Dragan, 1996). Once a chemical has damaged DNA by forming an adduct, or inducing a strand break, etc., and provided that the cell recognizes the damage, repair processes will intervene. Should the damage be misrepaired, or not recognized by the cell, then the outcome can be (1) aberrant transcription of the affected DNA if it is a structural gene lesion and on the transcribed strand, (2) altered expression if the damage is located within a regulatory sequence or (3) no effect if the damage is within noncoding or nonregulatory DNA. If the damage is at a critical gene site or extensive, then the cell may opt to apoptose. However, since the cell has two copies of each autosomal gene loci, compensation for the

Table 2 The stages of initiation, promotion and progression

Initiation	Promotion	Progression
Irreversible	Reversible	Irreversible
Cell not morphologically recognizable		Morphologically discernable
Single exposure to chemical is sufficient	Requires multiple exposure to chemical	Driven by multiple and varied exposure
Can occur spontaneously	Endogenous and exogenous chemicals can promote	Cells show altered sensitivity to endogenous chemicals
'Fixed' by cell division	Expanded through cell division	Driven by cell division and migration
No obvious dose–response threshold	Measurable dose–response threshold and maximum effect	Effect of individual chemicals difficult to discern
Involves simple mutation (transitions, transversions, frameshifts)	Does not necessarily involve mutation	Frequently involves complex mutation
Can involve single genes	Often involves multiple genes	Usually involves chromosomal alterations (clastogenicity, aneuploidy, polyploidy)
Proto-oncogenes and tumour-suppressor genes can be mutated	Altered levels of gene expression rather than mutation	Substantial alteration in gene expression due to multiple mutational events
Simple genotypic changes with corresponding phenotype change	Epigenetic phenotype changes	Complex phenotypic change
Limited growth advantage over surrounding cells	Chemical mediated selective growth advantage	Significant growth advantage due to genetic change and environment

(Adapted from Pitot and Dragan, 1996.)

damaged gene is often possible. The only cell directly affected is the recipient of the damage, the initiated cell. However, if division occurs, a daughter cell will inherit the mutation, which becomes fixed irrespective of whether it was misrepaired or remained as an unrepaired adduct in the parent cell, and will be inherited in all future generations of this lineage. If the cell type is generally quiescent, then the effect of the DNA lesion is unlikely to become apparent unless the cell experiences a proliferative stimulus (e.g. by promotion). However, if the cell is normally proliferative (e.g. a stem cell), then there is a much greater chance of neoplastic development, even in the absence of concerted promotion.

Promotion

Once the initiated cell has been encouraged to replicate, the initiated genetic damage is irreversibly fixed. Yet the phenotypic characteristics of the initiated cell remain insidious because it is surrounded by (and perhaps compensated by) normal cells. When the initiated cell is subjected to promotional influences, the effect is to encourage clonal expansion of the initiated cell. A key feature of the promotion response is its reversibility, i.e. if the promoting agent is removed, the initiated cell population is no longer encouraged to proliferate (Boutwell, 1964). The promotional stimulus is usually not a direct interaction of the promoting agent on DNA but acts to encourage growth and proliferation of the cell by exaggerating favourable environmental conditions, especially for initiated cells (see **Table 2**). This can be achieved by chemical interference with normal cellular signal transduction mechanisms with the consequence of enhanced transcription and translation.

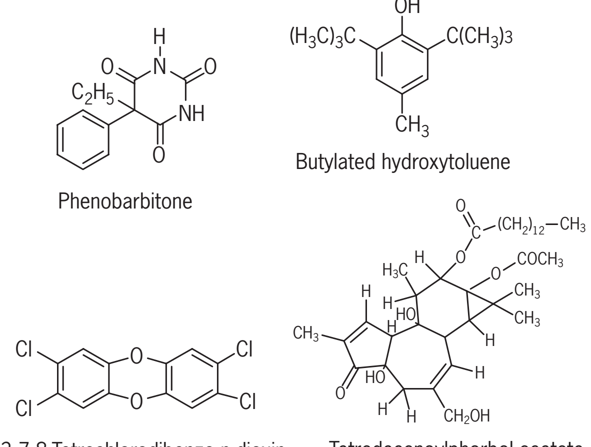


Figure 14 Examples of chemical promoters.

A number of well-characterized promoting agents, both natural and xenobiotic, are shown in **Figure 14**. A good example is tetradecanoylphorbol acetate (TPA), a naturally occurring component of croton oil. This chemical interacts with the key protein, protein kinase c, causing enhanced signalling via inositol triphosphate and diacylglycerol pathways, resulting in increased transcription and strong stimulation of growth. 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD) is another example of a xenobiotic with exceptional promoting activity, being particularly effective in the liver, lung and skin. There are also a number of very powerful naturally occurring promoting agents including androgens and oestrogens (Taper, 1978). In all cases removal of the agent can rapidly lead to regression of the preneoplastic lesion.

Another important characteristic of promoting agents is the existence of concentration thresholds. Thus a minimum dose of agent is required before effective promotion can be detected, after which there is a dose–response relationship before a maximum response is reached. This latter effect is probably due to saturation of physiological targets, usually endogenous receptors. Other physiological factors can alter the impact of promoting chemicals, e.g. diet and hormonal status will influence the promotion of numerous preneoplastic lesions including those in the mammary and hepatic tissues.

Progression

As the initiated cell population is expanded during promotion, individual cells acquire further genetic damage, which can be either agent-mediated or spontaneous, thereby introducing genetic heterogeneity into the promoted population. Acquisition of these changes is irreversible and progressive (**Table 2**). Ultimately, some promoted cells will acquire karyotype instability, which can bestow growth advantage over surrounding cells. In some cases this will remove environmental constraints over such cells, allowing them to escape proximal cellular and humoural influences. In many cases, such cells will have acquired clastogenic changes that bestow significant phenotypic change (Foulds, 1965; Welch and Tomasovic, 1985).

CRITICAL GENE TARGETS

The reaction between an activated carcinogen and DNA is concentration dependent. Since DNA comprises only four bases then, in theory, chemicals should damage DNA at any site within the genome. In practice there appears to be preferred sites at which attack occurs. The factors which govern the chance of attack at such sites would be expected to include accessibility, the presence of modifying influences (extent of methylation or association with nuclear proteins), etc. Furthermore, factors such as (1) DNA repair, (2) the presence of transcriptionally active or silent genes, (3) the role of the gene in normal cell function, (4) whether the chemical attacks a structural domain of the gene or (5) a regulatory domain and (6) whether the DNA target has any function whatsoever, will affect outcome. Although DNA is chemically the same throughout the genome, some sections (containing critical genes) are clearly more important to the correct functioning of the cell than are others. The number of currently recognized critical gene targets is only a fraction of the estimated 28 000–30 000 genes of the human genome, but is growing. These critical genes are involved in controlling the growth and differentiation of cells, damage or mutation of which is highly correlated with the neoplastic process; such genes are described as oncogenes or tumour-suppressor genes.

Oncogenes

Under normal circumstances oncogenes exist as proto-oncogenes (Garrett, 1986); they code for proteins that are involved in important pathways within the cell such as growth, signal transduction, cell cycle and nuclear transcription. Chemical-induced damage to key sites within the proto-oncogene can convert the sequence to one that codes for an aberrant protein. The product of this genetic change (conversion to an oncogenic variant of the gene) is a protein that has an altered sequence due to loss, gain or replacement of amino acids, and no longer responds to appropriate controlling stimuli. The activation of proto-oncogenes to oncogenes can be achieved by several well-recognized processes (Pitot, 1986). These may include (1) mutation of the structural gene, (2) mutation of regulatory sequences, (3) amplification of the structural gene, (4) translocation of the gene to a site of inappropriate gene regulation (e.g. within the control of a powerful promoter sequence) after clastogenic events or (5) epigenetic alteration of either structural or regulatory sequences leading to altered regulation of gene expression (e.g. altered gene methylation). Each of these mechanisms generates a functionally dominant gene and can be induced by chemical interaction at key sites within DNA. Examples of specific oncogenes are described in **Table 3** and can be grouped into functional biochemical classes, which code for key regulatory activities involved in correct functioning of the cell. These genes tend to be involved in pathways that mediate cell growth, differentiation and signalling. The ability of mutagenic chemicals to activate oncogenes is well established. A good example is the codon 12 K-*ras* mutations found in aberrant crypt foci in the colon of rats treated with azoxymethane (Shivapurkar *et al.*, 1994).

Tumour-suppressor Genes

Tumour-suppressor genes normally act to suppress the neoplastic phenotype. As with oncogenic activation, chemical damage to a tumour-suppressor gene can give rise to a protein with altered activity, in this case making it unable to operate as a suppressor of the cancer process (Knudson, 1993). Since these genes tend to be functionally recessive, full loss of phenotypic expression often requires mutation or loss of both alleles and chemicals tend to destroy or alter the normal cellular activity of the tumour-suppressor gene. Many of these genes have a nuclear location/function and are directly involved in maintaining genome integrity. Examples of tumour-suppressor genes are given in **Table 3**. Chemical-mediated damage of the *p53* gene is a frequently encountered event in chemical-induced neoplastic disease. For example, aflatoxin B_1 exposure is associated with a high incidence of hepatocellular carcinoma in parts of Africa. Examination of the *p53* gene in these tumours reveals characteristic mutations at specific

Table 3 Examples of critical genes involved in carcinogenesis

Gene	Function	Localization
Oncogenes		
SIS, *FGF*, *INT2*, *WNT1*	Growth factors	Extracellular
MET, *NEU*, *EPH*, *EGRF*, *FMS*, *KIT*, *HER2*, *RET*, *ROS*,	Receptor/protein tyrosine kinases	Extracellular/cell membrane
SRC, *ABL1*, *FPS*, *FGR*, *FYN*, *HCK*, *LCK*, *YES*	Nonreceptor tyrosine kinases	Cell membrane/cytoplasmic
MAS	Receptors lacking protein kinase activity	Cell membrane/cytoplasmic
RAS, *GIP2*, *GSP*	Membrane-associated G proteins	Cell membrane/cytoplasmic
BCR, *DBL*, *ECT2*	RHO/RAC-binding proteins	Cytoplasmic
RAF, *PIM1*, *BCR*, *EST*, *MOS*,	Cytoplasmic protein serine kinases	Cytoplasmic
STY	Protein serine, threonine and tyrosine kinase	Nuclear
BCL1, *CRK*, *ODC*, *NCK*	Cytoplasmic regulators	Cytoplasmic
MYC, *FOS*, *JUN*, *BCL3*, *CBL* *ERBA*, *ETS*, *HOX*, *MYB*, *MYCL*, *REL*, *TAL1*, *SKI*	Nuclear transcription factors	Nuclear
BCL-2	Mitochondrial membrane factor	Mitochondrial/cytoplasmic
Tumour-suppressor genes		
NF1	GTPase activation	Cell membrane/cytoplasmic
RB-1	Cell cycle-regulated nuclear transcription repressor	Nuclear
P53	Cell cycle-regulated nuclear transcription repressor	Nuclear
WT1	Zinc finger transcription factor	Nuclear
HMLH1	Mismatch DNA repair	Nuclear
BRCA1	DNA repair enzyme	Nuclear
APC	Regulates cytoskeletal networks	Cytoplasmic
DCC	Cell adhesion molecule	Plasma membrane
VHL	Signal transduction or cell–cell contact	Plasma membrane
NME	Cell receptor	Plasma membrane
CMAR/CAR	Cell attachment	Plasma membrane
WNT	Growth factor	Extracellular matrix
YES1	Tyrosine kinase	Plasma membrane

(Adapted from Hesketh, 1994.)

hot spots (e.g. codons 248/249). These tumours and p53 mutations can be reproduced *in vivo* in animals and *in vitro* in cells exposed to pure aflatoxin B_1 (Aguilar *et al*., 1993).

CONCLUSION

Cancer is often perceived as largely due to inherited defective genes, a view fuelled by recent discoveries in this area. However, the biggest cause of human cancer is the influence of chemical carcinogens. Indeed, the biggest contributors to the burden of human cancer (estimated to be about 90% of all cancers) are carcinogens contained within tobacco smoke, diet and the environment (Doll and Peto, 1981).

Carcinogenesis is a multistep process and many of these steps can be influenced by chemicals. Clearly a chemical with carcinogenic or mutagenic properties should be used with caution, especially in cases where it would come into contact with humans. Yet there will be some instances where the benefit of using the chemical will outweigh its immediate risk, e.g. as a chemotherapeutic agent. Under such circumstances its use must be carefully regulated and monitored.

As discussed, many chemicals are carcinogenic through metabolic bioactivation and genotoxic activity at the level of the genome. However, this primary activity is only one factor in a sequence of complex events. Of thousands of chemicals tested, a few hundred have been identified as carcinogenic in rodents and around 50 are human carcinogens. With the development of new pharmaceuticals and industrial chemicals, carcinogenicity testing has become an essential activity and has led to a clearer understanding of the mechanistic basis of the carcinogenicity process.

The value of understanding the mechanisms whereby a chemical causes cancer cannot be underestimated. It provides a rational basis for risk assessment allows predictions of possible outcomes and informs the development of alternatives or strategies, which reduce risk. In the case of

therapeutically useful agents, it indicates how improved pharmaceuticals can be conceived and employed. It is also increasingly evident that each chemical with carcinogenic properties is unique and that grouping chemical carcinogens on the basis of their structural attributes is not always appropriate and thus prediction of genetic toxicity should always be confirmed by experimental data. In this respect, the development of new technologies such as transgenic animals that have altered expression of proto-oncogenes and tumour-suppressor genes greatly add to the battery of systems available for helping us understand cancer. Such models will continue to provide substantial information on the mechanisms through which chemicals and biological processes interact to drive the neoplastic process.

REFERENCES

Aguilar, F., *et al*. (1993). Aflatoxin B1 induces the transversion of G→T in codon 249 of the p53 tumour suppressor gene in human hepatocytes. *Proceedings of the National Academy of Sciences of the USA*, **90**, 8586–8590.

Beatson, G. T. (1896). On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment with illustrative cases. *Lancet*, **2**, 104–107.

Berenblum, I. and Shubik, P. (1947). A new quantitative approach to the study of the stages of chemical carcinogenesis in the mouse's skin. *British Journal of Cancer*, **1**, 383–391.

Bolt, H. M. (1988) Roles of etheno–DNA adducts in tumorigenicity of olefins. *CRC Critical Reviews in Toxicology*, **18**, 299–309.

Boutwell, R. K. (1964). Some biological aspects of skin carcinogenesis. *Progress in Experimental Tumor Research*, **4**, 207–250.

Brand, K. G., *et al*. (1975). Etiological factors, stages, and the role of the foreign body in foreign body tumourigenesis. A review. *Cancer Research*, **35**, 279–286.

Carmichael, P. L. (1998). Mechanisms of action of antiestrogens: relevance to clinical benefits and risks. *Cancer Investigation*, **16**, 604–611.

Carothers, A. M., *et al*. (1994). Mutation and repair induced by the carcinogen 2-(hydroxylamino)-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (N-OH-PhIP) in the dihydrofolate reductase gene of Chinese hamster ovary cells and conformational modelling of the dG-C8-PhIP adduct in DNA. *Chemical Research in Toxicology*, **7**, 209–218.

Conney, A. H. (1982). Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: GHA Clowes Memorial Lecture. *Cancer Research*, **42**, 4875–4917.

Dogliotti, E. (1996). Mutational spectra: from model systems to cancer-related genes. *Carcinogenesis*, **17**, 2113–2118.

Doll, R. and Peto, R. (1981). *The Causes of Cancer* (Oxford University Press, Oxford).

Eling, T. E., *et al*. (1990). Prostaglandin H synthatase and xenobiotic oxidation. *Annual Review of Pharmacology and Toxicology*, **30**, 1–45.

Foulds, L. (1965). Multiple etiologic factors in neoplastic development. *Cancer Research*, **25**, 1339–1347.

Friedberg, E. C. (1985). *DNA Repair* (Freeman, New York).

Garrett, C. T. (1986). Oncogenes. *Clinica Chemica Acta*, **156**, 1–40.

Gonzalez, F. J. (1989). The molecular biology of cytochrome P450s. *Pharmacological Reviews*, **40**, 243–288.

Gooderham, N. J., *et al*. (1996). Heterocyclic amines: evaluation of their role in diet associated human cancer. *British Journal of Clinical Pharmacology*, **42**, 91–98.

Gooderham, N. J., *et al*. (1997). Assessing human risk to heterocyclic amines. *Mutation Research*, **376**, 53–60.

Harris, C. C. (1991). Chemical and physical carcinogenesis: advances and perspectives for the 1990s. *Cancer Research*, **51**, 5023s–5044s.

Hesketh, R. (1994). *The Oncogene Handbook* (Academic Press, London).

Jakoby, W. B., *et al*. (1982). *Metabolic Basis of Detoxification: Metabolism of Functional Groups*. 1–375 (Academic Press, New York).

Knudson, A. G. (1993). Antioncogenes and human cancer. *Proceedings of the National Academy of Sciences of the USA*, **90**, 10914–10921.

Lawley, P. D. (1994). Historical origins of current concepts of carcinogenesis. *Advances in Cancer Research*, **65**, 17–111.

Lipmann, M. (1993). Biophysical fibres affecting fibre toxicity. In: Wahrheit, D. B. (ed.), *Fibre Toxicology*. 259–303 (Academic Press, San Diego).

Lynch, A. M., *et al*. (1998). Genetic analysis of PhIP intestinal mutations in Muta^{TM} Mouse. *Mutagenesis*, **13**, 601–605.

Magee, P. N., *et al*. (1976). *N*-Nitroso compounds and related carcinogens. In: Searle, C. E. (ed.), *Chemical Carcinogens*. 491–625 (American Chemical Society, Washington, DC).

Miller, E. C. (1978). Some current perspectives on chemical carcinogenesis in humans and experimental animals: Presential Address. *Cancer Research*, **38**, 1479–1496.

Miller, E. C. and Miller, J. A. (1981). Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer*, **47**, 2327–2345.

Miller, J. A. (1970). Carcinogenesis by chemicals: an overview – GHA Clowes Memorial Lecture. *Cancer Research*, **30**, 559–576.

Morgenthaler, P. M. and Holzhauser, D. (1995). Analysis of mutations induced by 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) in human lymphoblastoid cells. *Carcinogenesis*, **16**, 713–718.

Okochi, E., *et al*. (1999). Preferential induction of guanine deletion at 5′-GGGA-3′ in the rat mammary glands by 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP). *Carcinogenesis*, **20**, 1933–1938.

Okonogi, H., *et al*. (1997). Agreement of mutational characteristics of heterocyclic amines in lacI of Big Blue mouse with

those in tumour related genes in rodents. *Carcinogenesis*, **18**, 745–748.

Parkinson, A. (1996). Biotransformation of xenobiotics. In: Klaassen, C. D., *et al*. (eds), *Casarett and Doull's Toxicology: the Basic Science of Poisons*. 113–186 (McGraw-Hill, New York).

Pitot, H. C. (1986). Oncogenes and human neoplasia. *Clinical Laboratory Medicine*, **6**, 167–179.

Pitot, H. C. and Dragan, Y. P. (1996). Chemical carcinogenesis. In: Klaassenm, C. D., *et al*. (eds), *Casarett and Doull's Toxicology: the Basic Science of Poisons*. 201–267 (McGraw-Hill, New York).

Porter, T. D. and Coon, M. J. (1991). Multiplicity of isoforms, substrates and catalytic and regulatory mechanisms. *Journal of Biological Chemistry*, **266**, 13469–13472.

Searle, C. E. (1984). *Chemical Carcinogens*, 2nd edn (American Chemical Society, Washington, DC).

Shivapurkar, N., *et al*. (1994). Sequence analysis of k-ras mutations in aberrant crypt foci and colonic tumours induced by azoyxmethane in Fisher-344 rats on high-risk diet. *Carcinogenesis*, **15**, 775–778.

Sims, P., *et al*. (1974). Metabolic activation of benzo[*a*]pyrene proceeds by a diol epoxide. *Nature*, **252**, 326–328.

Sky-Peck, H. H. (1986). Trace metals and neoplasia. *Clinical Physiology and Biochemistry*, **4**, 99–111.

Taper, H. S. (1978). The effect of estradiol-17-phenylpropionate and estradiol benzoate on *N*-nitrosomorpholine-induced liver carcinogenesis in ovariectomised female rats. *Cancer*, **42**, 462–467.

Thilly, W. G. (1990). Mutational spectrometry in animal toxicity testing. *Annual Review of Pharmacology and Toxicology*, **30**, 369–385.

Welch, D. R. and Tomasovic, S. P. (1985). Implications of tumour progression on clinical oncology. *Clinical and Experimental Metastasis*, **3**, 151–188.

Williams, R. T. (1971). *Detoxification Mechanisms*, 2nd edn (Wiley, New York).

Wogan, G. N. (1992). Aflatoxins as risk factors for hepatocellular carcinoma in humans. *Cancer Research*, **52**, 2114s–2118s.

Wogan, G. N., *et al*. (1971). Structure–activity relationships in toxicity and carcinogenicity of aflatoxins and analogues. *Cancer Research*, **31**, 1936–1942.

Yadollahi-Farsani, M., *et al*. (1996). Mutational spectra of the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo [4,5-*b*]pyridine (PhIP) at the Chinese hamster *hprt* locus. *Carcinogenesis*, **17**, 617–624.

FURTHER READING

Balmain, A. and Brown, K. (1988). Oncogene activation in chemical carcinogenesis. *Advances in Cancer Research*, **51**, 147–182.

Dipple, A., *et al*. (1985). Metabolism of chemical carcinogens. *Pharmacology and Therapeutics*, **27**, 265–296.

Guengerich, F. P. (1992). Metabolic activation of carcinogens. *Pharmacology and Therapeutics*, **54**, 17–61.

Hodgson, E. and Levi, P. E. (1994). *Biochemical Toxicology*, 2nd edn (Appleton and Lange, Norwalk, CT).

Levine, A. J. (1993). Tumour suppressor genes. *Annual Review of Biochemistry*, **62**, 623–651.

Timbrell, J. (2001). *Principles of Biochemical Toxicology*, 3rd edn (Taylor and Francis, London).

Vainio, H. (1992). *Mechanisms of Carcinogenesis in Risk Identification*. *IARC Publications*, No. 116 (International Agency for Research on Cancer, Lyon).

Chapter 21

The Formation of DNA Adducts

David H. Phillips
Institute of Cancer Research, Sutton, UK

CONTENTS

- DNA Damage by Carcinogens
- Adducts Formed by Chemical Carcinogens
- Methods of DNA Adduct Detection
- Evidence for the Biological Significance of DNA Adducts
- DNA Adduct Dosimetry
- Adducts Detected in Human Tissues
- Adducts in Urine
- Tobacco Exposure
- Endogenous DNA Adducts
- Risk Assessment
- Gene–Environment Interactions
- Conclusion

DNA DAMAGE BY CARCINOGENS

A property common to many chemical carcinogens is that they, or one or more of their metabolites, are DNA reactive. Cellular responses to DNA damage in mammalian cells include DNA repair, cytotoxicity, apoptosis, mutagenesis and transformation to malignancy. These processes are either fundamental to maintaining the integrity of the cell or they set the cell on a path to mortality or malignancy. Thus the study of reactions between carcinogens and DNA, and the biological consequences of these reactions, is central to understanding the early stages of the carcinogenic process.

The identification of DNA as the genetic material and the solving of its structure occurred about 50 years ago, but it was not immediately appreciated that carcinogens exert their biological effects by damaging it. Although the discovery that carcinogens could form covalent bonds with cellular macromolecules dates from about the same time, for some years afterwards the prevailing hypothesis was that the deletion of key proteins was critical to the carcinogenic process. However, the demonstration in the 1960s that the potency of a series of carcinogens correlated with their ability to bind to DNA *in vivo*, and not with the extent of binding to protein or RNA (Brookes and Lawley, 1964), led to the acceptance of DNA as the critical target in carcinogenesis. Subsequently, the discovery of several classes of genes that control cellular function and maintain cellular integrity and which are commonly mutated in tumours (oncogenes, tumour-suppressor genes and mismatch repair genes) has made it evident that if certain critical genes are modified by carcinogens, the mutations that may ensue from erroneous replication of the damaged gene template will contribute to the transformation of a normal cell into a malignant one. Finally, the types of genetic alterations commonly found in tumour cells – point mutations, deletions, translocations, gene amplifications – can also be induced in cells by treatment with DNA adduct-forming chemicals. Given the long latency of most types of human and experimental cancer, it is not possible to observe directly the biological consequences of DNA adduct formation *in vivo*. Nevertheless, the very strong correlation between this early biochemical event and the subsequent biological manifestation of malignancy in many different studies places the cause-and-effect association beyond reasonable doubt.

Carcinogen-induced DNA damage can take several forms. It can result in breaks in the sugar–phosphate backbone of the molecule, either in one of the two strands of the double helix, or in both. Covalent binding of the carcinogen results in the formation of a chemically altered base (or, occasionally, phosphate group) in DNA that is termed an adduct. As DNA adducts are usually studied by fragmenting the DNA either chemically or enzymatically, the term nucleotide adduct describes a fragment consisting of carcinogen–base–deoxyribose–phosphate, nucleoside adduct consists of carcinogen–base–deoxyribose and a base adduct is the carcinogen-modified base only. Some carcinogens are bifunctional and can give rise to both monoadducts and cross-links in DNA, the latter being either intra- or interstrand cross-links. Many cancer chemotherapeutic agents have this property, and it is widely held

that interstrand cross-links are cytotoxic (accounting for the therapeutic properties of the drugs), whereas the monoadducts and intrastrand cross-links are potentially mutagenic and carcinogenic (Lawley and Phillips, 1996).

Most chemical carcinogens are not biologically active as such, but undergo metabolic activation in mammalian cells to reactive intermediates that react with DNA (see also the chapter *Mechanisms of Chemical Carcinogenesis*). Why do mammalian cells do this? The answer is that it is an aberration of the general mechanisms that cells employ to rid themselves of toxins, generally by making them water soluble. Metabolism of xenobiotic (foreign) compounds is carried out in mammalian cells by broad spectrum oxidative enzymes (Phase I metabolism) that introduce polar groups (e.g. hydroxyl groups) into molecules and render the molecules suitable substrates for conjugation (Phase II metabolism) with one of a variety of hydrophilic groups. The resulting conjugate is substantially more water soluble than the parent compound and thus more readily excreted from the organism. Phase I metabolites are often formed through transient generation of reactive compounds such as epoxides, but are rapidly converted to hydroxyl groups by further enzymatic conversion or by nonenzymatic reaction with water. However, if they are slow to convert, their presence in the cell may lead to their reaction with DNA. Similarly, most Phase II metabolites are water soluble and chemically stable, ideal properties for their efficient elimination from the cell and subsequent excretion from the organism; however, some carcinogens are converted to conjugates that are reactive, and the loss of the conjugated function generates a highly reactive carbocation that reacts with DNA.

There have been two approaches to determining the pathways of activation of carcinogens. In the first, metabolites of the carcinogen are isolated and identified, and their abilities to induce tumours and other genotoxicity endpoints are investigated. In the second approach, DNA adducts are detected and identified, and the pathways by which they are formed are deduced by determining the metabolites that can also give rise to them, and the cofactors necessary for their formation.

An example of a carcinogen that is metabolically activated to a reactive Phase I metabolite is benzo[*a*]-pyrene (BaP), a polycyclic aromatic hydrocarbon (PAH) that is widespread in the environment through its formation during the incomplete combustion of organic material (wood, coal, petrol, tobacco, etc.). Although it is converted to many different metabolites, one cytochrome P450(CYP)-dependent pathway results in metabolic activation (**Figure 1**). The initial epoxide formed is rapidly metabolized further by epoxide hydrolase to a dihydrodiol (called the proximate carcinogen), but this metabolite then undergoes further metabolism to a dihydrodiol epoxide, BPDE (the ultimate carcinogen). This time, the molecule is not a good substrate for epoxide hydrolase and although it is chemically unstable (i.e. it is reactive), it is sufficiently long-lived in mammalian cells to be able to migrate to the nucleus and react with DNA, via formation of a carbocation, to form chemically stable DNA adducts (Phillips, 1983).

An example of a carcinogen activated by Phase II metabolism is tamoxifen (**Figure 2**). In the liver of rats, where it causes hepatocellular carcinoma, it is converted first to α-hydroxytamoxifen (the proximate carcinogen) by a CYP-dependent Phase I step. This metabolite then undergoes Phase II metabolism by sulfotransferase to a sulfate ester conjugate. This compound dissociates to form

CYP1A1
EH
7,8-Oxide
7,8-Dihydrodiol
CYP1A1
Reaction with DNA
7,8-Dihydrodiol 9,10-oxide

Figure 1 Major pathway of metabolic activation of benzo[*a*]pyrene. CYP1A1, cytochrome P450 1A1; EH, epoxide hydrolase.

Figure 2 Major pathway of metabolic activation of tamoxifen. CYP3A4, cytochrome P450 3A4; HST, hydroxysteroid sulfotransferase (SULT2A family).

Figure 3 DNA bases and sites at which they are bonded to carcinogens.

a reactive carbocation that reacts with DNA to form DNA adducts (Davis *et al*., 1998).

This chapter describes the formation of DNA adducts by carcinogenic chemicals, with particular emphasis on their detection in human tissues, and the role of such studies in investigating the aetiology of cancer, in monitoring human exposure to environmental carcinogens and in determining cancer risk.

ADDUCTS FORMED BY CHEMICAL CARCINOGENS

Most of the covalent binding of carcinogens to DNA involves the modification of the purine and pyrimidine bases, although some agents also react with the phosphodiester linkages. Guanine is the most commonly and extensively modified base, with interactions occurring at N^2, N-3, O^6, N-7 and C-8 (**Figure 3**). Adducts with adenine are formed at the N-1, N-3, N^6 and N-7 atoms. Pyrimidine adducts are formed at the O^2, N-3, N^4 and C-5 of cytosine and at the O^2 and O^4 of thymine.

Alkylating agents that react by a predominantly S_N2 mechanism show a greater affinity for the ring nitrogens in DNA bases, these being the most nucleophilic sites in DNA (Osborne, 1984). As the S_N1 character of the reaction increases (i.e. via generation of an electrophilic carbocation that is the reactive species), so the proportion of the reaction that occurs at exocyclic groups increases (Lawley, 1984). Thus the ratio of O^6/N-7 alkylation of guanine in DNA is 0.7 for *N*-ethylnitrosurea (ENU), 0.1 for *N*-methylnitrosourea (MNU) and 0.004 for methyl methanesulfonate (MMS) (Lawley, 1984).

Representative examples of active metabolites of carcinogens and of their reaction sites with DNA are shown in **Table 1**. More comprehensive reviews of carcinogen–DNA interactions can be found elsewhere (Osborne, 1984; Cooper and Grover, 1990; Hemminki *et al*., 1994). PAHs, activated by diol epoxide formation, mainly form stable DNA adducts at the exocyclic amino groups of guanine and adenine. The more distorted from planarity the ultimate carcinogen, the greater is the proportional reactivity with adenine relative to guanine (Dipple, 1994). Tamoxifen also reacts at these sites. Nitroaromatic polycyclic hydrocarbons, activated at the nitro group, react with the C-8 position of guanine. Aromatic amines and heterocyclic amines also modify predominantly the C-8 position of

Table 1 Some representative carcinogens, their active metabolites and sites of modification of DNA

Carcinogen	Major active metabolite	Site of modification of DNA
Benzo[*a*]pyrene (BP)	BP 7,8-diol 9,10-epoxide (HO, OH, O)	N^2-Guanine, N^6-adenine
Benzo[*c*]phenanthrene (BcPh)	BcPh 4,3-diol 2,1-epoxide (O, OH, OH)	N^6-Adenine, N^2-guanine
Aflatoxin B_1 (AfB_1) (O, O, O, O, O, CH_3)	AfB_1 8,9-epoxide (O, O, O, O, O, O, CH_3)	*N7*-Guanine
Tamoxifen (O, $N(CH_3)_2$)	α-Hydroxytamoxifen sulfate (O, $N(CH_3)_2$, $^{-}O_3SO$)	N^2-Guanine, N^6-adenine
2-Acetylaminofluorene (AAF) (NHAc)	*N*-Acetoxy-AAF (NAc, OAc)	*C8*-Guanine, N^2-guanine
Vinyl chloride (Cl)	Chloroethylene oxide (O, Cl)	3,N^4-Cytosine, 1,N^6-adenine, 3,N^2-guanine
2-Amino-1-methyl-6-phenylimidazo-[4,5-*b*]pyridine (PhIP) (CH_3, N, NH_2, N, N)	*N*-Acetoxy-PhIP (CH_3, N, NHOAc, N, N)	*C8*-Guanine
1-Nitropyrene (1-NP) (NO_2)	*N*-Hydroxy-1-aminopyrene (NHOH)	*C8*-Guanine

Figure 4 Conversion of the unstable aflatoxin-*N7*-guanine adduct in DNA to stable products.

guanine. Aflatoxin B_1, a mycotoxin, is metabolized to an epoxide that reacts at the N-7 position of guanine. This adduct structure is electronically charged and unstable. It can undergo one of three processes spontaneously: (1) hydrolysis to yield aflatoxin dihydrodiol and unmodified DNA; (2) depurination to yield an aflatoxin-base adduct, leaving an apurinic site in DNA; or (3) imidazole ring-opening to yield a stable adduct (**Figure 4**). Some compounds form cyclic adducts, in which two positions on the same base are modified by the same molecule. An example is vinyl chloride (and other vinyl halides), which forms etheno adducts with cytosine, adenine and guanine (**Figure 5**).

The large variety of DNA sites attacked by carcinogens leads to the question of whether some modifications are

1,N^6-Ethenoadenine 3,N^4-Ethenocytosine N^2,3-Ethenoguanine

Figure 5 Structures of etheno adducts formed by vinyl halides and by products of lipid peroxidation.

more biologically important than others. This is still a matter of debate, with some suggestions that adenine adducts formed by PAHs are more consequential than guanine adducts, even though benzo[*a*]pyrene forms very few of the former. Others have proposed that unstable, depurinating adducts are more important than stable adducts in causing mutations from which tumour initiation proceeds, although this theory has been challenged. Substitution at N-7 of guanine by simple alkylating agents appears to be ineffectual in causing mutations, whereas adducts at the O^6 position are highly promutagenic . On the other hand, aflatoxin B_1 appears to modify only N-7 in guanine, and is among the most potent carcinogens known (Osborne, 1984). Methylation of DNA at the O^6 position, a base-pairing position, will cause guanine to mispair with thymine, thereby causing a $G \rightarrow A$ transition after a round of replication (Lawley, 1984). The 7-position of guanine, however, is non-pairing; methylation produces little or no alteration to the tertiary structure of DNA, but substitution with a bulky molecule like aflatoxin B_1 would cause considerable distortion (and also leads to ring opening of the purine), which could in itself lead to replication errors by DNA polymerases.

There is also some evidence that DNA modification by carcinogens does not occur at random, but is influenced to some extent by DNA sequence (Osborne, 1984).

Additionally, the potential for an adduct to give rise to a mutation may be dependent, to some extent, on its sequence context, giving rise to the concept of hotspots for DNA damage and mutation in some genes. Some of these possibilities are discussed later in this chapter.

METHODS OF DNA ADDUCT DETECTION

The last 25 years have seen the development of a number of sensitive methods for the detection and characterization of DNA adducts in mammalian cells and tissues (Phillips, 1990; Strickland *et al*., 1993; Phillips *et al*., 2000; Poirier *et al*., 2000), the most important of which are reviewed briefly here.

Radiolabelled Compounds

Because only a very small proportion of an applied dose of a carcinogen becomes bound to DNA in the exposed cells or tissue, very sensitive methods of detection are required to study DNA adduct formation *in vivo*. Most of the early work on adducts was done using radiolabelled carcinogens and, although other methods now provide comparable or greater sensitivity, the method still has its uses. With compounds labelled either with ^{3}H or ^{14}C, at a position of the molecule where the isotope will not be lost as a result of metabolism, detection of radioactivity in DNA isolated from the exposed animal or cultured cells is the starting point for the characterization of the DNA binding. Sensitivities of detection of one adduct in 10^{8} nucleotides are achievable with ^{3}H labelling, although ^{14}C labelling is less sensitive because of the much longer half-life of ^{14}C than ^{3}H (Phillips *et al*., 2000).

Limitations to the use of the method are the high costs of synthesizing radiolabelled compounds and the difficulty in doing chronic, multidose, exposure studies due to the hazards of the use of radioactive materials in these circumstances. Furthermore, it is seldom possible to use radioactive test compounds in studies involving human subjects. However, the recent adaptation of accelerator mass spectrometry (AMS) enables isotope ratios to be measured with great sensitivity in biological samples. With this method, the binding of a radiolabelled carcinogen to DNA is detected not by means of its decay (thereby linking the sensitivity to the half-life of the isotope) but by measuring the abundance of the radioisotope relative to that of the natural isotope. For ^{14}C it is possible to detect one part in 10^{15} parts total carbon, and in practice limits of adduct detection of greater than one in 10^{11} nucleotides have been achieved (Phillips *et al*., 2000). Because of this high sensitivity, it has been possible to obtain ethical approval to give minute amounts of a radioactive carcinogen, e.g. the mutagen formed in cooked food, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP), to human subjects prior to surgery and to detect DNA adducts in the excised tissue. AMS does not give structural information on DNA adducts, and characterization requires chromatographic comparison with synthesized standards. Nevertheless, its ultra-sensitivity provides a means of establishing whether carcinogens thought to be nongenotoxic in their mechanism of action, such as 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), are truly devoid of DNA binding activity.

^{32}P Postlabelling

The ^{32}P postlabelling method of analysis comprises a procedure for introducing a radiolabel into a DNA adduct after it has formed, by enzymatic phosphorylation of the deoxyribose group of the nucleotide adduct (Phillips, 1997). Chromatographic separation of the labelled adducts followed by detection and quantitation by measuring ^{32}P decay provides a highly sensitive assay, requiring only small (1–10 μg) quantities of DNA. Adducts from different classes of carcinogens with diverse structures can be detected by this method, including PAHs, aromatic amines, heterocyclic amines, unsaturated aldehydes, simple alkylating agents, reactive oxygen species and UV radiation (Beach and Gupta, 1992). It is also able to detect adducts formed from complex mixtures of chemicals, such as tobacco smoke and fossil fuel products. It is sensitive enough to detect adduct levels as low as one per mammalian cell. A limitation of the method is that it does not provide structural information; identification of adducts is reliant on co-chromatography with characterized synthetic standards. Adduct levels may be underestimated if the DNA is not completely digested or if the nucleotide adduct is not efficiently ^{32}P labelled by polynucleotide kinase (Phillips *et al*., 2000).

Immunoassays

Antisera elicited against carcinogen–DNA adducts can be used in immunoassays to detect adducts in human or animal tissues. Antibodies have been raised against a variety of carcinogen-modified DNAs, including those containing adducts of PAHs, aromatic amines, methylating agents, tamoxifen, UV radiation and oxidative damage. Immunoassays are highly sensitive but generally require more DNA for analysis than ^{32}P postlabelling. The assay is relatively inexpensive to perform and can be automated. Various sensitive methods (radioactive, colorimetric, fluorescent and chemiluminescent) have been developed for detecting bound antiserum. When combined with histochemistry it can be used to localize adducts within biological samples. Antibodies raised against a particular adduct can show cross-reactivity with adducts formed by other carcinogens of the same class, which can obscure both the nature of the adducts detected and the levels at which they are present. Nevertheless, this cross-reactivity does at least afford the opportunity to use immunoaffinity chromatography as a means of extracting adducts of a

specific class, such as PAHs, for further analysis by other methods.

Mass spectrometry and other physicochemical methods

Mass spectrometry is the most chemically selective method for DNA adduct detection and it can provide unequivocal identification of the nature of an adduct. This selectivity comes at a price, that of sensitivity, which has limited its application to human DNA adduct studies. However, it is a method in which technological advances are being made rapidly, and it can be predicted with confidence that in the near future it will provide much valuable data on the nature of DNA modifications in human tissues, as well as in experimental studies (Phillips *et al*., 2000; Poirier *et al*., 2000). Most studies to date have used mass spectrometry combined either with gas chromatography (GC-MS) or liquid chromatography (LC-MS). Earlier methods required volatility to be a property of a molecule for its characterization by mass spectrometry, necessitating derivatization of polar species such as adducts, but softer ionization conditions (e.g. fast atom bombardment (FAB)) have overcome this limitation to some extent (Weston, 1993; Phillips and Farmer, 1995). With the advent of newer methods, such as electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), it will also become possible to investigate the presence of adducts in intact (high molecular mass) biomolecules including DNA and protein.

A number of carcinogens form adducts that are highly fluorescent, enabling their detection by fluorescence spectroscopy. These include adducts formed by PAHs and aflatoxins, cyclic (etheno) adducts and some methylated adducts. They can be analysed in intact DNA, DNA digests or as hydrolysis products. Low-temperature fluorescence spectroscopy (e.g. fluorescence line narrowing spectroscopy (FLNS) (Jankowiak and Small, 1991)) reveals considerable spectral fine structure that is lost at ambient temperature and that can be diagnostic for certain carcinogens, also revealing information about the conformation of the adducts.

Some adducts, notably the oxidative DNA lesion 8-hydroxy-2′-deoxyguanosine, are readily detected by high-performance liquid chromatography (HPLC) combined with electrochemical detection (ECD) (Halliwell and Dizdaroglu, 1992). The method has also been used to detect *N7*-methylguanine in combination with immunoaffinity purification.

Another specific method that can be used to detect adducts by means of their elemental content is atomic absorption spectrometry. An example of its use is the detection of platinum bound to DNA as a result of the treatment of cancer patients with platinum drugs (Weston, 1993). The method gives only the level of the element present, and no structural information.

EVIDENCE FOR THE BIOLOGICAL SIGNIFICANCE OF DNA ADDUCTS

While it is widely assumed that the formation of DNA adducts is an early and obligatory event in the process by which many carcinogens initiate tumours, it is by no means a sufficient event, and the long delay between carcinogen treatment and tumour appearance precludes a direct cause-and-effect demonstration. Nevertheless, it is the case that inhibition of DNA adduct formation will decrease the incidence of tumours formed subsequently, and increasing the adduct levels generally leads to a higher tumour yield.

That chemical modification of DNA can result in the same alterations as observed in mutated genes in tumours was observed with the H-*ras*-1 proto-oncogene transfected into NIH3T3 cells. Prior modification of the plasmid containing the gene resulted in mutations occurring in the DNA after transfection and replication of the host cells, manifested as the appearance of transformed foci. Mutations that activate *ras* genes occur in a few codons in the gene, so correlations between the sites of mutations in such experiments are less informative than in genes where there are many possible sites of DNA damage and mutation which can lead to altered function of the gene product. Such a gene is *p53*, where correlations can be usefully sought between the mutation spectra observed in different human tumours and clues sought to the nature of the initiating agent (Hainaut *et al*., 1998). This has led to strong evidence for the involvement of aflatoxin B_1 in the initiation of liver cancer in regions of high incidence in China. With respect to lung cancer, codons 157, 248 and 273 of the *p53* gene are frequently mutated in these tumours. $G \rightarrow T$ transversions are much more common in lung cancer of smokers than that of non-smokers, and these types of mutation are characteristic of bulky carcinogens, such as the PAHs which are present in tobacco smoke. When the sites of DNA adduct formation by benzo[*a*]pyrene diol epoxide, the reactive metabolite of benzo[*a*]pyrene, in the *p53* gene in HeLa cells and bronchial epithelial cells were determined, it was found that codons 157, 248 and 273 were preferentially modified (Denissenko *et al*., 1996).

Ultraviolet (UV) radiation causes DNA damage chiefly by dimerization of adjacent pyridines in the same DNA strand. The biological importance of these lesions is illustrated by the fact that sufferers of xerodema pigmentosum (XP), who have a deficiency in nucleotide excision repair mechanisms that remove pyrimidine dimers from DNA, are prone to sunlight-induced skin cancer. Moreover, mutations in the *p53* gene found commonly in such tumours, but rarely in tumours of internal organs, are tandem mutations occurring at pyrimidine pairs ($CC \rightarrow TT$ transitions), highly suggestive that they arose from UV-induced pyrimidine dimers (Dumaz *et al*., 1993).

Thus, there are examples of genetic changes in tumours that closely match the genetic changes that can be induced

experimentally in cellular DNA by specific genotoxic agents. These tumour-specific mutations in *p53* and the demonstration that chemically modified DNA transforms cells show that the mutations observed in human tumours could have arisen from the formation of carcinogen–DNA adducts *in vivo*. Clonal expansion of the mutated cells and the acquisition of further genetic alterations eventually leads to malignancy (Fearon and Vogelstein, 1990).

Another piece of evidence that strongly links DNA adduct formation to tumour initiation is the demonstration that XPA knockout mice, which are deficient in nucleotide excision repair, are highly sensitive to tumour induction by carcinogens that form stable adducts that would be removed from DNA in normal mice by this repair mechanism (van Steeg *et al*., 1998).

DNA ADDUCT DOSIMETRY

From a number of animal studies, it has been demonstrated that at chronic low doses, there is a linear relationship between the amount of carcinogen administered and the level of DNA adducts that results (Poirier and Beland, 1992). Adducts can be lost from DNA by depurination and by DNA repair, and can be diluted by DNA replication and cell division, so if the exposure is chronic, a steady-state level will be attained; typically this takes about 1 month of dosing. Where exposure is acute or of limited duration, it is clear from animal studies that a small proportion of the adducts persist in tissues for long periods, even in the presence of cell proliferation (e.g. in skin). Thus, the detection of DNA adducts can provide evidence of prior exposure to carcinogens even if the exposure was limited to a single dose.

It is also evident that DNA adducts may be formed in some tissues in the absence of tumour formation, indicating that adduct formation alone may not be sufficient for carcinogenesis (Poirier *et al*., 2000). Other tissue-specific events, such as cell proliferation, are required. Nevertheless, tumours do not form in the absence of adducts, and interventions that reduce adduct formation (e.g. co-administration of enzyme inhibitors or inducers) result in inhibition of carcinogenesis.

An analysis of adduct levels and tumour rates in experimental animals for 27 different chemicals has shown that the adduct levels required to produce a 50% incidence in liver tumours in rats or mice varies between 53 and 5543 adducts per 10^8 nucleotides (Otteneder and Lutz, 1999). This is a narrow range considering the diversity of the chemicals and their interactions with DNA. The analysis also shows that low levels of adducts are detectable in animal bioassays in which tumours were not observed. Thus it is accepted that there are uncertainties concerning the biological significance of low levels of DNA adduct formation, but there are not, as yet, sufficient data with which to define a threshold below which adduct levels can be deemed biologically irrelevant (Phillips *et al*., 2000).

ADDUCTS DETECTED IN HUMAN TISSUES

Studies of humans occupationally exposed to carcinogens have demonstrated the formation of DNA adducts in human tissues (see also the chapter *Occupational Causes of Cancer*). Many of these studies have involved exposure to PAHs with adduct formation being monitored in white blood cells or peripheral lymphocytes; for example, adduct levels are elevated in iron foundry workers, coke oven workers, aluminium plant workers, bus drivers exposed to diesel exhausts and roofers (Phillips, 1996). The same is true of residents of polluted regions of Poland, the Czech Republic and China (Perera, 2000). In some studies, dietary exposure to PAHs appears to be a stronger determinant of adducts in blood cells than does occupational exposure, so it is important that such factors be taken into consideration in interpreting the results of human biomonitoring studies.

Although DNA adducts of some sort have been detected in many studies of human tissues, there are still only a few studies in which the nature of the adduct has been unequivocally identified, and fewer still in which the origin of the DNA binding species can be defined. These limitations derive from the fact that many studies, particularly those employing ^{32}P postlabelling analysis, rely on the co-chromatography of the human adducts with synthetic standard adducts, rather than providing structural identification as such; although this can give reasonably reliable indications of the nature of the adducts formed *in vivo*, it cannot be considered sufficient evidence of identification. Thus those examples listed in **Table 2** all come from studies in which unambiguous physicochemical data (e.g. mass spectra or fluorescent spectra) on the properties of the adducts were obtained. In some cases, the nature of the DNA damaging agents can be deduced from the adduct, but in others, there is uncertainty or ambiguity because there is more than one potential source of the adduct.

As a general rule, several types of adduct have been detected in human DNA, at various levels. Thus 8-hydroxyguanine, originating from oxidative and free radical processes (see section on *Endogenous DNA Adducts*), is typically at levels of one in 10^4–10^5 nucleotides; etheno adducts, from lipid peroxidation or vinyl halides, are formed at between one in 10^7 and one in 10^8 nucleotides; O^6-methylguanine (formed by alkylating agents) is typically at one in 10^6–10^7 nucleotides; and bulky adducts (arising from smoking, pollution and diet) are often found at one in 10^7–10^8 nucleotides. With the exception of smoking-related adducts (see section on *Tobacco Exposure*) these can be regarded as approximate ‘background’ levels of DNA damage in human tissues due to environmental and/or endogenous DNA-damaging agents. Thus,

Table 2 Identified DNA adducts in human tissues

Adduct	Tissues
2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP, a food mutagen/heterocyclic amine) adducts	Colon
4-Aminobiphenyl (ABP, an aromatic amine) adducts	Bladder
4-(*N*-methyl-*N*-nitrosamino)-1-(3-pyridyl)-1-butanone (NNK, a tobacco-specific nitrosamine) adducts	Lung
Benzo[*a*]pyrene (BaP, a polycyclic aromatic hydrocarbon) adducts	Placenta, lung, leukocytes
Aflatoxin B_1 (AfB_1, a mycotoxin) adducts	Urine
Malondialdehyde-guanine	Liver, leukocytes
Thymine glycol	Placenta
5-Hydroxyethyluracil	Leukocytes
N7-(2-Hydroxyethyl)guanine	Liver
N^2,3-Ethenoguanine	Liver
8-Hydroxyguanine	Many tissues

when measuring adducts in a group of individuals suspected of being highly exposed to carcinogens and/or at elevated risk of cancer, it is important that such measurements be compared with the levels in a suitably selected control group of individuals.

DNA adducts have been detected in cancer patients undergoing chemotherapy (with, for example, platinum-based drugs, cyclophosphamide, melphalan, mitomycin C and the methylating agents dacarbazine and procarbazine) and studies have revealed considerable inter-individual variation in adduct levels among patients receiving similar dosage (Lawley and Phillips, 1996). It remains to be established whether such measurements may provide indications of therapeutic response.

ADDUCTS IN URINE

Because some adducts are chemically unstable and cause depurination of DNA, the excreted products, modified bases, can be detected in the urine (Shuker and Farmer, 1992). Examples of this include aflatoxin–guanine adducts where food contamination by this mycotoxin is endemic (e.g. in areas of Africa and China). Ethylated bases in urine can also serve as markers of exposure to ethylating agents. The method is of less use as a biomarker of exposure to methylating agents because there are instances in which methylated bases (e.g. 3-methylguanine) are ingested in the diet. Another caution is that carcinogen–purine adducts in urine may derive from both DNA and RNA. Urine contains significant numbers of exfoliated bladder epithelial cells, whose DNA can be isolated and analysed for the presence of adducts. An example is the case of a worker acutely exposed to MOCA (4,4′-methylenebis(2-chloroaniline)), whose urine samples yielded significant levels of adducts in the exfoliated urothelial cells for some weeks following the exposure.

TOBACCO EXPOSURE

The relationship between DNA adduct formation and tobacco smoke has been a fruitful area of research in which to explore the validity of the biomarker for a number of reasons (see also the chapter *Tobacco Use and Cancer*). First, a large proportion of the human population is regularly and habitually exposed to tobacco smoke because smoking is an addiction. Second, tobacco smoke contains at least 50 compounds that are known to be carcinogenic, including representatives of several distinct classes of compounds (PAHs, aromatic amines, *N*-nitrosoamines, azaarenes, aldehydes, other organic compounds and inorganic compounds). Most of these compounds are genotoxic carcinogens that form DNA adducts. Third, epidemiological studies have provided clear evidence that tobacco smoking causes not just lung cancer but also cancers in many other organs.

Many studies have compared DNA from smokers, ex-smokers and nonsmokers and have found that the levels of adducts in smokers are elevated in many target tissues: lung, bronchus, larynx, bladder, cervix and oral mucosa (Phillips, 1996). In some of these studies a linear correlation between estimated tobacco smoke exposure and adduct levels has been observed. In tissues of the respiratory tract adduct levels in ex-smokers tend to be intermediate between smokers and nonsmokers, indicating that adducts are removed through DNA repair and/or cell turnover. The half-life of adduct persistence appears to be between 1 and 2 years. This value is longer than would be predicted from adduct persistence studies in animals. A possible explanation is that the lungs of an ex-smoker continue to accumulate adducts after cessation of smoking owing to the continued presence in the lung of smoke and tar deposits.

For some of these studies specific adducts have been detected, but in others a more general measure of DNA damage has been made, namely aromatic/hydrophobic adducts detected by ^{32}P postlabelling, or PAH–DNA

adducts detected by immunoassay. Recent studies have found that when adduct levels are adjusted to take account of the level of tobacco smoke exposure, lung DNA from women smokers is more highly adducted that that of male smokers. This finding is interesting in view of the epidemiological evidence that women are at a 1.5–2-fold greater risk of lung cancer from smoking. It would appear that the adduct analysis provides biochemical, mechanistic evidence to support the morbidity data. The reason for the higher level of adduction in women could be related to levels of expression of metabolizing enzymes (e.g. CYP enzymes) that activate tobacco smoke carcinogens to DNA binding species (Mollerup *et al*., 1999).

Some, but not all, studies have shown elevated levels of lung adducts in cancer cases compared with controls. The relationship between adduct levels in target tissues (e.g. lung) and other tissues (e.g. blood) has been investigated to see whether the latter can serve as a more readily accessible surrogate source of DNA than the former. Results for smoking-related adducts have been inconsistent (Perera, 2000; Poirier *et al*., 2000), perhaps because other sources of exposure to some classes of carcinogens, such as the PAHs, which are also ingested as dietary contaminants, may contribute to the overall level of adducts in the blood but not to the same extent in the lung.

It should be emphasized that by measuring adducts in smokers at the time of cancer diagnosis, investigators are not looking at the biochemical events causal in the initiation of those tumours, as these would have occurred decades earlier. However, because smoking is addictive and habitual for most tobacco users, DNA adducts in tumour-adjacent tissue at the time of tumour manifestation can still serve as a useful biomarker that gives an indication of an individual's probable steady-state level of DNA damage maintained over a long period of time.

ENDOGENOUS DNA ADDUCTS

Thus far, the emphasis has been on exogenous, environmental sources of carcinogens, but there is a large body of evidence that shows that DNA is also subject to modification by a number of endogenous processes arising from normal metabolism, oxidative stress and chronic inflammation (Marnett and Burcham, 1993; Marnett, 2000). **Table 3** shows a list of endogenous DNA adducts that have been detected in human tissues.

The most abundant oxidized base is 7,8-dihydro-8-oxo-2′-deoxyguanosine (8-oxo-dGuo), but at least 30 other base modifications have been characterized in oxidized DNA. Some of these, including 8-oxo-dGuo, can be formed as a result of free radical attack on DNA (for example, by the hydroxyl radical) whereas others appear to a consequence of normal aerobic metabolism. Being the most abundant oxidative lesion, 8-oxo-dGuo is often used as a biomarker for oxidative DNA damage in humans.

Etheno adducts (see **Figure 5**) can be generated in DNA as a result of lipid peroxidation (Marnett and Burcham, 1993; Marnett, 2000), making them adducts of endogenous origin. Etheno bases are removed from DNA by a repair mechanism involving glycosylases, but the rate of removal appears to be slow. These lesions are efficient

Table 3 Endogenous DNA adducts detected in human tissues. (Adapted from Marnett and Burcham, 1993.)

Adduct	Tissue	Adduct levels (per 10^7 bases)	Method
7,8-Dihydro-8-oxo-2′-deoxyguanosine (8-oxo-dGuo)	Leukocytes	12 ± 7	HPLC-ECD
5-Hydroxy-2′-deoxycytidine (5-OH-dCyd)	Leukocytes	10 ± 5	HPLC-ECD
5-Hydroxy-2′-deoxyuridine (5-OH-dUrd)	Leukocytes	7 ± 6	HPLC-ECD
5,6-Dihydroxy-5,6-dihydro-2′-deoxyuridine (dUrdg)	Leukocytes	20 ± 15	HPLC-ECD
7,8-Dihydro-8-oxo-2′-deoxyadenosine (8-oxo-dAdo)	Leukocytes	230	GC-MS
5-(Hydroxymethyl)-2′-deoxyuridine (5-HmdUrd)	White blood cells	2300 ± 480	GC-MS
8-Hydroxy-6-methyl-1,N^2-propano-2′-deoxyguanosine (8-OH-6-Me-PdGuo)	Liver	6	^{32}P postlabelling
8-Hydroxy-1,N^2-propano-2′-deoxyguanosine (8-OH-PdGuo)	Liver	10	^{32}P postlabelling
1,N^6-Etheno-2′-deoxyadenosine (εdAdo)	Liver	0.7 ± 0.4	^{32}P postlabelling
3,N^4-Etheno-2′-deoxycytidine (εdCyd)	Liver	2.8 ± 0.9	^{32}P postlabelling
3-β-D-2′-Deoxyribofuranosylpyrimido [1,2-*a*]purin-10(3*H*)-one (M_1dGuo)	Liver, white blood cells, pancreas, breast	0.1–12	^{32}P postlabelling, MS, immunoassay

premutagenic lesions and, although they can also be formed by exogenous agents such as vinyl halides, they appear to have some credentials for consideration as endogenous origins of mutagenic and carcinogenic processes. Another promutagenic exocyclic base adduct that has been detected in human DNA is M_1dGuo (3-β-D-2′-deoxyribofuranosylpyrimido[1,2-*a*]purin-10(3*H*)-one), a product of the reaction of malondialdehyde (itself a product of lipid peroxidation and prostaglandin biosynthesis) with DNA (Marnett, 1999).

The use of ^{32}P postlabelling analysis has revealed the existence of a large variety of moderately polar lesions in DNA, termed I-compounds. The patterns observed are tissue- and species-specific and show, in experimental animals, significant increases with age. The levels of many of them are also dependent on the diets used to feed the animals. Paradoxically, their levels are actually lower in circumstances which result in tumour formation, such as feeding rats a choline-deficient diet or the administration of enzyme inducers, and they are lower in rat liver tumours than in normal liver. Furthermore, some I-compounds are subject to circadian rhythms.

There are many uncertainties about the role of endogenous DNA adducts in carcinogenesis. Some chemical carcinogens have been shown to cause oxidative damage to DNA as well as forming DNA adducts directly themselves. Other so-called nongenotoxic carcinogens, which do not directly damage DNA, are suspected of increasing levels of oxidative lesions in DNA, but it remains unclear whether this is a mechanism by which they induce tumour formation.

Several of the endogenous DNA adducts are found in human DNA at levels significantly higher than the levels of adducts from exogenous (environmental) carcinogens. Relatively little is known about the relationship between endogenous adduct levels and development of cancer. It might be expected that their levels would be increased in circumstances that ultimately lead to cancer. However, aside from evidence that oxidative DNA damage appears, in some instances, to be higher in tumour tissue than in normal adjacent tissue (e.g. in breast), other evidence for I-compounds suggests that their levels decrease with increasing cancer susceptibility. Until the nature of these compounds is more clearly understood, what their cellular function is remains a matter of speculation.

RISK ASSESSMENT

Elevated levels of adducts in human tissue are clearly biomarkers of exposure. Case-control studies, because they are retrospective, cannot establish causality (Perera, 2000). In order to determine the potential of DNA adducts as biomarkers of risk, it is necessary to examine their presence in subjects prior to the onset of disease, and then investigate at a later date whether those individuals in a cohort with higher adduct levels are the ones that subsequently develop tumours. To date only a few studies have done this. One study monitored aflatoxin B_1 exposure in Chinese men in Shanghai (Qian *et al.*, 1994); 18 244 volunteers each gave a single urine sample, which was stored for future analysis. Fifty-five men subsequently developed liver cancer and their urine samples were analysed along with matched control samples. Levels of aflatoxin adducts in the samples were significantly higher in the cases than in the controls, indicating that this parameter was indeed a biomarker of risk. Interestingly, the more classical methods of measuring exposure, based on estimating dietary consumption from food analysis and questionnaires, failed to identify the cases as being more exposed to aflatoxin B_1 than the controls. Thus the use of a DNA adduct biomarker in this instance revealed a link between exposure and risk of cancer that more classical epidemiological methods of exposure assessment failed to demonstrate.

In another such nested case-control study, the predictive value of DNA adducts in white blood cells for lung cancer risk was determined amongst male smokers. Those smokers who subsequently developed lung cancer had significantly higher levels of aromatic DNA adducts, determined by ^{32}P postlabelling, than those who did not (Tang *et al.*, 2000).

The limitations to the wider use of these methods for risk assessment are those which are generally encountered in prospective studies, namely that large numbers of subjects need to be recruited to the study and that it can take many years for the cases to appear and for the analyses on cases and matched controls to be conducted. In addition, the preservation and long-term storage of the biological samples of every individual recruited must be accomplished, adding considerably to the costs of the study.

In the field of genetic toxicology testing and regulatory affairs, the assessment of the potential carcinogenic and mutagenic properties of new compounds is of paramount importance. A number of *in vivo* and *in vitro* tests have been developed with different genetic endpoints, including bacterial mutagenicity, mammalian cell mutagenicity, micronucleus formation and aneuploidy (Phillips and Venitt, 1995), but no single test is reliably predictive (see also the chapter *Short Term Testing for Genotoxicity*). Many compounds are positive in some tests but negative in others, leaving uncertainties about the true characteristics of the compound. In these circumstances, investigating the compound for DNA adduct formation in suitable experimental systems may help to clarify the issue (Phillips *et al.*, 2000). A good example of such a compound, already mentioned, is tamoxifen, which is a potent liver carcinogen in rats but negative in most regulatory short-term tests for genotoxicity. Despite this, it gives rise to DNA adducts in the target tissue and is considered to be carcinogenic to rats by a genotoxic mechanism.

GENE–ENVIRONMENT INTERACTIONS

From early studies in which short-term explant cultures of human tissue or primary cultures of human cells were treated with carcinogens, it was evident that there was a wide range of DNA adduct formation with samples from different individuals. This variability is also observed in studies of DNA adducts formed *in vivo* in individuals apparently exposed to similar levels of carcinogens and points to the influence of genetic differences in carcinogen metabolism and/or DNA repair.

Detailed discussion of the potential role of polymorphisms in carcinogen-metabolizing enzymes in determining cancer susceptibility is beyond the scope of this chapter, but what can be commented on here are those studies where correlations have been sought between DNA adduct levels and genotype.

Polymorphisms in the *CYP1A1* gene have been extensively studied, although there is still disagreement as to whether these have functional consequences for the activity or inducibility of the enzyme. Two of these are the Msp1 polymorphism (loss/gain of a restriction site) and an exon 7 polymorphism that results in the coding for valine in place of isoleucine (Perera and Weinstein, 2000). One study has found that US smokers with the exon 7 variant allele had higher levels of DNA adducts in their white blood cells than smokers with the normal allele (Perera and Weinstein, 2000). Newborn babies with the *CYP1A1* Msp1 restriction site had higher levels of adducts in placenta and cord blood than those without it (Whyatt *et al*., 2000). Polymorphisms in other *CYP* genes may also be important but have been less well studied.

A number of polymorphisms in Phase II metabolizing enzymes are also of interest. The *null* genotype of *GSTM1* (in which glutathione *S*-transferase μ1 is absent) is associated with a greater risk of lung cancer, although, curiously, the gene does not appear to be expressed in the lung. Individuals with the *GSTM1 null* genotype have higher levels of lung DNA adducts (Perera and Weinstein, 2000). Polymorphisms in other members of the GST family, such as *GSTP1* and *GSTT1*, are also suspected of influencing carcinogen–DNA adduct levels in the lung. Combinations of polymorphisms in Phase I and II enzymes may additionally refine susceptibility to DNA adduct formation and cancer risk from carcinogen (e.g. tobacco smoke) exposure.

In studies of breast cancer risk, there is evidence that possession of the *GSTM1 null* genotype results in higher levels of PAH–DNA adducts in breast tumour tissue (Perera and Weinstein, 2000), and in another study possession of ‘slow’ alleles of *N*-acetyltransferase 2 (*NAT2*), which detoxifies carcinogenic aromatic amines, resulted in higher levels of aromatic/hydrophobic adducts (Pfau *et al*., 1998).

Currently there is much interest in how combinations of polymorphisms in different Phase I and Phase II metabolizing enzymes may, together, confer greater risk on certain individuals than others of DNA adduct formation and, ultimately, cancer risk. As yet, however, no correlations have been found that can account for the very wide range of interindividual variability in carcinogen–DNA adduct formation among human subjects.

CONCLUSION

The formation of DNA adducts by many carcinogens is causally associated with their mechanism of tumour initiation. Mutations in key genes as a consequence of adduct formation are found in many tumours and the altered proteins that they encode have functions that accord with the phenotypic differences between normal and malignant cells. While the formation of adducts is a necessary, but not sufficient, event for malignant transformation, enhancement of adduct formation will increase tumour formation, and inhibition of adduct formation will restrict it (see also the chapter *Mechanisms of Chemical Carcinogenesis*).

The detection and characterization of DNA adducts in mammalian tissues are research fields that are still undergoing rapid development, and in recent years it has become possible to detect adducts with high levels of sensitivity and/or selectivity. Among the many uses for DNA adduct determination, those currently of greatest interest include monitoring human exposure to environmental carcinogens, investigating the mechanism of activation and tumour initiation of carcinogens, monitoring DNA repair and investigating endogenous DNA damage and oxidative processes. DNA adduct detection has an important role to play in the burgeoning field of molecular epidemiology. It is also a supplementary procedure for assessing new compounds for genotoxic activity and can, potentially, provide valuable data for assessing patient response to cytotoxic chemotherapy.

REFERENCES

Beach, A. C. and Gupta, R. C. (1992). Human biomonitoring and the ^{32}P-postlabelling assay. *Carcinogenesis*, **13**, 1053–1074.

Brookes, P. and Lawley, P. D. (1964). Evidence for the binding of polynuclear aromatic hydrocarbons to the nucleic acids of mouse skin: relation between carcinogenic power of hydrocarbons and their binding to DNA. *Nature*, **202**, 781–784.

Cooper, C. S. and Grover, P. L. (eds) (1990). *Chemical Carcinogenesis and Mutagenesis.* Vols I and II. *Handbook of Experimental Pharmacology*, Vol. 94. (Springer, Berlin).

Davis, W., *et al*. (1998). The metabolic activation of tamoxifen and α-hydroxytamoxifen to DNA binding species in rat hepatocytes proceeds via sulphation. *Carcinogenesis*, **19**, 861–866.

Denissenko, M. F., *et al*. (1996). Preferential formation of benzo[*a*]pyrene adducts at lung cancer mutational hotspots in P53. *Science*, **274**, 430–432.

Dipple, A. (1994). Reactions of polycyclic aromatic hydrocarbons with DNA. In: Hemminki, K., Dipple, A., *et al*. (eds), *DNA Adducts: Identification and Biological Significance*. 107–129 (IARC, Lyon).

Dumaz, N., *et al*. (1993). Specific UV-induced mutation spectrum in the *p53* gene of skin tumors from DNA-repair-deficient xeroderma pigmentosum patients. *Proceedings of the National Academy of Sciences of the USA*, **90**, 10529–10533.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759–767.

Hainaut, P., *et al*. (1998). IARC Database of *p53* gene mutations in human tumors and cell lines: updated compilation, revised formats and new visualisation tools. *Nucleic Acids Research*, **26**, 205–213.

Halliwell, B. and Dizdaroglu, M. (1992). The measurement of oxidative damage to DNA by HPLC and GC/MS techniques. *Free Radical Research Communications*, **16**, 75–87.

Hemminki, K., Dipple, A., Shuker, D. E. G., Kadlubar, F. F., Segerback, D. and Bartsch, H. (eds) (1994). *DNA Adducts: Identification and Biological Significance.* (IARC, Lyon).

Jankowiak, R. and Small, G. J. (1991). Fluorescence line narrowing: a high-resolution window on DNA and protein damage from chemical carcinogens. *Chemical Research in Toxicology*, **4**, 256–269.

Lawley, P. D. (1984). Carcinogenesis by alkylating agents. In: Searle, C. E. (ed.), *Chemical Carcinogens*, 2nd edn. Vol. 1, 325–484 (American Chemical Society, Washington, DC).

Lawley, P. D. and Phillips, D. H. (1996). DNA adducts from chemotherapeutic agents. *Mutation Research*, **355**, 13–40.

Marnett, L. J. (1999). Lipid peroxidation–DNA damage by malondialdehyde. *Mutation Research*, **424**, 83–95.

Marnett, L. J. (2000). Oxyradicals and DNA damage. *Carcinogenesis*, **21**, 361–370.

Marnett, L. J. and Burcham, P. C. (1993). Endogenous DNA adducts: potential and paradox. *Chemical Research Toxicology*, **6**, 771–785.

Mollerup, S., *et al*. (1999). Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients. *Cancer Research*, **59**, 3317–3320.

Osborne, M. R. (1984). DNA interactions of reactive intermediates derived from carcinogens. In: Searle, C. E. (ed.), *Chemical Carcinogens*, 2nd edn. Vol. 1, 485–524 (American Chemical Society, Washington, DC).

Otteneder, M. and Lutz, W. K. (1999). Correlation of DNA adduct levels with tumor incidence: carcinogenic potency of DNA adducts. *Mutation Research*, **424**, 237–247.

Perera, F. P. (2000). Molecular epidemiology: on the path to prevention? *Journal of the National Cancer Institute*, **92**, 602–612.

Perera, F. P. and Weinstein, I. B. (2000). Molecular epidemiology: recent advances and future directions. *Carcinogenesis*, **21**, 517–524.

Pfau, W., *et al*. (1998). DNA adducts in human breast tissue: association with *N*-acetyltransferase-2 (NAT2) and NAT1 genotypes. *Cancer Epidemiology, Biomarkers and Prevention*, **7**, 1019–1025.

Phillips, D. H. (1983). Fifty years of benzo(*a*)pyrene. *Nature*, **303**, 468–472.

Phillips, D. H. (1990). Modern methods of DNA adduct determination. In: Cooper, C. S. and Grover, P. L. (eds), *Chemical Carcinogenesis and Mutagenesis*, Vol. 1. *Handbook of Experimental Pharmacology*, Vol. 94. 503–546 (Springer, Berlin).

Phillips, D. H. (1996). DNA adducts in human tissues: biomarkers of exposure to carcinogens in tobacco smoke. *Environmental Health Perspectives*, **104**, Supplement 3, 453–458.

Phillips, D. H. (1997). Detection of DNA modifications by the ^{32}P-postlabelling assay. *Mutation Research*, **378**, 1–12.

Phillips, D. H. and Farmer, P. B. (1995). Protein and DNA adducts as biomarkers of exposure to environmental mutagens. In: Phillips, D. H. and Venitt, S. (eds), *Environmental Mutagenesis.* 367–395 (Bios, Oxford).

Phillips, D. H., *et al*. (2000). Methods of DNA adduct determination and their application to testing compounds for genotoxicity. *Environmental Molecular Mutagenesis*, **35**, 222–233.

Phillips, D. H. and Venitt, S. (eds) (1995). *Environmental Mutagenesis*. (Bios, Oxford).

Poirier, M. C. and Beland, F. A. (1992). DNA adduct measurements and tumor incidence during chronic carcinogen exposure in animal models: implications for DNA adduct-based human cancer risk assessment. *Chemical Research in Toxicology*, **5**, 749–755.

Poirier, M. C., *et al*. (2000). Carcinogen macromolecular adducts and their measurement. *Carcinogenesis*, **21**, 353–359.

Qian, G. S., *et al*. (1994). A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China. *Cancer Epidemiology*, *Biomarkers and Prevention*, **3**, 3–10.

Shuker, D. E. G. and Farmer, P. B. (1992). Relevance of urinary DNA adducts as markers of carcinogen exposure. *Chemical Research in Toxicology*, **5**, 450–460.

Strickland, P. T., *et al*. (1993). Methodologies for measuring carcinogen adducts in humans. *Cancer Epidemiology, Biomarkers and Prevention*, **2**, 607–619.

Tang, D., *et al*. (2000). Aromatic DNA adducts as a predictor of lung cancer risk in a case-control study. *Proceedings of the American Association for Cancer Research*, **41**, 221.

van Steeg, H., *et al*. (1998). Use of DNA repair-deficient XPA transgenic mice in short-term carcinogenicity testing. *Toxicology and Pathology*, **26**, 742–749.

Weston, A. (1993). Physical methods for the detection of carcinogen–DNA adducts in humans. *Mutation Research*, **288**, 19–29.

Whyatt, R. M., *et al*. (2000). Association between polycyclic aromatic hydrocarbon–DNA adduct levels in maternal and newborn white blood cells and glutathione *S*-transferase P1 and CYP1A1 polymorphisms. *Cancer Epidemiology, Biomarkers and Prevention*, **9**, 207–212.

FURTHER READING

Cooper, C. S. and Grover, P. L. (eds) (1990). *Chemical Carcinogenesis and Mutagenesis*, Vols I and II. *Handbook of Experimental Pharmacology*, Vol. 94 (Springer, Berlin).

Groopman, J. D. and Skipper, P. L. (1991). *Molecular Dosimetry and Human Cancer: Analytical, Epidemiological and Social Considerations.* (CRC Press, Boca Raton, FL).

Hemminki, K., *et al.* (eds) (1994). *DNA Adducts: Identification and Biological Significance.* (IARC, Lyon).

Phillips, D. H. and Venitt, S. (eds) (1995). *Environmental Mutagenesis.* (Bios, Oxford).

Schulte, P. A. and Perera, F. P. (1993). *Molecular Epidemiology: Principles and Practice.* (Academic Press, San Diego).

Chapter 22

Physical Causes of Cancer

Jeffrey L. Schwartz
University of Washington, Seattle, WA, USA

CONTENTS

- Introduction
- Genotoxicity
- Carcinogenesis
- Risk Models and Protection Standards
- Summary and Conclusions
- Acknowledgements

INTRODUCTION

Unlike most chemical and viral carcinogens, physical carcinogens act by imparting energy into the biological material. If the energy imparted produces changes in the bonds holding molecules together, this will yield chemical changes and possibly biological effects. Radiation is the primary physical agent to which we are exposed in our environment. Radiation is a ubiquitous component of our environment. We are exposed to ultraviolet (UV) radiation from the sun, γ-rays from cosmic radiation and the decay of isotopes in building materials, air, water and food, and α-particles from radon and radon daughters that seep into our basements. We are exposed to X-rays and ultrasound from medical procedures, and both microwaves and radiofrequency (RF) radiation from various consumer products, including cellular telephones. Power generation leads to a host of different types of radiation exposure including electric and magnetic fields (EMFs). There is no way to avoid completely exposure to these potentially harmful levels of radiation. Understanding the risks associated with exposure can aid in developing appropriate protection standards and approaches to reducing risks.

Definitions

The amount of energy deposited in biological tissue, and therefore the types of changes seen in cells, depend on the nature of the radiation (**Table 1**). Ionizing radiation refers to those types of radiation that produce the ejection of an orbital electron from an atom or molecule and result in the formation of an ion pair. The ionization potential of most molecules in biological materials is 10–15 electronvolts (eV), so in order to be ionizing, the radiation must be able to impart at least that much energy. Ionizing radiation can be either electromagnetic (X-rays and γ-rays) or particulate (neutrons and α-particles). Energy loss varies with the energy of the incoming photon or particle, the charge of the particle and the character (atomic number, electron density) of the absorbing medium. The density of energy deposition along a track length has a profound influence on the subsequent biological effect. The spatial rate of energy loss along a track length is described by the term ‘linear energy transfer’ (LET). LET is defined as the energy lost (in kiloelectronvolts) per unit track length (in micrometres). X-rays and γ-rays are considered sparsely ionizing, low-LET radiation with ionization or ionization clusters being spaced relatively far apart. The LET for ^{60}Co γ-rays is 0.25 keV μm^{-1}, and that for 250-kV X-rays is about 3.0 keV μm^{-1}. Energetic particles tend to be more densely ionizing and have a higher LET than X-rays and γ-rays, although this is very dependent on the energy of the particle. The LET for a radon-derived α-particle is between 80 and 100 keV μm^{-1}. For energetic particles, the LET varies over the track length as the particle interacts and the energy spectrum changes. For high-energy particles, the density of ionization at the beginning of the track is fairly sparse and the LET is correspondingly low. As the particle loses energy, the density of ionization and the LET increases. At the end of the track, one may see a peak of ionization density (Bragg peak).

Table 1 Examples of ionizing and non-ionizing radiation

Ionizing radiation (> 10–15 eV)	Nonionizing radiation (<10 eV)
X-rays	UV
Gamma-rays	Microwaves
Alpha-particles	RF radiation
Neutrons	Ultrasound
	EMF

The major mode of energy loss for radiation having energy of 1–10 eV (nonionizing radiation) is excitation. Excitation refers to elevation of an electron to a higher ('excited') state. This state is transient, and when the electron returns to a ground state, the energy released can be in the form of visible light (fluorescence, phosphorescence) or chemical change (e.g. pyrimidine dimer formation). UV radiation is the major nonionizing radiation hazard in our environment. For UV radiation, both the wavelength and the fluence determine biological effects. UV radiation is subdivided into three wavelength bands, UVA (313–400 nm), UVB (290–315 nm) and UVC (220–290 nm). UVC is the most potent band for biological effects because DNA absorbs most strongly at the 254-nm wavelength. However, most UVC radiation is quickly absorbed in air. Thus UVB, which is also DNA damaging and potentially carcinogenic, is considered the greater environmental hazard.

For electromagnetic radiation that produces <1 eV, such as microwave and RF radiation, the energy deposited results primarily in molecular vibration and heat. These types of radiation are usually expressed by their frequency. Microwave radiation ranges from 300 MHz to 300 GHz, RF radiation from 300 Hz to 300 MHz and extremely low-frequency (ELF) radiation from 30 to 300 Hz.

Ultrasound consists of high-frequency acoustic waves too fast for us to hear. Human hearing cannot go beyond about 18 000 vibrations per second, or 18 kHz. The effects of ultrasound are usually classified into thermal, direct and cavitation effects. For nonthermal effects of ultrasound, cavitation is considered to be the most important. Under the right conditions, irradiation of a liquid with ultrasound leads to the formation and collapse of gas- and vapour-filled bubbles or cavities in the solution. The collapse of these bubbles can be violent enough to lead to chemical effects.

X- and γ-rays, UV radiation, microwaves and RF radiation are all forms of electromagnetic radiation. EMFs are not electromagnetic radiation, but mixtures of electric and magnetic fields. These fields emanate from electric power lines and all devices that use electricity. Electricity is usually delivered as alternating current. The resulting EMFs are of extremely low frequency and low energy. At the atomic level, weak electric and magnetic fields are too small to produce chemical changes by themselves. However, EMFs might act to modify biological processes by causing small changes in the frequency of events that trigger different signal transduction pathways. In this regard, all aspects of a field, including its frequency, amplitude and pattern, may be important.

Units of Dose and Activity

Radiation exposure is usually expressed either as energy incident on a surface or energy absorbed per gram of tissue. For ionizing radiation (**Table 2**), the gray (Gy) is the Système Internationale (SI) unit of dose that is most often used. It is equal to 1 J kg^{-1}. An older term still in use is the rad, which is equal to 0.01 Gy. For ionizing radiation, the unit of absorbed dose does not take into account the differences in efficiency with which one type of radiation might act. Therefore, to describe exposures to different types of radiation, a dose equivalent is used. The dose equivalent is calculated by multiplying the absorbed dose by a quality factor that takes into account the biological effectiveness of the radiation. The quality factor for ^{60}Co γ-rays is 1. For some energies of α-particles, the quality factor can be as high as 100. The original term used to compare radiation of different qualities was the rem (roentgen equivalent in man), which was equal to the dose in rad times a quality factor. The present SI term is the sievert (Sv), which is equal to the dose in gray times a quality factor.

Table 2 Ionizing radiation units

Type of unit	Unit	Definition
Dose	Gray (Gy)	1 J kg^{-1}
	Rad	1 rad = 0.01 Gy
Dose-equivalent	Sievert (Sv)	Dose × quality factor
	Rem	1 rem = 0.01 Sv
Activity	Bequerel (Bq)	1 dps[a]
	Curie (Ci)	1 Ci = 3.7×10^4 MBq

[a] Disintegration per second.

Radioactive isotopes decay, producing ionizing radiation at a rate specific for the type and concentration of the isotope. The intensity of the source (activity) is determined by the rate of nuclear transformations per unit time. The SI term used to describe activity is the becquerel (Bq), which is equal to one disintegration per second. The older term sometimes used is the curie (Ci), which was originally defined as the activity associated with 1 g of ^{222}Ra and later was defined as 3.7×10^{10} disintegrations per second; 1 Ci is equal to 3.7×10^4 MBq. Activity is a rate measure. It is not equivalent to dose and, without knowledge of the nature of the decay produced, it is not possible to assign dose or biological effect solely on the basis of activity measures.

Exposure to UV radiation is expressed in joules per square centimetre (J cm^{-2}) . Dose for UV raditation, is in joules. The UV index is another term used to describe UV dose rate. The UV index is a forecast of the amount of skin damaging UV radiation expected to reach the Earth's surface at the time when the sun is highest in the sky. The UV index ranges from 0 at night time to 15–16 in the tropics at high elevations under clear skies. The higher the UV index, the greater is the dose rate of skin-damaging UV radiation.

Incident microwave or RF energy is expressed as energy flux or power density in watts per square centimetre (W cm^{-2}). The specific absorption rate (in watts per

kilogram) is the measure of energy absorbed. Ultrasound exposure is similarly expressed in watts per square centimetre and also by their frequency in hertz. EMFs are described by frequency, amplitude and pattern. Magnetic field strength, a key component of EMF, is expressed in tesla or gauss ($1\,T = 10^4\,G$).

Sources of Exposure

Ionizing radiation exposure varies widely for different populations. In part, exposure levels depend on altitude and latitude. The atmosphere attenuates dose from extraterrestrial sources of radiation; thus higher elevations receive greater doses of γ-rays from cosmic radiation. In the USA, the highest background exposures in a heavily populated area occur in Denver, which experiences about 0.25 mSv per year more than the US average. Higher exposure levels also accompany high-altitude flight. For flights at around 40 000 feet, the average exposure is between 0.005 and 0.01 $mSv\,h^{-1}$.

The Earth's magnetic field acts to deflect the protons from cosmic radiation towards the poles. Thus the polar regions tend to receive greater doses of ionizing radiation. There are also areas in the world where the natural background is higher than the average. These areas include the monozite regions of India and Brazil. Background radiation levels in these areas may exceed 10 times the world average.

The average yearly dose of ionizing radiation in the USA is 3.6 mSv (360 mrem). To put this number in perspective, for a typical dental X-ray, the equivalent dose is 0.01–0.02 mSv and a chest X-ray is about 0.1 mSv. Cosmic rays and γ-rays from naturally occurring isotopes account for about 1 mSv per year (**Table 3**). Radon accounts for about 2 mSv per year. That is not to imply that we are exposed to large levels of radon. Instead, the α-radiation associated with radon decay has a high quality factor of around 20, hence the dose equivalent is high. As for anthropogenic radiation exposure, medical exposures account for about 0.5 mSv per year, while exposures from consumer products make up about 0.1 mSv per year. There are also exposures associated with a variety of other activities, including the use of coal-fired power plants, tobacco products and, as mentioned above, travel at high altitudes.

Table 3 Sources of ionizing radiation exposure (in the USA)

Source	Amount (mSv per year)
Natural	3.0
Radon	2.0
Cosmic	0.3
Terrestrial	0.3
Internal	0.4
Artificial	0.6
Medical	0.5
Consumer	0.1
Other	
Work, nuclear power generation, fallout	<0.01
Total	3.6

There are substantial daily and seasonal variations in UV exposure levels. The atmosphere also acts to attenuate dose of UV radiation from the Sun. Most UV radiation is absorbed by the ozone layer or reflected back into space. Atmospheric ozone, which is principally located in the stratosphere, is a strong absorber of UVB radiation. Ozone is formed when molecular oxygen is split by UV radiation. The singlet oxygen atom that results can combine with other molecules of molecular oxygen to form ozone. Ozone production is normally balanced by its destruction through similar photochemical events. In addition, ozone is depleted by the chlorine contained in chlorofluorocarbons (CFCs), a family of commonly used industrial compounds. In part due to these CFCs, the amount of ozone screening the Earth has decreased by an average of 3% per year over the last decade, and a polar ozone hole over the South Polar region has developed. The reduction in ozone has resulted in a corresponding increase in UV levels.

There are a myriad of consumer products that generate radiation. Microwave ovens and cellular telephones provide a common source of microwave and RF radiation. There are a number of medical and consumer devices that use ultrasound. All electrical devices also generate EMFs. Transmission power lines generate both strong electric fields and strong magnetic fields. Distribution lines generate weak electric fields but can generate strong magnetic fields. Exposure levels to EMFs around the home are in the range 0.1–2.5 mG. For homes near power lines, these levels may be as high as 5–10 mG. Immediately under the power line, magnetic field levels of 60–100 mG may be found. The Earth's magnetic field has a strength of about 500 mG, although, unlike EMFs, it is not changing directions 50–60 times per second and is therefore not comparable to electrically generated EMFs. EMFs also show intermittent spikes in frequency (transients) and harmonics, which are multiples of the standard frequency.

GENOTOXICITY

Cancer is a genetic disease and therefore, at some point, carcinogen exposure must result in genetic alterations (see also the chapter *Mechanisms of Chemical Carcinogenesis*). These can be produced directly through an interaction of the radiation with DNA, indirectly, through some intermediate molecule that in turn damages DNA, or radiation may produce epigenetic alterations in cell growth or metabolism that ultimately lead to the transformed phenotype.

Ionizing Radiation

Ionizing radiation induces a variety of different types of DNA alterations, including nucleotide-base alterations and breaks in the sugar–phosphate backbone (Schwartz, 1995). Normally, cells handle these base alterations and single-strand breaks very well using base excision repair processes. Base excision repair is an error-free process that removes damaged bases and then takes advantage of an intact complementary strand to fill in missing information. Defects in excision repair can lead to mutagen sensitivity and cancer susceptibility (see also the section *The Molecular Basis of Cell and Tissue Organisation*).

For X-rays, about one in 10–20 breaks span both strands of the DNA molecule. This ratio of double-to single-strand breaks increases with increasing LET. DNA double-strand breaks can form directly as a result of radiation exposure, or can develop during excision repair due to the formation of overlapping gaps. DNA double-strand breaks are repaired by homologous and nonhomologous recombination processes. These are inherently error-prone and can lead to large deletions and rearrangements. DNA double-strand breaks have been shown to be the primary lethal and mutagenic lesion induced by ionizing radiation. (See also the chapter *Genomic Instability and DNA Repair*.)

Most of the genetic alterations seen in irradiated cells are large deletions and rearrangements. This is mostly a reflection of the nonhomologous repair of DNA double-strand breaks. Because radiation involves the loss of so much genetic material, it is a relatively weak mutagen. Most of the genetic changes lead to cytotoxicity. The nature of the alterations induced by radiation is very dependent on gene locus. Some loci can tolerate the loss of large amounts of genetic material. These loci would tend to be more sensitive to mutation induction and would show primarily large deletion. Other loci, because of their proximity to essential genes, are less tolerant of large deletion formation. These loci will be less sensitive to mutation induction and show far smaller types of genetic changes.

The dose response for mutation induction by X-rays and γ-rays has a quadratic portion to it, suggesting that most mutations develop from the interaction of two breaks (**Figure 1**). Fractionating the dose or reducing the dose rate will in general lead to a reduction in the effectiveness of the radiation by allowing time for repair of these sublethal lesions. As LET increases, both the frequency and the shape of the dose response change. High-LET radiation such as α-radiation is more effective at inducing mutations and their dose response is linear, suggesting that a single α-track can produce enough damage for lesion interaction. Fractionation of high-LET radiation usually has little effect on radiation-induced genotoxicity. The lack of sparing at low dose rates is presumed to be due to the inability of cells to repair damage induced by high-LET radiation.

Recent studies have demonstrated that ionizing radiation can induce genetic instability in exposed cells.

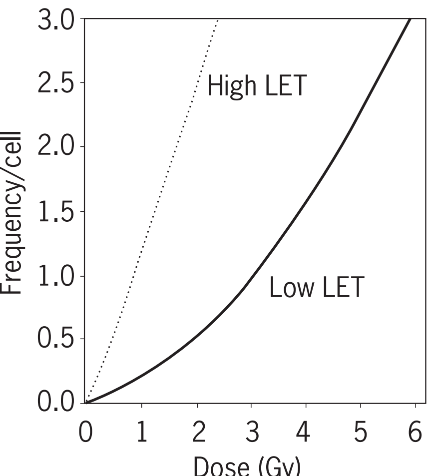


Figure 1 Examples of low- versus high-LET dose responses.

Delayed appearance of elevated frequencies of gene and chromosome mutations has been reported for cells exposed both *in vitro* and *in vivo*. The target for the effect is unknown. Instability is a frequent byproduct of radiation exposure. Rates as high 3% Gy^{-1} have been reported. (See also the chapter *Genomic Instability and DNA Repair*.)

Exposure to ionizing radiation can also induce alterations in gene expression as part of a generalized stress response. This stress response can affect DNA repair, cell toxicity and cell growth. As such, they may lead to more permanent genetic effects. While DNA damage can initiate these stress responses, other targets of radiation have also been identified. Alterations in redox levels or energy levels or specific molecular alterations in key protein molecules may trigger these responses. For example, direct activation of TGF-β by ionizing radiation can influence cell growth and apoptosis.

UV Radiation

The excitations induced by UV exposure will also produce DNA base changes and breaks in the sugar–phosphate backbone. The most important alterations appear to be cyclobutanepyrimidine dimers and pyrimidine-(6,4)-pyrimidone photoproducts. Replication of DNA containing these lesions leads to C to T transitions at dipyrimidine sites. Nucleotide excision repair processes normally repair these base alterations efficiently and without error. Like ionizing radiation, UV radiation can also induce alterations in gene expression as part of a generalized stress response. (See also the chapter *Genomic Instability and DNA Repair*.)

RF and Microwave Radiation

RF and microwave radiation do not produce enough energy to induce chemical change. They can lead to heat, however, and excessive heat may lead to more permanent

genotoxic changes (Moulder *et al*., 1999). Heat can also lead to the induction of a stress response, the spectrum of which overlaps with that seen with ionizing and UV radiation. It remains to be shown whether subthermal levels of microwave or RF exposure such as most of us experience can lead to permanent genetic effects. Most studies report no evidence that prolonged exposures to subthermal levels of microwave or RF radiation lead to genetic or epigenetic changes. The few positive studies suggest that the effects, if any, are small.

Ultrasound

Cavitation is considered to be the primary nonthermal effect of ultrasound that might lead to biological effects (Miller, 1987). Cavitation can lead to cell membrane damage by mechanical forces produced by cavitation. Ultimately, this can cause cell lysis. Cavitation can also lead to the formation of free radicals that can further damage cells. Alterations in cell morphology, membrane transport and cell growth following ultrasound exposure have all been reported, but the effects are small and transient. Ultrasound-induced free radicals have the potential to damage DNA. However, most reports on DNA damage or mutation following ultrasound exposure are negative. The few positive studies that demonstrate DNA strand breakage and gene mutations are for very high intensity exposures. Almost all of these studies are *in vitro* studies. It is less clear whether similar effects will be seen *in vivo*.

EMF

In vitro studies with EMF exposure have focused on endpoints commonly associated with cancer, including cell proliferation, signal transduction alterations and differentiation inhibition as well as on more traditional DNA damage endpoints (Moulder *et al*., 1999). Effects on gene expression, cell growth and signal transduction were reported for magnetic flux densities $>100\,\mu T$ or internal electric field strengths $>1\,mV\,m^{-1}$, but no consistent alterations have been reported for lower intensities. There are no reported direct effects on DNA. Disruption of normal circadian rhythm of melatonin was one postulated mechanism of EMF action. Studies on EMF effects suggest some effect on melatonin, but the significance of this observation for carcinogenesis is not obvious.

CARCINOGENESIS

It has clearly been established that ionizing and UV radiation are both carcinogens. In contrast, the epidemiological evidence for an association between cancer and microwave and RF radiation, ultrasound or EMF exposure is weak and inconsistent.

Ionizing Radiation

Information on cancer induction by human populations comes from epidemiological studies of exposed human populations (National Research Council, 1990, 1999; Schwartz, 1995). For ionizing radiation, the largest single group of exposed individuals are those Japanese exposed to the atomic bombs at Hiroshima and Nagasaki. About 280 000 individuals survived the immediate effects and about 80 000 have been followed for long-term effects. Other exposed populations include early radiation workers such as radium dial painters, uranium miners and populations exposed for therapeutic reasons. These and other smaller groups of exposed individuals serve as the human database for estimating risk for developing cancer following ionizing radiation exposure. (See also the chapter *Identifying Cancer Causes through Epidemiology*.)

At the molecular, cell and tissue levels, cancers induced by ionizing radiation are so far indistinguishable from those that occur spontaneously. Unlike with many chemical mutagens and UV radiation, there are no specific mutations associated with radiation exposure that are also seen in radiation-induced tumours. Recent studies suggest that the initiating event in radiation carcinogenesis may be the induction of genetic instability. (See also the chapter *Molecular Epidemiology of Cancer*.)

The types of tumours seen following radiation exposure are usually the same as seen spontaneously. The effect of radiation on cancer induction is usually inferred from the increase in frequency over background. Radiation-induced tumours appear in almost all tissues of the body, but sensitivity varies greatly for specific tissues and organs (**Table 4**). In general, the thyroid, female breast and certain blood forming organs are considered the most sensitive to induction by ionizing radiation in humans, while kidney, bone, skin, brain and salivary gland are considered the least sensitive. The lymphatic system, lung, colon, liver, and pancreas are among those tissues with moderate sensitivity. The differential sensitivity probably reflects a complex number of factors; it is not simply a reflection of spontaneous frequencies.

The types of cancers observed are also related to the nature of exposure. Many different types of cancer have been seen in the atomic bomb-exposed individuals who in general experienced total body irradiation. Early radiation workers exposed occupationally showed increases in skin

Table 4 Relative ionizing radiation sensitivity of tissues

Sensitive	Moderate	Resistant
Thyroid	Lung	Kidney
Breast	Colon	Bone
Blood-forming tissues	Liver	Skin
	Pancreas	Salivary gland
	Lymphatic system	Brain

cancer and leukaemias, reflecting the nature of their exposure. Many worked with radium, a bone-seeking isotope. Others studied relatively low-energy X-rays where skin would be the primary tissue exposed. Radium dial painters were a group (mostly women) who painted watch dials with a radioactive solution. In the course of their work, they ingested radium. The radium deposited in bone, resulting primarily in higher incidences of bone cancers. Uranium mineworkers inhaled radon gas. The α-radiation exposure resulting from that inhaled radon led to a higher incidence of lung cancer. Thorotrast was a contrast agent used in the late 1920s and early 1930s. It is a colloidal preparation of thorium-232 dioxide that tends to concentrate in the liver. Thorium is an α-emitter. Excess liver cancers and leukaemias were seen in these patients. During the late 1930s, patients in the UK with ankylosing spondylitis were treated with radiation to reduce bone pain. An excess of leukaemias has been reported in these individuals. In the past, radiation has been used to reduce thymus size in children and also for the treatment of tinea capitus. The critical organ exposed was the thyroid. These exposed children later showed excesses in the incidence of thyroid cancers.

The latent period for cancer induction following radiation varies with tumour type, radiation type, dose and dose rate. Leukaemias have the shortest latent periods (mean, 5–10 years), which no doubt accounts for their being seen with most radiation exposures. Solid tumours show latencies of between 20 and 30 years.

In general, the dose response for the induction of tumours by radiation follows a sigmoidal response. At low doses, there is little induction. A steep increase and then saturation or even decrease in tumour frequency at high doses follow. The dose response will vary depending on tissue type, dose rate and tumour latency time. As mentioned above, most of the available data on human carcinogenesis come from individuals exposed to relatively high doses of radiation. No human data exist for proven carcinogenic effects of radiation below 0.1 Sv. The dose response at low levels of exposure remains unknown. One predicts response in this region by extrapolation. Animal studies have suggested that the shape of this crucial portion of the curve may vary greatly from tissue to tissue and animal to animal. At present, linear extrapolation of high-dose effects is used to estimate risk for low doses of radiation. This is considered the most conservative approach. However, there is growing evidence for non-linearity at low doses (Pollycove, 1998). Some investigators have suggested that cells can adapt to low dose exposure and thus linear extrapolation overestimates risk at low doses.

As the LET of the radiation increases, the carcinogenic effects seen also increase in severity up to about $100\,\text{keV}\,\mu\text{m}^{-1}$ (**Figure 2**). As LET increases beyond this point, the effect on carcinogenesis usually declines. This reflects 'overkill,' where individual ionization events are not less effective, but the increasing amounts of the energy released per ionization event are wasted.

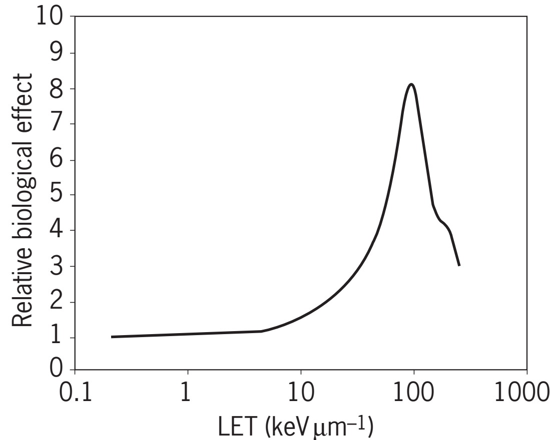


Figure 2 Variation of biological effect with LET. RBE is defined as the ratio of doses (standard: experimental) that yield the same biological effect.

As with other biological effects, protracting the dose over long periods of time reduces cancer incidence, and the effectiveness of fractionated or low-dose-rate irradiation in causing tumours is different for low- and high-LET radiation. Most studies on fractionated low-LET radiation have resulted in a reduction of tumour incidence for a given total dose. Presumably, fractionation of the dose allows for time to repair sublethal and subcarcinogenic damage. In contrast, fractionation of high-LET radiation or low-dose exposure usually has little effect on radiation-induced tumour formation. For some types of radiation, such as fission-spectrum neutrons, reducing the dose rate actually leads to increased transformation and more tumours. This is known as the inverse dose-rate effect. The largest inverse dose-rate effects are seen for fission-spectrum neutrons, with monoenergetic neutrons yielding reduced enhancements, and charged particles having LETs $>120\,\text{keV}\,\mu\text{m}^{-1}$ producing little or no enhancement. The inverse dose-rate effect is most prominent at low doses (<20 cGy) and low dose rates ($<0.5\,\text{cGy}\,\text{min}^{-1}$).

Sensitivity to tumour induction varies for different species and strains of animals, is different for males and females and also shows interindividual variability (Schwartz, 1995). This variability suggests that the initial damage, which is presumed to be the same for a given dose and type of radiation, is subject to a number of host factors that modify response. Presumably these include repair capability, presence of endogenous viruses, cell proliferation status, endocrine levels, immune competence, age of irradiation and factors associated with genetic susceptibility. There are a number of genetic syndromes that show radiation sensitivity and cancer susceptibility (Murnane and Kapp, 1993).

Radiation risk is defined as the increase in the number of cancer deaths over that expected for an unirradiated population. It is expressed in units per person exposed

per gray of radiation. Estimates based on linear extrapolation of the atomic-bomb data and on other more limited data from pooled results of various partial body exposures give total cancer mortality risks for a general population exposed to whole body radiation of $(1–4) \times 10^{-2}$ per person-Gy.

UV Radiation

There is extensive epidemiological evidence supporting the direct role that sunlight plays in human skin cancer (van der Leun, 1984; de Gruijl, 1999; Green *et al*., 1999). (See also the section *Systemic Oncology*.) Patients who develop skin cancer generally have decreased melanin pigmentation. Melanin normally acts to protect skin from UV radiation by absorbing it. People with light complexions and who sunburn easily have a higher incidence of tumours. Basal cell carcinomas, the most common skin cancers in Caucasians, are found primarily on sun-exposed areas such as the head and neck where a dose–response relationship exists. There is even stronger evidence for the role of sunlight in causing squamous cell carcinomas of the skin. Although both types of tumours are more prevalent in geographical areas of high sun light exposure, there is a much greater increase in squamous cell carcinomas with decreasing latitude and increasing sunlight exposure. A reasonable correlation also exists between sunlight exposure and melanoma, but the relationship is not as clear as with basal and squamous cell carcinomas. Unlike basal and squamous cell carcinomas, melanomas occur most frequently on the upper back in males and lower extremities in females. Melanoma incidence does not follow a pattern of increased risk with cumulative UV exposure, whereas the incidences of basal and squamous cell carcinomas do increase with cumulative exposure. The risk of skin cancer is highly dependent on UV wavelength. (See also the chapter *Identifying Cancer Causes through Epidemiology*.)

Microwave and RF Radiation

Studies on microwave and RF radiation and cancer have included analyses of a wide variety of different populations, including radar laboratory workers, foreign service workers, military personnel and electrical workers (National Research Council, 1997; Moulder *et al*., 1999; National Institute of Environmental Health Sciences, 1999). In addition, there have been a number of animal studies that have looked at the effects of exposure. For many of the epidemiological studies, there is no precise information on dose. Often occupation is used to define exposure groups. In general these studies either find no effect of nonthermal levels of microwave or RF radiation on tumour induction, or show weak and inconsistent results. There is no evidence for any single type of tumour being induced by exposure, and no strong evidence for any dose–response relationship.

There have also been a number of long-term exposure studies with mice and rats. Too many of the animal studies suffer from poorly controlled exposures where heat stress is a component of exposure. Still, taken together, these studies do not support an effect of nonthermal levels of RF exposure on cancer induction. There is also mixed evidence for a tumour-promoting effect of RF exposure. Some groups have reported that RF exposure shortens tumour latency time and increases tumour frequency in carcinogen-exposed mice and in lymphoma-prone transgenic mice. Others see no effect on the promotion of spontaneous or chemically induced tumours. Hence it is not possible to conclude that nonthermal exposures to microwave or RF radiation have any effect on cancer induction or progression. (See also the chapter *Non-Genotoxic Causes of Cancer*.)

Ultrasound

There is no reported evidence for cancer induction by ultrasonic exposure. There are some suggestions that ultrasound might encourage neoplastic growth and promote metastases, but no strong evidence for either effect (Miller, 1987). (See also the chapter *Non-Genotoxic Causes of Cancer*.)

EMFs

The evidence for carcinogenic effects of EMFs is weak. The strongest evidence comes from epidemiological studies that observe associations between EMF exposure and leukaemia. The initial study by Wertheimer and Leeper (1979) suggested a causal association between risk of childhood leukaemia and exposure to magnetic fields. Wire code classifications were used to estimate exposure. Subsequent studies on this association have come to mixed conclusions. In general there appears to be at most a small increased risk of childhood leukaemia associated with EMF exposure (National Research Council, 1997; Moulder *et al*., 1999; National Institute of Environmental Health Sciences, 1999). There were also initial suggestions of an association between EMF exposure and brain cancers and lymphomas in children, but subsequent studies have not borne out this risk. Epidemiological reports of adult cancer induction by EMF were based primarily on occupational exposures. As with childhood leukaemia, there were mixed results for an association between EMF exposure and chronic lymphocytic leukaemia, suggesting at most a weak risk. The evidence for increased risk based on residential studies of adults is even weaker. (See also the chapter *Non-Genotoxic Causes of Cancer*.)

There have been numerous animal carcinogenicity studies of EMFs. The animal studies are all negative. There are no data to support any association between leukaemia and EMF exposure. Similarly, no evidence for any significantly increased frequency of tumours or changes in tumour latency or size was observed. The lack of any experimental data to back up the epidemiological data suggests that

there might be other factors that explain the increased risk of cancer. However, none has yet been identified.

RISK MODELS AND PROTECTION STANDARDS

There are a number of ways to model risk. Absolute risk refers to the number of cancers induced over spontaneous levels. Relative risk is a multiplicative increase over spontaneous. As cancer latency periods are generally long, time-dependent relative risk models are also used. With ionizing radiation, there are a number of other dependent variables that are considered in developing risk models. These include dose and $(dose)^2$, age at exposure, time since exposure and gender.

Risk estimate development for ionizing radiation is handled by the BEIR (Biological Effects of Ionizing Radiation) Committee in the USA and by UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation) for the United Nations. The NCRP (National Council for Radiation Protection) in the USA and the ICRP (International Commission of Radiation Protection) for the United Nations use BEIR/UNSCEAR information to develop appropriate radiation protection standards. Total cancer mortality risks for a general population exposed to whole-body radiation is currently based on linear extrapolation of the atomic-bomb data and on other more limited data from pooled results of various partial body exposures. Risks are 5×10^{-2} per person-Sv for low dose and low dose rates and 10×10^{-2} per person-Sv for high dose and high dose rates. The values for the working (adult) population are about 80% of those for the general (adult and child) population. Exposures to the general public are limited to 1 mSv per year while occupational standards are 50 mSv per year. Embryo exposure is limited to 0.5 mSv per month.

There are no corresponding limits for UV or ultrasound exposures, although protective measures are usually required for working with UV- or ultrasonic-producing equipment. For microwave radiation, the recommended exposure limit is $10\,mW\,cm^{-2}$. It is based on thermal effects of microwaves. RF limits are one-tenth of the microwave standard. The International Commission on Non-Ionizing Radiation Protection has set up guidelines for EMF exposure. Magnetic field exposures are limited to 1 G for the general public (10 G for short-term exposure) and 5 G (50 G for short-term exposure) for the occupationally exposed.

SUMMARY AND CONCLUSIONS

As mentioned at the beginning of this chapter, radiation is a ubiquitous component of our environment. There is no way to avoid exposure to radiation. Furthermore, as our technology advances, our exposure to various forms of radiation increases in both amount and complexity.

It has clearly been established that ionizing and UV radiation are both carcinogens. They represent the primary physical carcinogens in our environment and most efforts at reducing cancer risks are appropriately focused on these agents. There remain questions as to mechanisms of carcinogenesis for ionizing radiation, and in particular the effects of low-level exposures. Most ongoing studies in this area are attempting to address these questions. The answers may have a profound effect on radiation protection standards.

In contrast to ionizing and UV radiation, the evidence for an association between cancer and microwave and RF radiation, ultrasound or EMF exposure is weak and inconsistent. The levels of exposure to these types of radiation are increasing, as are the numbers of individuals exposed. There is also some uncertainty as to potential interactions between different types of radiation. Hence there are likely to be continuing investigations into the effects of these different types of radiation exposures.

ACKNOWLEDGEMENTS

This work was supported by the National Cancer Institute (NCI) and the National Aeronautics and Space Administration (NASA) through NIH grant CA-73931 and by a grant from the Department of Energy (DE-FC03-00ER62908).

REFERENCES

de Gruijl, F. R. (1999). Skin cancer and solar UV radiation. *European Journal of Cancer*, **35**, 2003–2009.

Green, A., *et al*. (1999). Sun exposure, skin cancers and related skin conditions. *Journal of Epidemiology*, **9**, S7–S13.

Miller, D. L. (1987). A review of the ultrasonic bioeffects of microsonation, gas-body activation, and related cavitation-like phenomena. *Ultrasound in Medicine and Biology*, **13**, 443–470.

Moulder, J. E., *et al*. (1999). Cell phones and cancer: what is the evidence for a connection? *Radiation Research*, **151**, 513–531.

Murnane, J. P. and Kapp, L. N. (1993). A critical look at the association of human genetic syndromes with sensitivity to ionizing radiation. *Seminars in Cancer and Biology*, **4**, 93–104.

National Institute of Environmental Health Sciences (1999). *Health Effects from Exposure to Power-line Frequency Electric and Magnetic Fields*. NIH Publication No. 99-4493. (NIEHS, Washington, DC).

National Research Council (1990). Committee on the Biological Effects of Ionizing Radiation (BEIR V). *Possible Health*

Effects of Exposure to Low Levels of Ionizing Radiation: BEIR V. (National Academy Press, Washington, DC).

National Research Council (1997). Committee on the Possible Effects of Electromagnetic Fields on Biologic Systems. *Effects of Exposure to Residential Electric and Magnetic Fields.* (National Academy Press, Washington, DC).

National Research Council (1999). Committee on Health Risks of Exposure to Radon (BEIR VI). *Health Effects of Exposure to Radon: BEIR VI*. (National Academy Press, Washington, DC).

Pollycove, M. (1998). Nonlinearity of radiation health effects. *Environmental Health Perspectives*, **106**, Supplement 1, 363–368.

Schwartz, J. L. (1995). Radiation carcinogenesis and the development of radiation injury. In: Arcos, J. C. and Arcos, M. F. (eds). *Chemical Carcinogenesis: Modulation and Combination Effects*. 473–508 (Bikrkhäuser, Springer International, Boston).

van der Leun, J. C. (1984). UV-carcinogenesis. *Photochemistry Photobiology*, **39**, 861–868.

Wertheimer, N. and Leeper, E. (1979). Electrical wiring configurations and childhood cancer. *American Journal of Epidemiology*, **109**, 273–284.

FURTHER READING

Hall, E. J. (2000). *Radiobiology for the Radiobiologist*, 5th edn. (Lippinicott, Williams & Wilkins, Philadelphia).

Websites

www.epa.gov
www.lowdose.org
www.osha.gov

Chapter 23

Non-genotoxic Causes of Cancer

Susan K. Murphy and Randy L. Jirtle
Duke University Medical Center, Durham, NC, USA

CONTENTS

- Introduction
- Epigenetic Characteristics of DNA
- Molecular Modulators of Epigenetic Signals
- Epigenetic Characteristics of Tumour Cells
- Imprinted Genes as Cancer Susceptibility Loci
- Non-genotoxic Mediators of Carcinogenesis
- Diagnostic and Therapeutic Potential of Epigenetic Abnormalities
- Conclusion
- Acknowledgements

INTRODUCTION

Genotoxic agents are characterized by their ability to induce DNA changes that alter the coding information within the primary nucleotide sequence either directly or through a reactive metabolite. These nucleotide mutations are widely accepted as the molecular events responsible for inactivating both alleles of a tumour-suppressor gene according to the Knudson 'two-hit model' of carcinogenesis (Knudson, 1971). However, agents that induce epigenetic alterations to the genome, or DNA structural changes in the absence of nucleotide sequence alterations, have increasingly been demonstrated to also play a fundamental role in cancer formation. These non-genotoxic agents induce heritable changes in the DNA that can disrupt gene regulatory regions. Epigenetic changes are potentially even more potent than genetic mutations in causing cancer since the resulting chromatin structural alterations can exert regional influences, thereby disrupting the normal transcriptional activity of multiple genes. This chapter will focus on the mechanisms of chromatin structure modulation and the potential role of epigenetics in the formation and treatment of cancer.

EPIGENETIC CHARACTERISTICS OF DNA

The double helical structure of DNA is remarkable in its simplicity, yet correct gene function requires not only that the base sequence is faithfully transcribed, but also that expression is both spatially and temporally regulated in a tightly controlled manner. The process of development in multicellular organisms depends on the differential repression or activation of particular genes in a cell type-specific manner, and this programming information must be maintained throughout the life of the individual. This is referred to as 'cellular memory,' and is controlled by epigenetic mechanisms (Riggs and Porter, 1996). Epigenetic regulation is therefore fundamentally important to the control of gene expression. Two major classes of epigenetic modification are instrumental in determining this complex level of gene regulation: histone acetylation and cytosine methylation.

Chromatin Structure

The human genome consists of approximately three billion base pairs of DNA divided amongst 22 pairs of autosomes and two sex chromosomes. If left in its native form, the genome would be several metres in length. The problem of packaging and organizing the DNA within the nucleus is overcome by the coordinated compaction of the genome by specialized DNA-binding proteins, including the histone proteins. 'Chromatin' refers to the combination of DNA together with these proteins and is organized into two major types of subchromosomal domains. Heterochromatin is tightly compacted and less transcriptionally active while euchromatin is less compacted and more likely to be transcribed. The chromatin structure also varies with the phases of the cell cycle. It is relaxed and transcriptionally active during interphase whereas it is condensed and inactive during mitosis. (See chapter on *Regulation of the Cell Cycle*.)

Histones are a family of proteins that provide the scaffolding for chromatin assembly and, consistent with this function, are among the most highly conserved proteins throughout evolution. The histones share the same basic structure consisting of a globular head and a positively charged, nonglobular tail. Histones H2A, H2B, H3 and H4

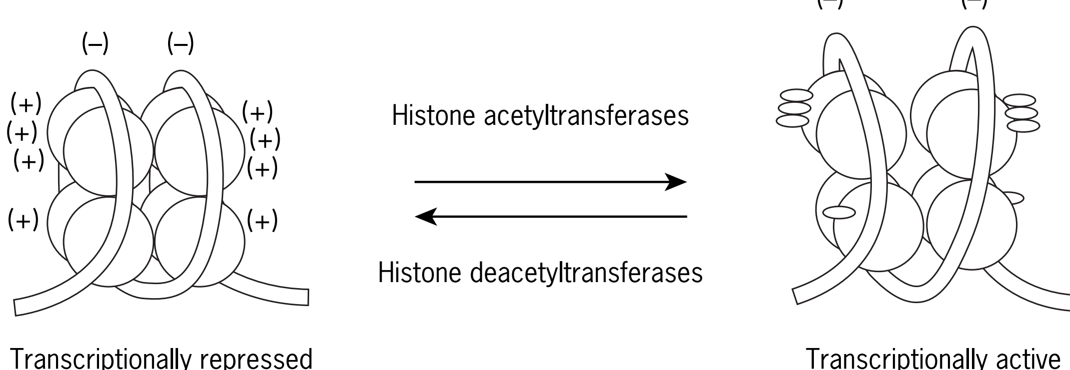


Figure 1 Histone acetylation and gene transcription. The nucleosomal core consists of eight histone proteins: two copies each of histones H2A, H2B, H3 and H4 (spheres). Approximately 150 base pairs of DNA are wrapped around each histone octamer. In transcriptionally inactive chromatin, the histones lack acetyl groups and are tightly compacted with the DNA. In transcriptionally active chromatin, histones H3 and H4 are acetylated (ovals) on their N-terminal tails. This post-translational modification neutralizes the inherent positive charge of the histone tails that in turn is thought to decrease the affinity of the histones for the negatively- charged DNA phosphate backbone. This contributes to an open chromatin structure in which the DNA is more loosely wrapped around the histones, and is receptive to interaction with transcription factors.

together form a heterooctamer core around which approximately 150 base pairs of DNA are wrapped to form discrete units called nucleosomes (**Figure 1**). Internucleosomal segments are bound by histone H1 and link each nucleosome to the next. The nucleosomes are further bundled into higher order structures to form compacted and organized chromatin. Histone–DNA interactions are modulated in part by reversible acetylation of the ε-amino groups of the histones lysine side chains. The lysine residues carry positive charges thought to contribute to the histone's affinity for negatively charged DNA. Acetylation of these lysine side chains neutralizes the positive charge thereby decreasing the ability of histones to interact with DNA, resulting in a more open chromatin configuration and increased transcription (**Figure 1**). Acetylation of histones by histone acetyltransferase (HAT) is reversed by the action of histone deacetylase (HDAC). The latter enzyme reduces transcriptional activity by promoting chromatin condensation and inhibiting access of the transcription machinery to the DNA. Histone modifications at distinct lysine residues may also allow for the recruitment of proteins capable of regulating transcription in a gene-specific manner.

CpG Methylation

Genome structure is also influenced by cytosine methylation, the only known biological base modification of DNA. Indeed, methylated cytosine has been referred to as the 'fifth base' because of its ability to convey heritable information. The extent of cytosine methylation plays a major role in the organisation of the genomic DNA. Densely methylated DNA is located in condensed

Cytosine → 5-Methylcytosine (DNMT; SAM → SAH)

Figure 2 Methylation of cytosine. Cytosine methylation is performed by DNMT (DNA methyltransferase), and occurs predominantly at CpG dinucleotide pairs, although cytosines in the context CpNpG are sometimes also methylated. The DNMT enzyme transfers a methyl group from *S*-adenosylmethionine (SAM) to the 5-carbon position of cytosine (boxed CH_3), forming 5-methylcytosine and leaving *S*-adenosylhomocysteine (SAH).

heterochromatin while sparsely methylated DNA is located in the more relaxed euchromatin. Methylation is the best studied epigenetic modification that occurs in cancer, and will be the main focus of the remainder of this chapter.

The genomic methylation patterns in gametic DNA are erased by a genome-wide demethylation shortly after fertilization. This is followed by *de novo* re-establishment of the methylation patterns after implantation. The mechanisms involved and the proteins guiding the erasure and resetting of the methylation patterns in the genome during embryogenesis are not yet fully understood. Several known DNA methyltransferases (DNMTs) in mammals, including DNMT1, DNMT3a and DNMT3b, catalyse the transfer of a methyl group from *S*-adenosylmethionine (SAM) to the 5-carbon position of cytosine (**Figure 2**).

This reaction occurs most commonly when cytosine is in a CpG dinucleotide sequence. All three enzymes are capable of performing both *de novo* and maintenance methylation (see below), but DNMT3a and DNMT3b appear to act by transferring methyl groups to previously unmodified CpG sequences, a process that occurs predominantly during embryogenesis.

In contrast, DNMT1 is thought to be the major maintenance DNA methyltransferase enzyme. The palindromic nature of the CpG target of DNMT1 is a key feature in the heritability of the DNA methylation profile during cell replication (**Figure 3**). During DNA synthesis, the nascent daughter strand is methylated by the methyltransferase enzyme positioned at the replication fork. DNMT1 preferentially recognizes the hemimethylated state of the two strands, and copies the methylation pattern of the parent strand on to the daughter strand. Thus, somatic cell methylation profiles represent epigenetic information that is faithfully replicated from one generation to the next.

The incidence of CpG dinucleotides in the genome is about 5–10-fold less than the approximate 6% frequency expected from the random distribution of the 16 possible dinucleotide combinations. The mammalian genome is thought to have progressively lost many of the methylated cytosines within CpG dinucleotide pairs during the course of evolution. This most likely occurred by endogenous deamination of methylated cytosines to form thymine (**Figure 4**). Although 60–80% of the CpGs within the mammalian genome are methylated, cytosines that reside in CpG islands, ranging from 200 to 4000 bases in length, are protected from methylation. They are therefore resistant to mutation by methyl group-driven deamination. These CpG islands are located near the promoter regions of approximately 50–60% of the genes within the genome, including all housekeeping genes. It is not unusual to find CpG islands also located within the 5′ coding region of genes or even in downstream introns.

The biological function of CpG methylation is not clearly understood. Methylation of genomic DNA is a modification employed by numerous species, including bacteria, plants and mammals; however, methylation is not detectable in yeast, *Drosophila* or *Caenorhabditis elegans*. Methylation serves as a host defence mechanism in prokaryotes to protect against the introduction of foreign DNA. DNA methylation in eukaryotes is proposed to similarly serve in host defence by protecting cells from transcription and transposition of endogenous retroviral sequences, and/or to reduce transcriptional 'noise' from very large genomes (Baylin *et al*., 1998; Robertson and Wolffe, 2000). The CpG islands associated with many of the genes located on the inactive X chromosome in females are also extensively methylated whereas the same alleles on the active X chromosome are unmethylated. Furthermore, the silenced allele of imprinted genes usually exhibits parent of origin-dependent dense methylation of at least one associated CpG island.

Hypermethylation of promoter region CpG dinucleotides is strongly correlated with the transcriptional silencing of genes. The causal relationship between cytosine methylation and gene silencing in mammals is supported by studies both *in vitro* and *in vivo*. Transfection experiments using reporter constructs with a methylated promoter

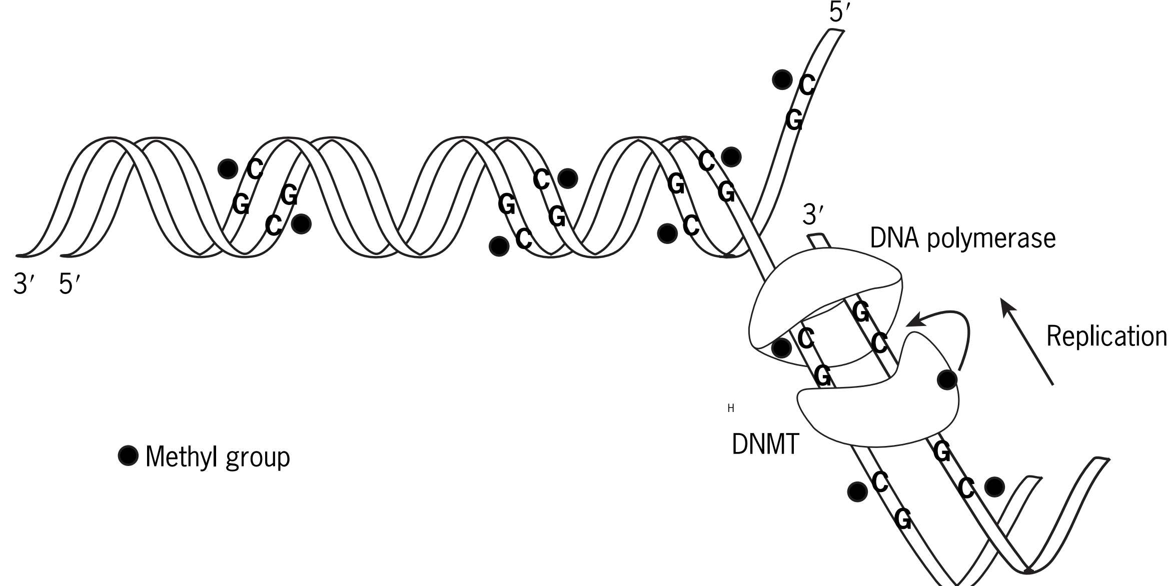


Figure 3 Replication and methylation of hemimethylated DNA. During DNA replication, the newly synthesized daughter strand is methylated by DNMT (DNA methyltransferase) within 1 min of synthesis. DNMT1 recognizes the hemimethylated state of the parent/daughter strand duplex and copies the methylation pattern of the parent strand CpG dinucleotide on to the daughter strand.

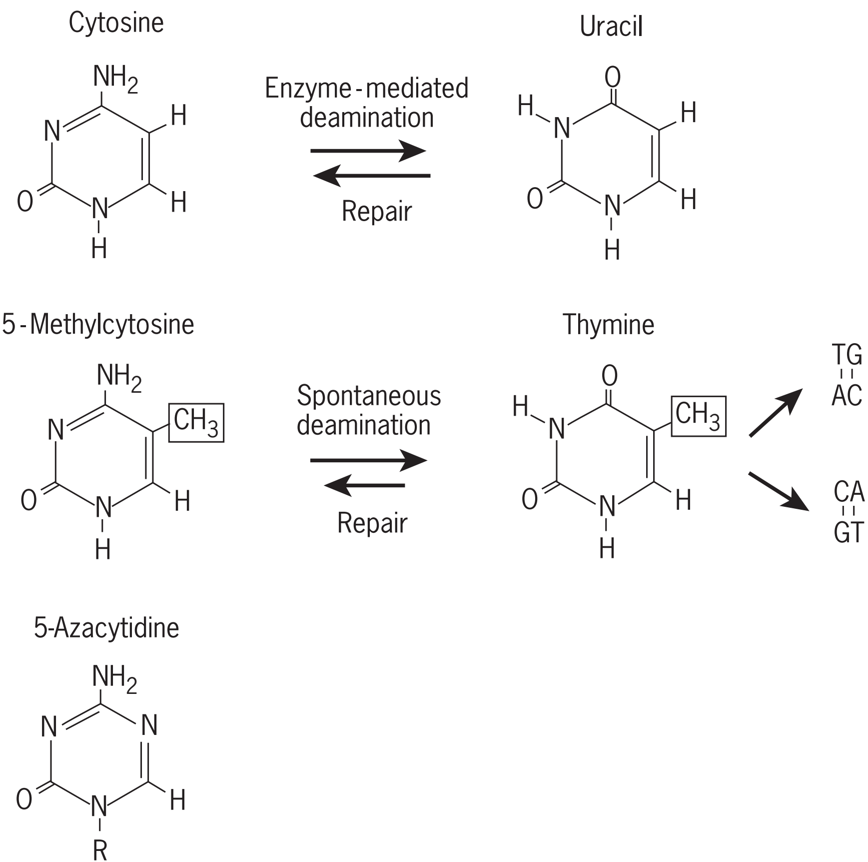


Figure 4 Structures of cytosine, uracil, 5-methylcytosine, thymine and 5-azacytidine (5-azaC). Cytosine is subject to DNMT mediated deamination to form uracil. DNA containing uracil is repaired by uracil DNA glycosylase. Spontaneous deamination of 5-methylcytosine produces thymine and repair of the resultant T–G mismatch is inefficient. Subsequent replication results in the production of either a TpG or CpA dinucleotide pair, depending on the strand affected by the original deamination event. When incorporated into either DNA or RNA, 5-azaC forms a covalent bond to DNMT (DNA methyltransferase) that results in sequestration of the enzyme followed by a progressive depletion of DNA methylation.

region show reduced transcription relative to that for constructs with unmethylated promoters. The DNMT inhibitor 5-azacytidine (5-azaC) (**Figure 4**) causes transcriptional reactivation of endogenous genes with hypermethylated promoters. Furthermore, homozygous disruption of *Dnmt1* in mice results in a three fold reduction in genomic 5-methylcytosine content, embryonic death (Li *et al*., 1992), and biallelic expression of imprinted genes (Li *et al*., 1993). These results emphasize the importance of cytosine methylation in gene regulation and embryogenesis. The deviations from normal methylation patterns frequently observed in cancer cells further suggest that epigenetic perturbations are mechanistically involved in oncogenesis.

MOLECULAR MODULATORS OF EPIGENETIC SIGNALS

Several protein complexes play fundamental roles in transcriptional control by recognizing and binding to methylated DNA. They can block the formation of transcription initiation complexes at methylated promoters by steric hindrance. The methyl CpG proteins (MeCP1 and MeCP2) are also capable of directing transcriptional repression through coordinated chromatin alterations. MeCP1 is a ubiquitously expressed protein that binds densely methylated DNA in a sequence-independent manner. It forms a complex with the methyl binding domain protein (MBD2), which further associates with HDACs. The HDACs are thought to contribute to the MeCP1-mediated transcriptional repression by causing chromatin condensation (Hendrich and Bird, 2000).

MeCP2 recruits several other proteins including a repressor of transcription, Sin3A, and an HDAC complex that induces chromatin condensation by deacetylating histone proteins (**Figure 5**) (Jones *et al*., 1998; Nan *et al*., 1998). Unlike MeCP1, MeCP2 can bind to DNA via a single methylated CpG dinucleotide; however, the efficiency of repression coordinated by the MeCP2 complex is dependent on the density of methylated CpGs (Magdinier *et al*., 2000). Dense promoter methylation may in itself

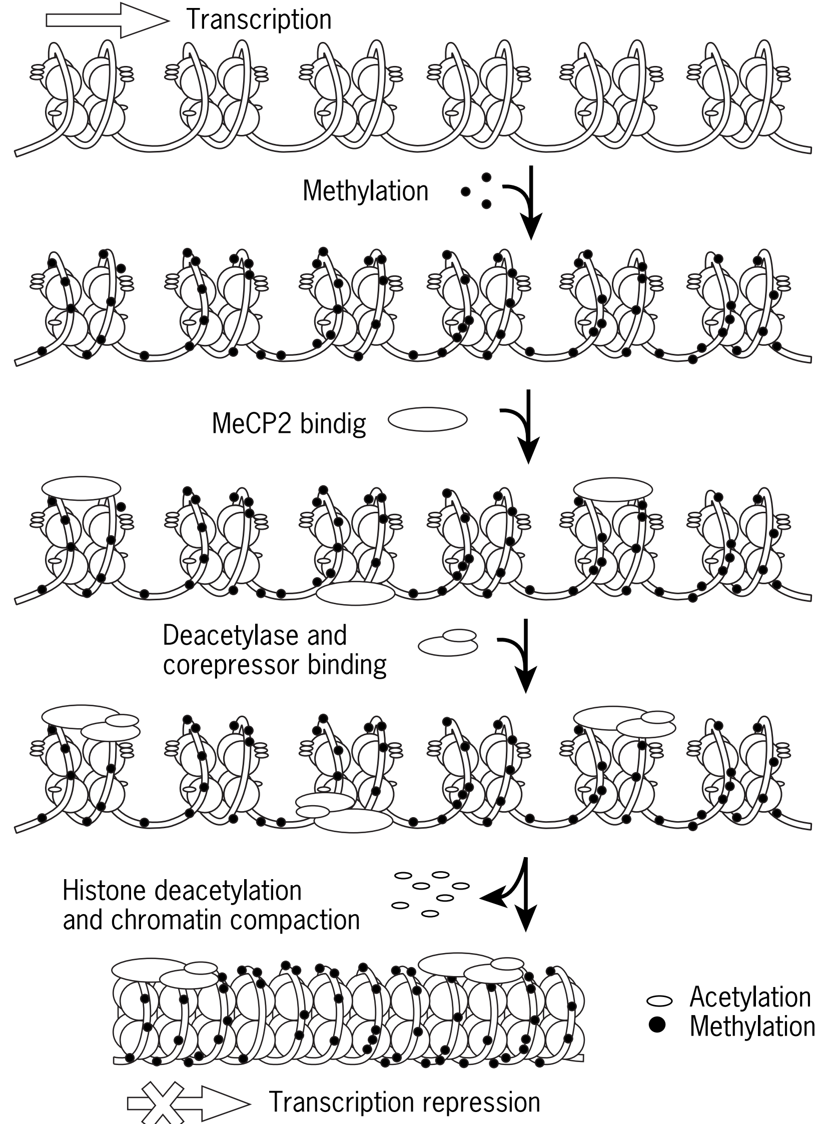


Figure 5 Model for the epigenetic inactivation of DNA transcription. A schematic representation of transcriptionally active open chromatin with acetylated histones is shown at the top. Cytosine methylation of the normally unmethylated promoter region is followed by recognition and binding by MeCP2 (methyl CpG protein). MeCP2 recruits HDAC (histone deacetylase) and other corepressor proteins that function to locally deacetylate histones followed by condensation of the chromatin; the result is promoter silencing. (Adapted from Jones and Laird, 1999.)

further inhibit the access of HAT activity to the hypoacetylated histones, preventing conversion to a more active chromatin configuration.

Recent studies supporting a link between DNA methylation and chromatin structure formation have shown that DNMT1 also interacts with HDAC2 to repress transcription (Fuks *et al*., 2000; Robertson and Wolffe, 2000; Robertson *et al*., 2000; Rountree *et al*., 2000). During DNA replication, unincorporated histones arrive at the replication fork in an acetylated state. DNMT1 and HDAC2 are both positioned at the replication foci when heterochromatin is replicated during late S phase. This association with HDAC2 could provide a mechanism whereby histones are deacetylated during assembly to facilitate the faithful structural reproduction of condensed chromatin from one cell to the next during cell division (Rountree *et al*., 2000).

Cytosine methylation and histone acetylation are proposed to act synergistically in the progressive silencing of genes. One model that accounts for tumour-suppressor gene silencing by epigenetic mechanisms invokes abnormal hypermethylation of the promoter CpG island followed by recruitment of methyl-DNA-binding proteins, including complexes such as MeCP2 that direct HDACs to the area of hypermethylation. Histone deacetylation, probably accompanied by other chromatin remodeling

events, can then cause local chromatin condensation and subsequent transcriptional repression (**Figure 5**) (Jones and Laird, 1999; Rice and Futscher, 2000). Thus, epigenetic events are poised to function mechanistically in carcinogenesis by silencing tumour-suppressor genes.

EPIGENETIC CHARACTERISTICS OF TUMOUR CELLS

There is widespread documentation of significant changes in the epigenome of cancer cells. These changes include an overall level of genomic hypomethylation coupled with gene-specific hypermethylation. Hypomethylation events are more generalized, and could lead to the activation of endogenous retroviral elements and any associated proto-oncogenes. However, gene-specific hypomethylation is unlikely to play a major role in oncogenesis since promoter CpG islands are normally unmethylated, with the notable exceptions of imprinted alleles and genes on the inactive X chromosome. In contrast, promoter-specific hypermethylation can lead to the silencing of tumour-suppressor genes.

Inherent Mutability of Methylated Cytosine

The methylation of CpG dinucleotides creates mutagenic susceptibility targets that can subsequently undergo endogenous deamination to form TpG (CpA on the opposite strand) dinucleotide pairs (**Figure 4**). In contrast to cytosine deamination that results in DNA containing uracil, 5-methylcytosine deamination creates a C to T base substitution that is not as readily recognized by DNA repair proteins as being misplaced in the DNA strand. This contributes to inefficient repair of these lesions, and subsequent accumulation of this type of mutation in the genome. Deamination of 5-methylcytosine in *p53* and *HPRT* is a frequent mutational event associated with human cancers (Pfeifer *et al*., 2000). Although the spectrum of mutations within *p53* varies between different forms of cancer, 50% of all point mutations in colon cancer arise from transition mutations of normally methylated CpG dinucleotides. In contrast, only 10% of liver and lung cancers contain these same mutations. In addition to the potential direct mutational inactivation of a gene by a CpG to TpG transition, these mutated sequences could hinder the interaction between DNA and specific proteins involved in transcriptional regulation.

Alterations in CpG Island Methylation

The genome regions subject to hypermethylation in cancer cells are the CpG islands associated with gene promoters. A study undertaken to identify all differentially methylated CpG islands in cancer estimated that of the 45 000 CpG islands in the human genome, 600 exhibit methylation patterns in tumors different from those in normal tissues (Costello *et al*., 2000). These methylation changes appear to occur early in the neoplastic process, and some are even cancer-type specific, suggesting that CpG island hypermethylation is mechanistically involved in carcinogenesis rather than being a consequence of neoplastic transformation. An increase in DNA methylation also occurs with ageing. Consequently, the increased cancer predisposition observed with ageing may be partially attributable to the age-dependent increases in genome methylation (see below).

A tumour-suppressor gene inactivated by CpG island hypermethylation would have a number of important characteristics. These include (1) dense methylation of the normally unmethylated CpG island present in the promoter region, (2) an absence of coding region mutations in the tumour, (3) a deficiency of gene-specific transcripts in the tumour, (4) gene reactivation in the tumour with DNMT inhibitors (e.g. 5-azaC) and (5) loss of gene function from hypermethylation comparable to that seen for inactivating mutations (Baylin *et al*., 1998). It is important to note that alterations in CpG methylation are mitotically heritable (**Figure 3**), and can potentially result in the clonal expansion of neoplastic cells if the epigenetic changes provide a selective growth advantage.

Hypermethylation of Genes in Cancer

The normally unmethylated promoter for many genes is increasingly methylated during neoplastic progression. **Table 1** lists genes whose promoter regions are hypermethylated in cancer, and a number of them are described in further detail below.

p16 located at human chromosome 9p21 is the best documented example of tumour suppressor gene inactivation by promoter CpG island hypermethylation. *p16* is involved in the process of cell cycle regulation. *p16* encodes for a cyclin-dependent kinase inhibitor that blocks cell cycle progression by inhibiting the action of cyclin D. Cyclin D binds to cyclin-dependent kinases and forms a complex that phosphorylates the Rb protein, allowing progression through the cell cycle. Thus, *p16* inactivation would enable cells to proceed inappropriately through cell cycle check points.

p16 was initially identified as a tumour suppressor in melanoma; however, its function as a tumour suppressor was questioned early on because many tumours with LOH at this gene locus lacked mutations in the remaining allele. This apparent discrepancy was ultimately explained by the finding that the promoter of the nonmutated allele was often hypermethylated. Promoter region hypermethylation is also frequently the only detectable alteration, suggesting that the two ‘hits’ required to inactivate a tumour suppressor can both be epigenetic in origin. *p16* promoter methylation also appears to be an early event in carcinogenesis since it is

Table 1 Genes subject to hypermethylation cancer

Gene	Locus	Function	Tumour type(s)
Tumour-suppressor genes			
APC	5q21	Regulation of β-catenin; cell adhesion	Colorectal, gastrointestinal
BRCA1	17q21	DNA damage repair	Breast, ovarian
CDH1 (E-cadherin)	16q22.1	Homotypic epithelial cell–cell adhesion	Bladder, breast, colon, liver
LKB1	19p13.3	Serine, threonine kinase	Hamartomatous colon, papillary breast
MLH1	3p21.3	DNA mismatch repair	MSI positive colorectal and endometrial
p15 (CDKN2B)	9p21	Cyclin-dependent kinase inhibitor	Acute leukaemias, Burkitt lymphoma, multiple myelomas
p16 (CDKN2A)	9p21	Cyclin-dependent kinase inhibitor	Lung, gliomas, breast, colon, bladder, nasopharyngeal, melanomas, prostate, thymomas, multiple myelomas, lymphomas
PTEN	10q23.3	Regulation of cell growth and apoptosis	Prostate
RB	13q14.2	Sequesters E2F transcription factor	Retinoblastoma
VHL	3p25	Inhibits angiogenesis, regulates transcription	Renal cell carcinoma
Candidate tumour-suppressor and other genes			
14-3-3-σ (stratifin)	1p	Cell cycle control	Breast
CALCA (calcitonin)	11p15.2-p15.1	Reduces serum calcium	Various carcinomas, leukaemia
CD44	11p13	Metastasis suppressor	Prostate
DAP-kinase 1	9q34.1	γ-IFN-induced promoter of apoptosis	Burkitt lymphoma, other B cell malignancies
EDNRB	13q22	Endothelin receptor	Prostate
ER	6q25.1	Oestrogen-induced transcriptional activation	Breast, colon, lung, leukaemia
FHIT	3p14.2	Dinucleoside oligophosphate hydrolase	Oesophageal
GST-π	11q13	Cellular detoxification	Prostate
H-cadherin	16q24.2-q24.3	Cell adhesion	Lung
HIC1	17p13.3	Zinc finger protien	Brain, breast, colon, renal
IGF2	11p15.5	Growth factor, paternally expressed	Rhabdomyosarcoma, Wilms tumour
MDG1	1p33-p32	Fatty acid-binding protein	Breast
MYO-D1	11p15.4	Myogenesis	Bladder, lung
O^6-MGMT	10q26	DNA repair	Brain, colon, lung, lymphomas
p73	1p36	Apoptotic response	Neuroblastoma, T and B cell malignancies
RASSFIA	3p21.3	Ras effector homolog	Lung
RAR-β2	3p24	Retinoic acid receptor/ transcription factor	Breast, colon, pancreas
TIMP-3	22q12.3	Inhibitor of matrix metalloproteinases	Brain, renal

already present in the preinvasive stages of cancer. In colon cancer, where *p16* is frequently hypermethylated, there is also a paucity of mutations in the *Rb* tumour-suppressor gene, indicating that *p16* inactivation alone is sufficient to circumvent the cell cycle G_1 block (Baylin *et al*., 1998). *p16* hypermethylation has now been implicated in many forms of cancer, including those that originate in the breast, bladder, brain, colon, oesophagus, head and neck, and lung (Liggett and Sidransky, 1998).

p15 and *p16* are positioned 15 kb apart at chromosome location 9p21, and both encode for cyclin-dependent kinase inhibitors. Whereas *p16* inactivation is found principally in solid tumours, loss of *p15* function occurs mainly in haematopoietic cell cancer. The constitutive expression of *p16* differs markedly from that of *p15*, which is more limited and regulated by transforming growth factor-β (Baylin *et al*., 1998). Hypermethylation in this chromosomal region can affect both *p15* and *p16* promoter

CpG islands simultaneously or it can be specific to either gene promoter. For example, in Burkitt's lymphomas both *p15* and *p16* are hypermethylated, whereas only *p15* is commonly hypermethylated in adult acute myelogenous leukaemia (AML), paediatric AML, adult acute lymphocytic leukaemia (ALL) and paediatric B cell ALL (Baylin *et al*., 1998). *p15* promoter hypermethylation is also present in the myelodysplastic state preceding leukaemia, indicating that DNA methylation of the *p15* promoter is an early carcinogenic event.

hMLH1 on chromosome 3p21.3 encodes for a protein essential for DNA mismatch repair. (See chapter on *Genetic Instability and DNA Repair*.) Disruption of *hMLH1* expression is frequently found in patients with hereditary nonpolyposis carcinoma, and also to a lesser extent in sporadic colon cancers (Tycko, 2000). The inactivation of *hMLH1* results in microsatellite instability (MIN + phenotype) which is characterized by errors in replication at one to four base pair repetitive microsatellite DNA sequences. *hMLH1* hypermethylation and the concomitant loss of gene expression is estimated to be present in approximately 70% of sporadic MIN + colorectal carcinomas whereas it is infrequent in MIN− tumours (Herman *et al*., 1998). *DNMT* inhibitors can restore DNA mismatch repair activity by causing the demethylation of the hypermethylated *hMLH1* promoter (Herman *et al*., 1998).

E-CADHERIN is located at chromosomal position 16q22.1. E-CADHERIN is involved in the calcium-dependent regulation of cell growth and differentiation by virtue of its ability to mediate homotypic cell-to-cell adhesion. Disruptions in these cellular adhesions are prominent in cancer cells with reduced *E-CADHERIN* expression, providing support for its normal role in preventing tumour invasion and metastasis. LOH at this chromosome location occurs often in breast cancer, and inactivating mutations in *E-CADHERIN* are also found in a variety of other solid tumours. Furthermore, decreased expression of *E-CADHERIN* in both breast and prostate cancers correlates with hypermethylation of the CpG island that encompasses the transcription start site (Baylin *et al*., 1998).

ER (Oestrogen receptor) at chromosome location 6q25 is expressed in a wide variety of tissues, and it encodes for a transcription factor that is activated only upon ligand binding. The presence of ER protein in breast cancer is prognostic for increased survival and lower risk of relapse. (See chapter on *Signalling by Steroid Receptors*.) Hypermethylation of the *ER* promoter has been found in a variety of tumour types including breast, colon and lung cancer as well as leukaemia (Baylin *et al*., 1998). An increase in *ER* promoter methylation in the normal colon also occurs with ageing, and this epigenetic change may predispose humans to colon cancer (Baylin *et al*., 1998; Issa, 2000).

GELSOLIN encodes for a calcium-dependent actin filament severing and nucleating protein whose expression directly correlates with the induction of cellular differentiation (Hoshikawa *et al*., 1994). *GELSOLIN* expression is always decreased during malignant transformation (Hoshikawa *et al*., 1994). Epigenetic silencing of *GELSOLIN* is a common feature of most human breast malignancies, a condition that is reversed in cultured cells by treatment with either HDAC inhibitors such as trichostatin A or DNMT inhibitors such as 5-azaC (Tycko, 2000). The ability to reverse silencing of *GELSOLIN* by either one of these agents alone is unusual since many other genes silenced by hypermethylation require treatment first with a DNMT inhibitor followed by a HDAC inhibitor to achieve substantial gene reactivation (Cameron *et al*., 1999).

O^6-MGMT on chromosome 10q26 encodes for the enzyme methylguanine–DNA methyltransferase. This ubiquitous protein is responsible for repairing mutagenic and carcinogenic O^6-alkylguanine adducts. The CpG island associated with the promoter of *O^6-MGMT* is hypermethylated and transcriptionally silenced in multiple solid tumour types, including those originating in the brain, colon and lung (Herman and Baylin, 2000). *O^6-MGMT* loss of function in colon cancer also plays a major role in determining the type and extent of mutations found in the oncogene K-*RAS*. G to A transition mutations result from unrepaired O^6-alkylguanine adducts, and they are the major type of K-*RAS* mutations present in human colon cancer. The finding that 70% of these K-*RAS* mutations are associated with a hypermethylated *O^6-MGMT* promoter suggests that epigenetic inactivation of this gene is involved in the formation of this disease (Herman and Baylin, 2000)

p73 is located at chromosome position 1p36, and encodes for a protein proposed to act as a neuroblastoma tumour suppressor. *p73* maps to an area of 1p36 commonly found to have LOH in neuroblastomas; however, the remaining allele in these tumours frequently lacks mutations. Wild-type p73 is reported to interact with c-Abl in promoting apoptosis, while a truncated p73 functions as an anti-apoptotic protein during neural development in mice. Several studies have concluded that the maternal allele of *p73* is functionally inactivated by the epigenetic process of genomic imprinting (see below). Although it is presently unclear if imprinting at the *p73* locus is mechanistically involved in carcinogenesis, *p73* promoter hypermethylation has clearly been documented in both T cell acute lymphoblastic leukaemias and Burkitt lymphomas (Herman and Baylin, 2000).

14-3-3-σ (***STRATIFIN***) on chromosome 1p is a member of the 14-3-3 family of proteins that mediate signal transduction by binding to other proteins containing phosphoserine motifs. *14-3-3-σ* transcription is induced by DNA damage through a p53 response element in the promoter. It is involved in the maintenance of cell cycle arrest at G_2, and it prevents entry into mitosis by sequestering the cdc2–cyclinB1 mitotic initiation complex in the cytoplasm. The promoter CpG island of *14-3-3-σ* is frequently hypermethylated with transcriptional silencing in 91% of

breast cancers while point mutations in the coding region and LOH occur infrequently (Ferguson *et al*., 2000). This transcriptional silencing is relieved by treatment of 14-3-3-σ-deficient breast cancer cell lines with 5-azaC.

BRCA1 (breast cancer type 1) is located on human chromosome 17q21. It encodes for an ~220-kDa nuclear phosphoprotein component of the RNA polymerase II holoenzyme complex. *BRCA1* is implicated in diverse cellular functions, including cell cycle control, chromatin structure-imposed transcriptional regulation and DNA damage response. Approximately 50% of familial forms of breast and ovarian cancer harbour germ-line mutations in *BRCA1*. Whereas 40–80% of sporadic breast carcinomas and 30–60% of sporadic ovarian carcinomas show LOH at this chromosomal location and have decreased levels of *BRCA1* mRNA, somatic mutations in *BRCA1* are very rare (Bianco *et al*., 2000; Esteller *et al*., 2000a).

The promoter of *BRCA1* is contained within a 2.7-kb CpG island. The 5′ end of this large CpG island is unusual in that it is normally methylated in somatic cells but not in the gametes. A repeat element located within this CpG island may be responsible for facilitating methylation of this region. *BRCA1* expression in normal breast, cervical and kidney cells is enhanced up to fivefold by treatment with DNMT inhibitors, suggesting that methylation of the 5′ portion of the CpG island in normal tissues may function in regulating *BRCA1* expression (Magdinier *et al*., 2000). Interestingly, *BRCA1* promoter methylation is significantly increased in medullary and mucinous subtypes of breast carcinomas relative to ductal subtypes; *BRCA1* methylation is also observed in conjunction with LOH in ovarian cancer (Esteller *et al*., 2000a). Evidence also suggests that *BRCA1* hypermethylation is specific to breast and ovarian cancer since the *BRCA1* promoter is not abnormally methylated in primary colorectal and liver carcinomas or in leukaemias (Bianco *et al*., 2000; Esteller *et al*., 2000a).

APC (adenomatous polyposis coli) tumour-suppressor gene is located at chromosome 5q21. It encodes for a large protein that normally associates with and negatively regulates signalling of the cell adhesion protein, β-catenin. Familial adenomatous polyposis (FAP) is an inherited disorder arising from germ-line mutations in *APC*. (See chapter on *Inherited Predispositions to Cancer*.) The majority of these mutations produce truncated proteins that are incapable of mediating β-catenin degradation, thereby resulting in the activation of genes involved in cell growth (Esteller *et al*., 2000b). *APC* is genetically inactivated early in the genesis of both familial and sporadic forms of colorectal carcinoma. The identification of FAP cases where either one or both alleles of *APC* lacked genetic lesions suggested an alternative form of *APC* repression. Methylation analyses confirmed that the *APC* promoter is commonly hypermethylated in colorectal adenomas and other cancers associated with the gastrointestinal tract, but not in brain, head and neck, lung and ovarian cancer (Esteller *et al*. 2000b).

HIC1 (hypermethylated in cancer) on chromosome 17p13.3 encodes for a zinc finger protein transcription factor that is a candidate tumour suppressor activated by p53 (Baylin *et al*., 1998). The normally unmethylated *HIC1* contains a CpG island that spans the entire gene, and hypermethylation of this CpG island has been documented in many solid tumours including those of the breast, colon, kidney and lung as well as in leukaemia. *HIC1* hypermethylation occurs in both acute lymphoblastic leukaemias (53%) and acute myeloid leukaemias (10%) (Issa *et al*., 1997). Furthermore, *HIC1* promoter hypermethylation is more prominent during blast crisis in chronic myelogenous leukaemia and the progression stage of acute lymphocytic leukaemia. This suggests that *HIC1* methylation is a later event in the genesis of leukaemia.

RAR-β2 (retinoic acid receptor beta) is located on chromosome 3p24. It is one of three known retinoic acid receptors that together belong to the nuclear receptor superfamily. *RAR-β* encodes for a putative tumour-suppressor protein that functions as a transcription factor when bound to retinoic acid (RA). RAR transcription factors bound to RA heterodimerize with retinoid X receptors, and then bind to retinoic acid responsive elements (RAREs) within the regulatory regions of RA-inducible genes. Transfection of *RAR-β* cDNAs into non-expressing tumour cells results in growth suppression, whereas transgenic mice expressing *RAR-β* antisense RNAs demonstrate enhanced tumorigenesis. These results demonstrate that *RAR-β* normally has an antiproliferative function. RAR-β levels are decreased in a number of cancers, including those that develop in the breast, cervix, head and neck, lung and ovary. Analysis of primary breast tumours indicates that approximately one-third have hypermethylation at the *RAR-β2* promoter (Sirchia *et al*., 2000). Loss of both RAR-β and ER function in breast cancers is correlated with resistance to RA treatment. The *RAR-β2* promoter is abnormally hypermethylated in several forms of cancer including breast cancer, colon carcinoma and pancreatic carcinoma, but its methylation status appears to not be related to the presence of functional ER protein (Sirchia *et al*., 2000).

IMPRINTED GENES AS CANCER SUSCEPTIBILITY LOCI

Genomically imprinted genes provide strong evidence that transcriptional silencing results from DNA methylation. This subset of genes is normally monoallelically expressed in a parent of origin-dependent manner. Imprinted gene expression in somatic cells depends upon the sex of the parent from which the allele originated, but not the sex of the individual. Every imprinted gene thus far examined has been associated with at least one differentially methylated CpG island. The epigenetic changes that confer the heritable

imprint mark have not yet been unambiguously defined, but cytosine methylation is the strongest candidate. Histone acetylation has also been proposed to be mechanistically involved in imprinted gene regulation since parental-specific differences in acetylation are associated with imprinted genes (Hu *et al*., 1998). Approximately 40 imprinted genes have been identified in humans to date, and estimates predict the presence of 100–200 imprinted genes in the entire genome (Barlow, 1995). The normal silencing of the imprinted allele is equivalent to a first 'hit' in the 'two-hit hypothesis' for carcinogenesis (**Figure 6**) (Knudson, 1971). A single genetic or epigenetic alteration in the expressed allele could therefore completely abrogate function of an imprinted gene. Since most imprinted genes are involved in cell growth and all are functionally haploid, they represent unique susceptibility loci for cancer development.

Human *IGF2* and *H19* are located in a chromosomal region (11p15.5) harbouring a cluster of imprinted genes. *IGF2* encodes for a potent mitogenic factor involved in cell growth and embryonic development whereas *H19* transcripts are non-coding. The reciprocally imprinted *IGF2* and *H19* genes are expressed from the paternally and maternally inherited alleles, respectively. The epigenetic regulation of this locus has been intensively studied in normal and malignant tissues. IGF2 overexpression occurs commonly in cancer (for a review, see Reik *et al*., 2000), and loss of imprinting is one mechanism responsible for the dysregulation of this influential growth factor (Rainier *et al*., 1993).

Human *IGF2* has four promoters that function in a tissue-specific manner (Vu and Hoffman, 1994). Promoters 2 to 4 are contained within a CpG island, and transcripts derived from these promoters are monoallelically expressed. In contrast, the P1 promoter 20 kb upstream of P2 is not associated with this CpG island, and drives biallelic expression of IGF2 in the liver (Vu and Hoffman, 1994; Baylin *et al*., 1998). During the ageing process, the P2 to P4 promoters of the normally unmethylated paternal allele are subject to incremental increases in methylation, and methylation of these promoters is also enhanced in cancer. Hypermethylation of this region in tumour cells results in decreased expression originating from the P3 promoter with the switching of *IGF2* transcription regulation to the non-imprinted P1 promoter (Vu and Hoffman, 1994). The ability of *IGF2* to undergo such promoter switching can be regarded as an epigenetic mechanism by which tumour cells can gain a selective growth advantage. Interestingly, the mouse *Igf2* does not have the equivalent of the human P1 promoter, and

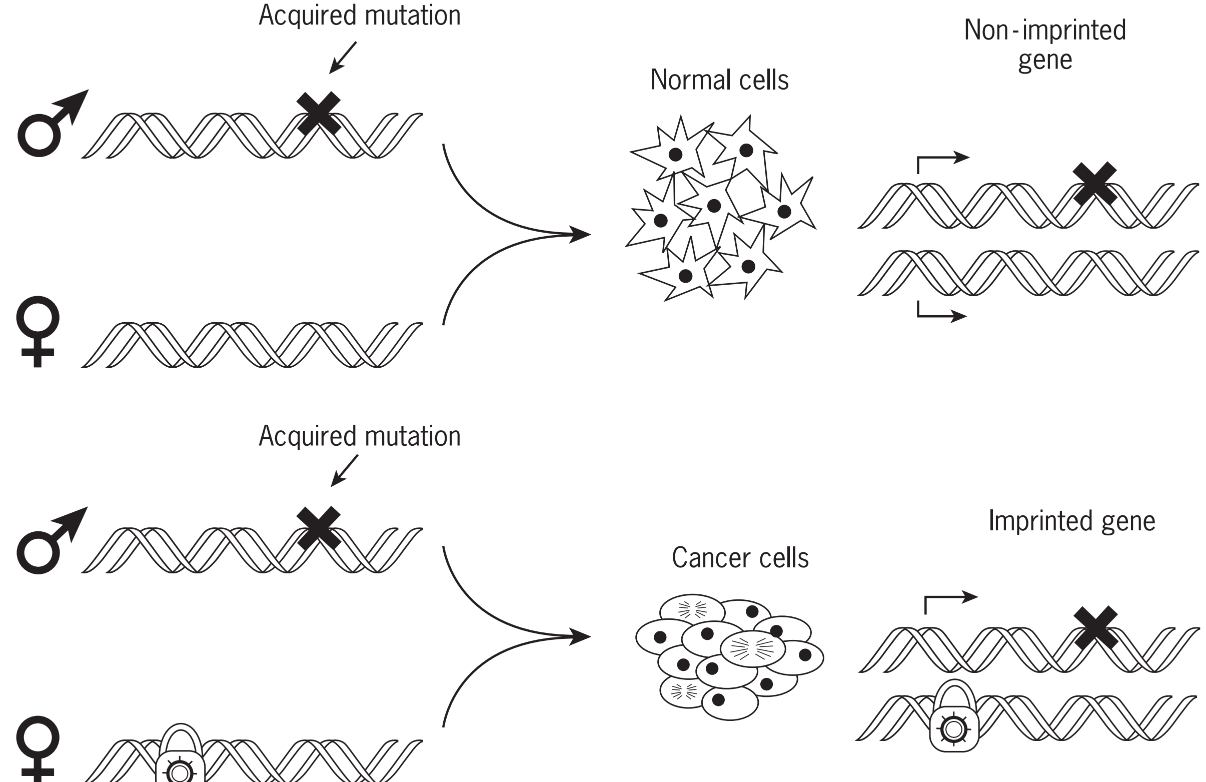


Figure 6 Imprinted genes as susceptibility loci in cancer. For most nonimprinted genes, an acquired mutation (X) does not contribute directly to carcinogenesis because of the presence of a second transcriptionally active wild-type allele. In contrast, acquisition of a genetic or epigenetic mutation on the active allele of an imprinted tumour-suppressor gene (paternal in this example) can directly result in cancer by inactivation of the single functional copy of the gene.

therefore biallelic expression cannot occur via this mechanism in mice. This fundamental difference between species in regulating *IGF2* expression is important to consider when extrapolating carcinogenic risk estimates from mice to humans. (See chapter on *Advantages and Limitations of Models for Cancer and Malignant Cell Progression*.)

The epigenetic control of *IGF2* imprinting has recently been further refined with the demonstration that imprinted expression of both *IGF2* and the adjacent, reciprocally imprinted *H19* depends on the presence of differentially methylated CTCF (vertebrate enhancer blocking protein) DNA-binding sites between these two genes (**Figure 7**) (Bell and Felsenfeld, 2000; Hark *et al*., 2000; Kanduri *et al*., 2000). The mechanism of CTCF-mediated transcriptional repression also involves CTCF interaction with HDACs (Lutz *et al*., 2000). In the case of *IGF2*, methylation of the paternal allele at the CTCF recognition sites prevents CTCF binding. In contrast, CTCF proteins bind to the unmethylated maternal allele, creating a physical boundary. This prevents the enhancer elements located downstream of *H19* from interacting with the *IGF2* promoter. Consequently, *IGF2* and *H19* are reciprocally imprinted, and transcribed from the paternal and maternal alleles, respectively. Deletion of this CTCF binding region on the maternal allele results in biallelic expression of *IGF2* (Thorvaldson *et al*., 1998; Srivastava *et al*., 2000). In colorectal cancers with loss of imprinting for *IGF2*, abnormal methylation of the maternal CTCF binding sites has been observed (Nakagawa *et al*., 2001). Together with the demonstration that methylation of CpGs within the CTCF binding site block CTCF binding (Kanduri *et al*., 2000), these results suggest a mechanistic link between the CTCF epigenetic control elements, loss of *IGF2* imprinting, and carcinogenesis.

We have recently provided evidence that another pair of human imprinted genes utilizes a similar mechanism to control their reciprocal imprinting pattern. The *DLK1/GTL2* domain at chromosome 14q32 encodes for the paternally expressed *DLK1* and the maternally expressed *GTL2* genes. *DLK1* is involved in several cellular differentiation processes including adipogenesis, haematopoiesis and

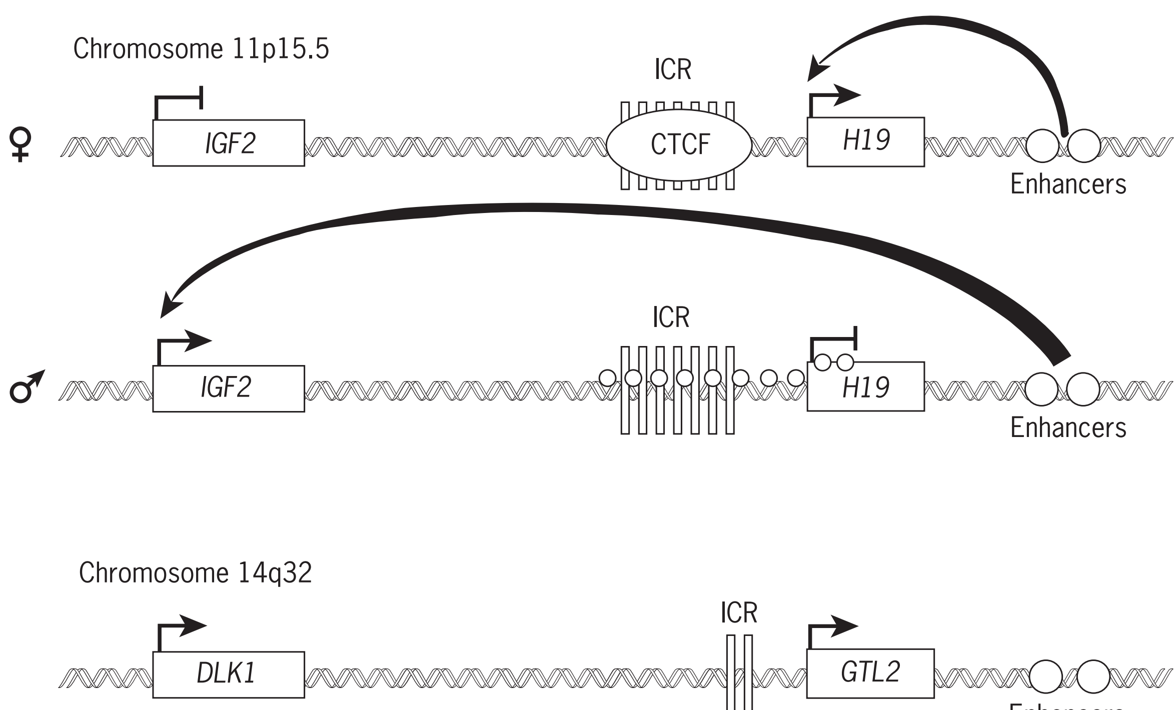


Figure 7 Cancer predisposition and methylation-dependent chromatin boundary elements. The reciprocally imprinted *IGF2* and *H19* genes at 11p15.5 are separated by approximately 100 kb of DNA that contains a differentially methylated ICR (imprint control region). The ICR contains seven methylation-sensitive CTCF (vertebrate enhancer blocking protein) binding sites (vertical bars). CTCF proteins are blocked from binding the methylated (small circles) paternal allele but the unmethylated maternal allele binds CTCF resulting in the formation of a chromatin boundary element. This boundary is thought to divert the enhancer elements, located downstream of *H19*, away from the maternal *IGF2* promoter and toward the available *H19* promoter, resulting in maternal *H19* expression. In contrast, the paternal *IGF2* is expressed because its promoter is subject to enhancer influence since methylation of the paternal ICR prevents CTCF binding and boundary formation. A second set of reciprocally imprinted genes was identified at 14q32 (Wylie *et al*., 2000). *DLK1* and *GTL2* share spatial, structural, and expression characteristics with the *IGF2/H19* locus and there are two differentially methylated consensus CTCF binding sites between these two genes.

neuroendocrine differentiation, and may play an important role in neuroendocrine tumorigenesis (Laborda, 2000). Like *IGF2* and *H19*, *DLK1* and *GTL2* are also separated by differentially methylated consensus CTCF binding sites. Furthermore, enhancer element sequences are located downstream from *GTL2* that are identical with those found downstream of *H19* (**Figure 7**) (Wylie *et al*., 2000). These findings suggest that this mechanism of epigenetic regulation may be commonly employed to coordinate the expression of juxtapositioned reciprocally imprinted genes. Thus, specific mutation of these CTCF binding sites and/or the CTCF binding proteins could potentially alter the expression of a number of imprinted genes in the genome.

NON-GENOTOXIC MEDIATORS OF CARCINOGENESIS

Non-genotoxic carcinogens induce cancer by causing epigenetic changes in the genome rather than by directly mutating the DNA base composition. Furthermore, both endogenous and exogenous factors can cause these epigenetic changes. Oxidation of guanine by endogenous chemicals often results in the formation of 8-hydroxyguanine that can contribute directly to genetic alteration because it is a potentially miscoding base, and also impair normal epigenetic DNA modification by impeding the methylation of adjacent cytosines. This is also true for photodimers, abasic sites, base alkylations and other oxidative DNA lesions that not only directly damage DNA, but also lead to heritable alterations in normal genomic methylation patterns (Wachsman, 1997; MacPhee, 1998). Additionally, a number of exogenous chemical agents are now known that induce both genome wide and gene specific chromatin structure changes by altering DNA methylation directly or by modifying DNMT activity (**Table 2**) (Baylin *et al*., 1998). Below is a brief description of specific compounds and factors that cause cancer through non-genotoxic mechanisms.

Table 2 Agents that influence DNA methylation

Induce hypomethylation	Induce hypermethylation
1,3-Bis(2-chloroethyl)-1-nitrosourea	3′-Azidodideoxythymidine (AZT)
4-Nitroquinoline 1-oxide	3-Deazaadenosine
4-Nitrosodiethylamine	5-Fluorouracil
4-Nitroso-*N*-ethylurea	5-Fluorodeoxyuridine
5-Azadeoxycytidine	Aphidicolin
5-Fluorodeoxycytidine	Butyrate
6-Thioguanine	Cisplatinum
7,12-Dimethylbenz[*a*]-anthracene	Colchicine
	Etoposide
Aflatoxin B_1	Doxorubicin
Benzo[*a*]pyrene	Nalidixic acid
Butyrate	Propionate
Butyryl-cAMP	Trapoxin
Bromobenzene	Trichostatin
Cyclophosphamide	Vinblastine
Ethionine	Vincristine
Hydralazine	
Lead nitrate	
MNNG, other alkylating agents	
N-Methyl-*N*-nitrosourea	
Novobiocin	
Oestradiol	
Procainamide	
Pseudoisocytidine	
Retinoic acid	
Teniposide	
Toposide	
Topoisomerase II inhibitors	

(Adapted from Holliday, 1991 and Wachsman, 1997.)

Nickel

Studies with carcinogenic nickel have led to development of a model for the potential epigenetic mechanisms by which non-genotoxic carcinogens inactivate tumor suppressor genes (Costa, 1995). Water-insoluble Ni^{2+} is a non-genotoxic carcinogen that localizes to the nucleus following its cell entry by phagocytosis. There it is thought to act through its affinity for heterochromatic regions of the genome. Chromosomal damage in these regions presumably results from oxidation that occurs when Ni^{2+} binds to chromatin proteins. These genotoxic events are normally not detrimental since the affected heterochromatin is usually genetically inactive; however, nickel also appears to cause harmful epigenetic modifications. This is postulated to result from nickel binding to oxygen atoms in the DNA phosphate backbone within heterochromatin and inducing a localized increase in DNA methylation. If this nickel-induced DNA methylation spreads outward to encompass adjacent euchromatic regions, tumour-suppressor genes within the proximity of the epimutated heterochromatin can be aberrantly silenced (Costa, 1995).

Peroxisome Proliferators

Peroxisomes are cytoplasmic organelles that are present in all eukaryotic cells except red blood cells. They are bounded by a single membrane and function in the metabolism of many substrates including long-chain fatty acids, sterols, dicarboxylic acids, prostaglandins, xenobiotics and oxygen free radicals. They are also involved in the synthesis of cholesterol, ether lipids, carbohydrates and bile acids (Masters and Crane, 1998). While peroxisomes are present in cells throughout the body, their density is cell-type dependent, with the largest number present in mammalian liver and kidney (Masters and Crane, 1998).

Peroxisomes are induced by a group of compounds collectively called peroxisome proliferators. These agents markedly increase peroxisome number, and they can also stimulate replicative DNA synthesis while suppressing apoptosis (Masters and Crane, 1998). Although this class of agents induces peroxisome proliferation in mice and rats through its interaction with the peroxisome proliferator activated receptor alpha (Peters *et al*., 1997), guinea pigs do not exhibit this effect, and neither do dogs, marmosets or humans. There is strong evidence that peroxisome proliferators are hepatocellular carcinogens in rodents; however, there is a substantial species-specific difference in the carcinogenic response to these agents (Roberts, 1999). They are considered to be non-genotoxic carcinogens since they act independently of covalent DNA binding and without evidence of genetic mutation. The mechanism by which peroxisome proliferators cause cancer is still under active investigation, but it is thought to involve increased oxidative stress caused by prolonged agent exposure (Masters and Crane, 1998). Importantly, accumulated data indicate that peroxisome proliferators do not constitute a serious carcinogenic risk to humans and other primates.

Diet

S-Adenosylmethionine (SAM) is the methyl donor for various methylation reactions in mammalian cells, including DNA methylation (**Figure 2**). A dietary deficiency in SAM precursors or cofactors involved in SAM biosynthesis (such as folate) leads to genomic hypomethylation, and diets deficient in these components have been widely used to induce liver tumours in rodents (Laird, 2000). In humans, dietary deficiency of methyl donors correlates with an increased risk for liver and colon tumours (Giovannucci *et al*., 1993). This hypomethylation is thought to be caused by either a deficiency in methyl group donors or by DNMT-enhanced cytosine deamination, but the mechanism has not yet been established (Laird and Jaenisch, 1996; Zingg and Jones, 1997). With the increasing awareness of the importance of DNA methylation in cancer formation, there is an expanding commercial interest in human dietary supplements fortified with methylation precursors such as SAM and folate. Studies to determine the specific contribution of diet to changes in DNA methylation patterns are certainly called for, and will help clarify the need for dietary intervention as a preventative or therapeutic measure against cancer.

Ageing

DNA hypomethylation was originally suspected to be responsible for the gene expression changes often observed with the ageing process. Interestingly, age-related decreases in DNA methylation occur primarily in the coding and intronic regions of genes, and they correlate poorly with observed reductions in gene expression (Issa, 2000). This disparity was clarified by studies showing that reduced gene expression that occurs with age results from a progressive increase in gene specific promoter methylation rather than generalized genomic hypomethylation.

These age-related increases in promoter CpG island methylation occur in a number of genes involved in cancer, including *IGF2*, *Versican*, *PAX6*, and *N33* in colon cancer and *HIC1* in prostate cancer (Issa, 2000). It is likely that many other genes will also fall into this category, because several studies designed to isolate differentially methylated CpG islands in cancer have identified a number of CpG islands that exhibit increased methylation with both ageing and neoplastic transformation (Issa, 2000). Not all age-dependent hypermethylation events result in cancer. The *ER* gene is hypermethylated in nearly all primary colon cancers, yet the normal colon of patients both with and without colon cancer has about the same yearly increase in *ER* promoter CpG island methylation (Issa, 2000). Since age-related hypermethylation varies among individuals of the same age, it is likely that genetic predisposition to epimutations, as well as exposure to environmental factors are involved in cancer formation. Thus, there is now compelling evidence of a mechanistic link between the ageing process and tumorigenesis in that age-related promoter hypermethylation frequently occurs in genes known to be involved in cancer formation (Issa, 2000).

DIAGNOSTIC AND THERAPEUTIC POTENTIAL OF EPIGENETIC ABNORMALITIES

Cancer Diagnosis

Sequencing data obtained from the human genome project are currently undergoing analysis to construct a human epigenetic map based on CpG content. This knowledge coupled with cross-species comparisons of the epigenome will be invaluable in deciphering the epigenetic elements involved in gene regulation (e.g., Hardison, 2000; Killian *et al*., 2000; Vu and Hoffman, 2000; Wylie *et al*., 2000). Epigenetic alterations in genes are early oncogenic events in some cancers, and detection of these early abnormalities may aid in protecting people from cancer through dietary alterations or pharmacological intervention (Laird, 1997). With increasing awareness of the importance of epigenetics in cancer formation, and the advent of laboratory techniques such as bisulfite DNA sequencing, methylation-sensitive PCR (**Figure 8**) and gene expression profiling by DNA microarrays, it is likely that methylation profiles will ultimately be used to predict an individual's predisposition to cancer, assist in cancer diagnosis and determine optimal therapeutic approaches (Esteller *et al*., 2001).

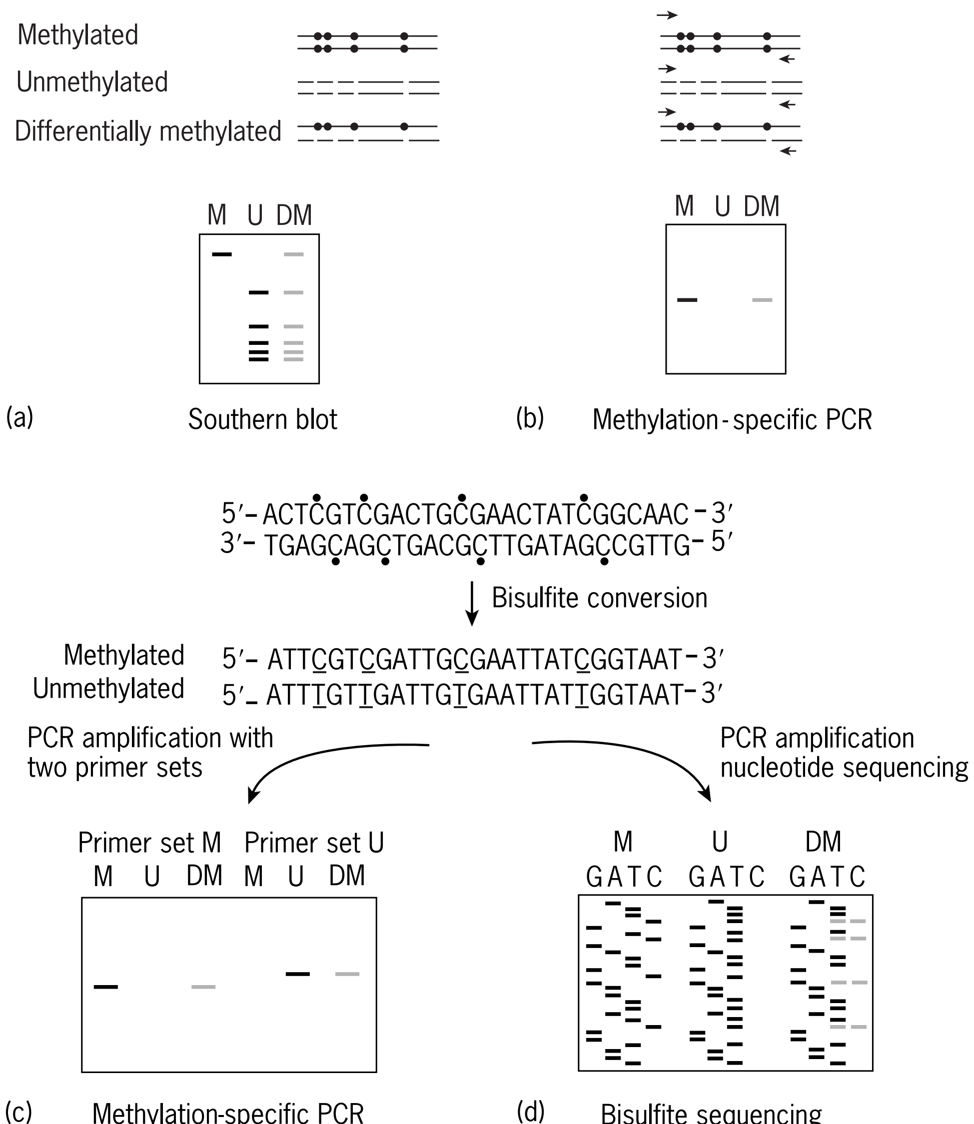


Figure 8 Methods used to analyse CpG methylation. Cytosine methylation is indicated by the circles. (a) Southern blotting depends on methylation-sensitive restriction enzymes to discriminate between methylated and unmethylated alleles. DNA is digested and fractionated on an agarose gel followed by blotting with a probe specific to the region of interest. Methylated recognition sites are resistant to digestion and will yield larger DNA fragments on the blot than unmethylated DNA. (b) Methylation-specific PCR requires digestion with methylation sensitive restriction enzymes, followed by PCR amplification. Methylated (uncut) DNA will yield an amplification product whereas unmethylated DNA will not be amplified. Bisulfite conversion of unmethylated cytosines precedes analysis by either a modification of methylation-specific PCR (c) or bisulfite sequencing (d). Sodium bisulfite treatment of DNA leads to the conversion of unmethylated cytosines to uracils while methylated cytosines are protected from conversion. Subsequent PCR using two independent primer sets designed to bisulfite protected and bisulfite converted sequence amplify methylated (M) and unmethylated (U) alleles, respectively. Bisulfite sequencing is the most direct means of analysing the methylation status of individual cytosines. Fully methylated cytosines are evident in the C lane of a sequencing gel using this method, while unmethylated cytosines are converted to thymines in the PCR amplification step prior to sequencing. Alleles having differential cytosine methylation (DM) are evident by the presence of bands in both the C and T lanes.

Cancer Treatment

A promising feature of alterations in DNA methylation patterns and chromatin structure in cancer cells is their potential for reversibility, because these modifications occur without changing the primary nucleotide sequence. The two major pharmacological targets associated with these epigenetic changes are DNMT

and HDAC. The DNMT inhibitor 5-azaC is structurally similar to cytosine (**Figure 4**), but when incorporated into DNA it forms a stable covalent bond with DNMT that inhibits further methylation by the sequestered enzyme. Consequently, overall genomic hypomethylation develops with subsequent rounds of DNA replication.

5-AzaC is efficacious in treating patients with acute leukaemia. It has also undergone clinical testing for the treatment of solid tumours; however, 5-azaC produces a high level of normal tissue toxicity and mutagenicity. These untoward side effects are not due to the resulting hypomethylation, but are attributed to the presence of the incorporated DNMT–5-azaC complexes in the genomic DNA (Laird, 1997). More specific strategies to inhibit the action of DNMT are being developed, including the use of antisense molecules. In this approach, antisense DNAs complementary to the DNMT mRNA inhibit methyltransferase activity by preventing DNMT translation. HDAC inhibitors, such as trichostatin A and sodium butyrate, have been shown to increase the level of histone acetylation in cultured cells, and to cause growth arrest, differentiation and apoptosis. Consequently, they are currently being tested in clinical trials as therapeutic agents for cancer.

Refinement in our understanding of the specific contributions of methylation and histone deacetylation to tumour-suppressor gene silencing in each type of cancer may make custom-designed treatments for gene reactivation possible. For example, *GELSOLIN* silencing can be reversed by treatment with HDAC inhibitors alone, while other tumour-suppressor genes achieve higher levels of reactivation when DNMT and HDAC inhibitors are used together. *RAR-β2* provides another example of the specificity of epigenetic reactivation. Loss of *RAR-β* expression by promoter hypermethylation can result in tumour resistance to treatment with all-*trans*-retinoic acid. Thus, demethylating agents in combination with all-*trans*-retinoic acid have been proposed for the treatment of cancers lacking *RAR-β* expression because of promoter hypermethylation (Côté and Momparler, 1997). Combining trichostatin A with all-*trans*-retinoic acid not only reactivated *RAR-β* in breast cancer cells, but also significantly reduced cell proliferation (Sirchia *et al*., 2000).

Another novel therapeutic approach proposes to use genetically engineered proteins to reactivate genes with epigenetically silenced promoters. In this approach, a chimeric fusion protein containing DNA-binding zinc finger motifs joined to protein domains having, or capable of recruiting HAT activity for example, might be used to target and alleviate localized areas of chromatin condensation in the promoter regions of silenced tumour suppressor genes. The specificity of the target sequence binding originates from the customized site-specific DNA contacts of the zinc finger domain. Once bound to the target sequence, the chromatin modifier activity would act locally to alleviate chromatin condensation and promote gene reactivation. These proteins could be tailored to virtually any sequence, and would be tethered to the particular protein domains required for the activation of the affected gene promoter (e.g. Liu *et al*., 2001). Combinations of these chimeric proteins might also expedite a positive clinical outcome by targeting multiple silenced genes simultaneously. Epigenetic cancer therapy also has major potential advantages over conventional therapeutic approaches. First, intact copies of tumour-suppressor genes do not need to be transfected into cells because they are already present in the cancer cell genome; they only need to be reactivated. Second, if gene-specific approaches are used to reactivate epigenetically silenced tumour-suppressor genes there should be little normal tissue toxicity, enabling them to be safely combined with more conventional therapies.

CONCLUSION

Recent years have seen a shift in thinking regarding the molecular basis for gene inactivation in cancer toward accommodating both genetic and epigenetic mechanisms. There has been a dramatic increase in the number of literature reports documenting hypermethylation of specific genes in numerous types of cancer, and it is clear that abnormal promoter hypermethylation is a prominent aetiological event. The specific mechanisms leading to epigenetic inactivation of genes in cancer must be further defined in addition to the roles of repressor complex components that coordinate silencing of these specific genes. With this knowledge, it may be possible to implement strategies in susceptible individuals to prevent deleterious epigenetic alterations that would otherwise lead to cancer. There is also the exciting possibility of developing novel therapeutic approaches for specifically alleviating abnormal promoter hypermethylation in tumours. The sequences involved in establishing the epigenetic profile of chromatin are clearly of fundamental importance to oncogenesis. Cross-species sequence comparisons in the future will greatly facilitate our ability to move from a single-gene to a genome-wide approach to identify conserved regulatory elements and determine their role in the epigenetic control of gene expression.

ACKNOWLEDGEMENTS

Research in the Jirtle laboratory was supported by NIH grants CA25951 and ES08823, DOD grant DAMD17-98-1-8305, Sumitomo Chemical Company, Ltd, and AstraZeneca Pharmaceuticals, Ltd.

REFERENCES

Barlow, D. P. (1995). Gametic imprinting in mammals. *Science*, **270**, 1610–1613.

Baylin, S. B., *et al.* (1998). Alterations in DNA methylation: a fundamental aspect of neoplasia. *Cancer Research*, **72**, 141–196.

Bell, A. C. and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*, **405**, 482–485.

Bianco, T., *et al*. (2000). Tumour-specific distribution of *BRCA1* promoter region methylation supports a pathogenic role in breast and ovarian cancer. *Carcinogenesis*, **21**, 147–151.

Cameron, E. E., *et al.* (1999). Synergy of demethylation and histone deacetylase inhibition in the reexpression of genes silenced in cancer. *Nature Genetics*, **21**, 103–107.

Costa, M. (1995). Model for the epigenetic mechanism of action of nongenotoxic carcinogens. *American Journal of Clinical Nutrition*, **61(suppl.)**, 666S-669S.

Costello, J. F., *et al*. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature Genetics*, **25**, 132–138.

Côté, S. and Momparler, R. L. (1997). Activation of the retinoic acid receptor gene by 5-aza-2′-deoxycytidine in human DLD-1 colon carcinoma cells. *Anticancer Drugs*, **8**, 56–61.

Esteller, M., *et al*. (2000a). Promoter region hypermethylation and *BRCA1* inactivation in sporadic breast and ovarian tumors. *Journal of the National Cancer Institute*, **92**, 564–569.

Esteller, M., *et al*. (2000b). Analysis of *Adenomatous Polyposis coli* promoter hypermethylation in human cancer. *Cancer Research*, **60**, 4366–4371.

Esteller, M., *et al*. (2001). A gene hypermethylation profile of human cancer. *Cancer Research*, **61**, 3225–3229.

Ferguson, A. T., *et al*. (2000). High frequency of hypermethylation at the 14-3-3 locus leads to gene silencing in breast cancer. *Proceedings of the National Academy of Sciences of the USA*, **97**, 6049–6054.

Fuks, F., *et al*. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genetics*, **24**, 88–91.

Giovannucci, E., *et al*. (1993). Folate, methionine, and alcohol intake and risk of colorectal adenoma. *Journal of National Cancer Institute*, **85**, 875–884.

Hardison, R. C. (2000). Conserved noncoding sequences are reliable guides to regulatory motifs. *Trends in Genetics*, **16**, 369–372.

Hark, A. T., *et al*. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature*, **405**, 486–489.

Hendrich, B. and Bird, A. (2000). Mammalian methyltransferases and methyl-CpG-binding domains: proteins involved in DNA methylation. *Current Topics in Microbiology and Immunology*, **249**, 35–54.

Herman, J. G. and Baylin, S. B. (2000). Promoter-region hypermethylation and gene silencing in human cancer. *Current Topics in Microbiology and Immunology*, **249**, 35–54.

Herman, J. G., *et al*. (1998). Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *Proceedings of the National Academy of Sciences of the USA*, **95**, 6870–6875.

Holliday, R. (1991). Mutations and epimutations in mammalian cells. *Mutation Research*, **250**, 351–363.

Hoshikawa, Y., *et al*. (1994). Trichostatin A induces morphological changes and gelsolin expression by inhibiting histone deacetylase in human carcinoma cell lines. *Experimental Cell Research*, **214**, 189–197.

Hu, J. F., *et al*. (1998). The role of histone acetylation in the allelic expression of the imprinted human insulin-like growth factor II gene. *Biochemical Biophysical Research Communications*, **251**, 403–408.

Issa, J. P. (2000). CpG-island methylation in aging and cancer. *Current Topics in Microbiology and Immunology*, **249**, 101–118.

Issa, J. P., *et al*. (1997). *HIC1* hypermethylation is a late event in hematopoietic neoplasms. *Cancer Research*, **57**, 1678–1681.

Jones, P. A. and Laird, P. W. (1999). Cancer epigenetics comes of age. *Nature Genetics*, **21**, 163–167.

Jones, P. L., *et al*. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, **19**, 187–191.

Kanduri, C., *et al*. (2000). Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation sensitive. *Current Biology*, **10**, 853–856.

Killian, J. K., *et al*. (2000). *M6P/IGF2R* imprinting evolution in mammals. *Molecular Cell*, **5**, 707–716.

Knudson, A. G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the USA*, **68**, 820–823.

Laborda, J. (2000). The role of the epidermal growth factor-like protein dlk in cell differentiation. *Histology and Histopathology*, **15**, 119–129.

Laird, P. W. (1997). Oncogenic mechanisms mediated by DNA methylation. *Molecular Medicine Today*, **3**, 223–229.

Laird, P. W. (2000). Mouse models in DNA-methylation research. *Current Topics in Microbiology and Immunology*, **249**, 119–134.

Laird, P. W. and Jaenisch, R. (1996). The role of DNA methylation in cancer genetics and epigenetics. *Annual Review of Genetics*, **30**, 441–464.

Li, E., *et al*. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.

Li, E., *et al*. (1993). Role for DNA methylation in genomic imprinting. *Nature*, **366**, 362–365.

Liggett, W. H. and Sidransky, D. (1998). Role of the *p16* tumor suppressor gene in cancer. *Journal of Clinical Oncology*, **16**, 1197–1206.

Liu, P., *et al*. (2001). Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. *Journal of Biological Chemistry*, **276**, 11323–11334.

Lutz, M., *et al*. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Research*, **28**, 1707–1713.

MacPhee, D. G. (1998). Epigenetics and epimutagens: some new perspectives on cancer, germ line effects, and endocrine disrupters. *Mutation Research*, **400**, 369–379.

Magdinier, F., *et al*. (2000). Regional methylation of the 5′ end CpG island of *BRCA1* is associated with reduced gene expression in human somatic cells. *FASEB Journal*, **14**, 1585–1594.

Masters, C. and Crane, D. (1998). On the role of the peroxisome in cell differentiation and carcinogenesis. *Molecular and Cellular Biochemistry*, **187**, 85–97.

Nakagawa, H., *et al*. (2001). Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of *H19*-associated CTCF-binding sites in colorectal cancer. *Proceedings of the National Academy of Sciences of the USA*, **98**, 591–596.

Nan, X., *et al*. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386–389.

Peters, J. M., *et al*. (1997). Role of PPAR alpha in the mechanism of action of the nongenotoxic carcinogen and peroxisome proliferator, Wy-14,643. *Carcinogenesis*, **18**, 2029–2033.

Pfeifer, G. P., *et al*. (2000). Mutation hotspots and DNA methylation. *Current Topics in Microbiology and Immunology*, **249**, 1–20.

Rainier, S., *et al*. (1993). Relaxation of imprinted genes in human cancer. *Nature*, **362**, 747–749.

Reik, W., *et al*. (2000). Igf2 imprinting in development and disease. *International Journal of Developmental Biology*, **44**, 145–150.

Rice, J. C. and Futscher, B. W. (2000). Transcriptional repression of *BRCA1* by aberrant cytosine methylation, histone hypoacetylation and chromatin condensation of the *BRCA1* promoter. *Nucleic Acids Research*, **28**, 3233–3239.

Riggs, A. D. and Porter, T. N. (1996). Overview of epigenetic mechanisms. In: Russo, V.E.A., *et al*. (eds), *Epigenetic Mechanisms of Gene Regulation*. 29–45 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Roberts, R. (1999). Peroxisome proliferators: mechanism of adverse effects in rodents and molecular basis for species differences. *Archives of Toxicology*, **73**, 413–418.

Robertson, K. D. and Wolffe, A. P. (2000). DNA methylation in health and disease. *Nature Reviews in Genetics*, **1**, 11–19.

Robertson, K. D., *et al*. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature Genetics*, **25**, 338–342.

Rountree, M. R., *et al*. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genetics*, **25**, 269–277.

Sirchia, S. M., *et al*. (2000). Evidence of epigenetic changes affecting the chromatin state of the retinoic acid receptor *β2* promoter in breast cancer cells. *Oncogene*, **19**, 1556–1563.

Srivastava, M., *et al*. (2000). H19 and Igf2 monoallelic expression is regulated in two distinct ways by a shared *cis* acting regulatory region upstream of H19. *Genes and Development*, **14**, 1186–1195.

Thorvaldsen, J. L., *et al*. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes and Development*, **12**, 3693–3702.

Tycko, B. (2000). Epigenetic gene silencing in cancer. *Journal Clinical Investigation*, **105**, 401–407.

Vu, T. H. and Hoffman, A. R. (1994). Promoter-specific imprinting of the human *insulin-like growth factor-II* gene. *Nature*, **371**, 714–717.

Vu, T. H. and Hoffman, A. R. (2000). Comparative genomics sheds light on mechanisms of genomic imprinting. *Genome Research*, **10**, 1660–1663.

Wachsman, J. T. (1997). DNA methylation and the association between genetic and epigenetic changes: relation to carcinogenesis. *Mutation Research*, **375**, 1–8.

Wylie, A. A., *et al*. (2000). Regulatory motifs of the novel imprinted domain, *DLK1/GTL2*, mimic those of *IGF2/H19*. *Genome Research*, **10**, 1711–1718.

Zingg, J. and Jones, P. A., (1997). Genetic and epigenetic aspects of DNA methylation on genome expression, evolution, mutation, and carcinogenesis. *Carcinogenesis*, **18**, 869–882.

FURTHER READING

Baylin, S. B., *et al*. (2001). Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Human Molecular Genetics*, **10**, 687–692.

Cheung, P., *et al*. (2000). Signaling to chromatin through histone modifications. *Cell*, **103**, 263–271.

Murphy, S. K. and Jirtle, R. L. (2000). Imprinted genes as potential genetic and epigenetic toxicological targets. *Environmental Health Perspectives*, **108(Suppl. 1)**, 5–11.

Reik, W. and Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*, **2**, 21–32.

Wade, P. A. (2001). Transcriptional control at regulatory checkpoints by histone deacetylases: molecular connections between cancer and chromatin. *Human Molecular Genetics*, **10**, 693–698.

Websites

http://www.geneimprint.com (Genomic Imprinting website).

http://www3.mdanderson.org/leukemia/methylation (MD Anderson resource on CpG island methylation in ageing and cancer).

http://dnamethsoc.server101.com/ (DNA Methylation Society homepage).

http://www.cancer.med.umich.edu/learn/1learn.htm (University of Michigan Comprehensive Cancer Center homepage).

http://condor.bcm.tmc.edu/ermb/tgdb (The Tumor Gene Database).

http://nhgri.nih.gov/histones/ (The Histone Sequence Database).

Chapter 24

Infectious Agents and Cancer

Robert Newton
ICRF Radcliffe Infirmary, Oxford, UK

CONTENTS

- Introduction
- Viruses
- Bacteria
- Helminths
- The Future

INTRODUCTION

Infections play an important role in cancer. With the exception of canine venereal transmissible sarcoma, cancer is not itself contagious, but the underlying cause can be. Not only is the process whereby infection leads to cancer important for gaining insights into oncogenesis, but also control or elimination of infection holds promise for cancer prevention.

Historical Perspective

Transmissible agents have a venerable part in the history of cancer research. In 1911, Peyton Rous, often considered to be the father of tumour virology, was the first to demonstrate the acellular transmission of a sarcoma between chickens (the term ‘virus’ had not yet been coined). The research community was not receptive to the notion that a chronic disease may have an infectious cause and it was to be 55 years before Rous received the Nobel Prize for his seminal discovery. In the 1930s, Shope discovered oncogenic pox viruses and papillomaviruses in rabbits. In 1936, Bittner demonstrated that predisposition to breast cancer in C3H mice was transmitted in breast milk. In 1951, Gross discovered the first murine leukaemia virus and, in 1960, Hilleman identified SV40 virus as a contaminant of polio vaccine grown in monkey kidney cultures. However, the concept that infections might cause chronic diseases, such as cancer, can be traced back even further. For example, in the nineteenth century, the simple epidemiological observation that cancer of the uterine cervix was relatively common in prostitutes, but unknown in celibate nuns, led to the suggestion that the cause might be linked to sexual behaviour and perhaps even be sexually transmitted. In 1905, several years before Rous published his work on chicken sarcomas, Goebel drew attention to ‘the occurrence of bladder tumours due to bilharziasis’ (shistosomiasis).

In 1964, the first human tumour virus (the Epstein–Barr virus (EBV)) was discovered using electron microscopy, in Burkitt lymphoma cells, by Epstein, Achong and Barr. Later, EBV was also detected in undifferentiated nasopharyngeal carcinoma and subsequently in several other tumours. By the 1970s, cancer viruses were in fashion. President Nixon ‘declared war’ on cancer (National Cancer Act, 1971) and funding was increased for the National Cancer Institute’s ‘special virus cancer programme.’ Although the 1970s saw many important developments, including the discovery of oncogenes and tumour-suppressor genes (*TP53*), no new cancer viruses were identified and interest began to wane. This was to change in the early 1980s, with several major discoveries. In 1980, Poiesz, Gallo and colleagues discovered the human T cell leukaemia virus, which is associated with endemic leukaemia/lymphoma, particularly in Southern Japan and the Caribbean. In 1981, the large-scale prospective epidemiological studies of Beasley *et al.*, in Tiawan, confirmed the long-suspected causal association between the hepatitis B virus and liver cancer. In 1983, zur Hausen and colleagues isolated HPV 16 from a human cervical cancer specimen, Marshall and Warren identified *Helicobacter pylori* (later associated with gastric cancer), and the HIV (discovered by Barré-Sinoussi) emerged as an important cause of several cancers. Hepatitis C virus, a cause of liver cancer, was discovered in 1989 and, in 1994, Chang and Moore identified the Kaposi sarcoma-associated herpesvirus (HHV-8).

The Global Burden of Cancers Caused by Infections

It is estimated that approximately 15% of cancers (between about 1.2 and 1.5 million cases per year, worldwide) are attributable to viral (11%), bacterial (4%) and helminth (0.1%) infections (**Table 1**). Collectively, infectious agents are the most important known cause of cancer after tobacco.

Table 1 Cancers attributable to infections (these are conservative estimates, adapted from Parkin *et al.*, 1999)

Infection	Cancer(s)	No. of cases worldwide
Human papillomaviruses	Cervical cancer	360 000
	Other female genital cancers	
Hepatitis B virus	Liver cancer	230 000
Epstein–Barr virus	Burkitt lymphoma	100 000
	Hodgkin disease	
	Nasopharyngeal cancer	
Epstein–Barr virus/HIV	Non-Hodgkin lymphoma	9 000
Human herpesvirus-8/HIV	Kaposi sarcoma	45 000
Human T cell leukaemia virus	Leukaemia	3 000
Hepatitis C virus	Liver cancer	110 000
Helicobacter pylori	Gastric carcinoma	350 000
	Gastric lymphoma	
Schistosomes	Bladder cancer	10 000
Liver flukes	Cholangiocarcinoma	1 000

A better understanding of the role of infectious agents in the aetiology of cancer is a public health imperative, because such cancers are theoretically preventable by vaccination or early treatment of infection. Furthermore, cancer-causing infections often cause substantial morbidity and mortality from non-malignant conditions. Therefore, an additional benefit of any scheme to reduce the burden of cancers caused by infections would also involve a reduction in the incidence of other diseases.

The majority of infection attributable cancers (perhaps 1 million cases per year) occur in the developing world, reflecting the higher prevalence of the major causative agents, particularly hepatitis B, human papillomaviruses (HPV), *H. pylori* and human immunodeficiency virus (HIV). It is conservatively estimated that if these infectious diseases were controlled, up to one in four cancers in developing countries and one in 10 cancers in developed countries might be prevented. This chapter briefly reviews the association between certain infections and cancer, outlines the mechanism by which disease might be caused (if known) and presents comments on the potential for prevention of such tumours.

VIRUSES

Human Papillomaviruses (HPV)

The papillomaviruses are double-stranded DNA viruses. About 100 subtypes have been distinguished to date, many of which can infect humans, and at least as many again await formal characterization. Several specific subtypes have been associated with cancer in humans. Indeed, more cancers are attributable to HPV infection than to any other transmissible agent (**Table 1**). Most of these are cervix cancers, but other anogenital tumours and cancers at distant sites may also be caused by HPV.

After cancer of the breast, cervix cancer is the most common female malignancy worldwide. The risk of disease is increased among women reporting multiple sexual partners, early age of first sexual intercourse and among those whose male partners have multiple partners, all features that implicate a sexually transmitted aetiological agent. Of the known HPV types, about 30 can infect the female genital tract. Some of these are associated with benign lesions, such as warts (e.g. HPV 6 and 11) while others, so-called 'high-risk' types, are associated with invasive cancer and advanced precancerous lesions (e.g. HPV 16, 18, 31, 33, 45, 51, 52, 58, 59).

HPV infection is one of the most common sexually transmitted infections of women and probably also of men. Viral DNA is detectable in a large proportion of women shortly after becoming sexually active and the main determinant of infection is the number of sexual partners. The prevalence of infection varies between populations, but is of the order of 20–30% in women aged 20–24 years, declining to 5–10% in women over the age of 40 years. Follow-up of young, sexually active cohorts of women suggest that the incidence of HPV infection is about 15% per year, with more than 50% of women becoming infected at some stage in their lives. Most infections are cleared spontaneously, but a small proportion become persistent and it is these that carry the risk of neoplastic change. Much less is known about HPV infection in men and their reservoir of infection has not been clearly identified, although the glans penis and internal meatus are the most likely locations.

HPV infection of the genital tract can be latent, or associated with cellular alterations known as cervical intraepithelial neoplasia (CIN) (sometimes called squamous intraepithelial lesions (SIL)), graded according to severity from 1 to 3, depending on the degree of nuclear and cytoplasmic change. All grades are considered to be manifestations of HPV infection; grade 1 lesions are generally

benign and often resolve, whereas high-grade lesions are more likely to persist and progress. The time taken to progress from one grade to another and then to invasive disease is unknown, but has been estimated to be several years, possibly more than a decade. The features associated with HPV-mediated progression of CIN are not clear, but persistent infection is more likely in older women infected with 'high-risk' HPV subtypes. The quality of the immune response to infection is likely to be important and some HLA associations have been reported. Other possible factors include high parity, concomitant venereal infections and tobacco use (see IARC (1995) and Herrero and Muñoz (1999) for a more thorough review).

The evidence for a causal role of HPV infection in the aetiology of CIN and invasive cancer of the cervix is overwhelming. Virtually all squamous cell cancers and more than 90% of adenocarcinomas of the cervix contain HPV-DNA. A recent study by Walboomers *et al*. (1999) identified HPV-DNA in 99.7% of almost 1000 invasive cancers in a worldwide study. HPV 16 accounts for about 55% of tumours and is particularly dominant in Western countries. HPV 18 and 45 are relatively more important in tropical areas and account for about 15 and 10% of cervix cancers, respectively. Data from case-control and cohort studies consistently suggest that the risk of invasive cervix cancer or CIN is very high in association with HPV infection (relative risks of greater than 50 in HPV infected women compared with uninfected women). Furthermore, the risk increases with increasing viral load, as measured by the amount of HPV-DNA.

Molecular analyses support the epidemiological evidence. The high-risk HPV subtypes exhibit transforming potential and can immortalize cells, processes fundamental to the development of malignancy. The HPV genome contains two oncogenes, E6 and E7, and despite the frequent loss of much of the viral genome in cervical cancer cells, these regions are consistently maintained and expressed. The protein products of the E6 and E7 oncogenes affect the normal function of cellular proteins, essential for regulating cell growth. TP53 and Rb are important tumour-suppressor proteins and loss of their function is a common theme in most human cancers, regardless of origin. E6 interferes with p53 and E7 with the Rb protein and this is likely to be central to the oncogenic activity of certain HPVs (Phillips and Vousden, 1999).

There is some evidence that high-risk HPV types also play a role in the aetiology of cancers at other anogenital sites (IARC, 1995). These include cancers of the anus, penis, vagina and certain histological subtypes of vulval cancer. However, these tumours are sufficiently rare (particularly in comparison with cancer of the cervix), as to have only a limited impact on public health. A proportion of tumours of the head and neck (including conjunctiva), oesophagus, lung, bladder and prostate may also be associated with HPV infection, although the evidence remains scant. Of more public health importance is the suggestion that HPV subtypes 5 and 8 may cause a proportion of squamous cell carcinomas of the skin. Epidermodysplasia verruciformis is a rare hereditary disease characterized by impaired cell mediated immunity. Patients with this condition suffer from multiple skin warts, which often progress to squamous carcinoma. HPV 5 and 8 have been consistently identified in these tumours (McGregor and Proby, 1996); similar subtypes have been found in warts and malignant skin lesions of renal transplant and other immunosuppressed patients. The role of HPV 5 and 8 in the aetiology of squamous cell skin cancers occurring in the general population is uncertain.

Prevention of HPV-associated cancers, particularly cervix cancer, is a public health priority. In the long-term, vaccination holds the most promise for the eradication of such tumours, but of more immediate value is screening. Cervical cancer screening programmes rely on the detection of treatable precancerous lesions by exfoliative cervical cytology, and are known to be effective at reducing the incidence of and mortality from invasive disease. Several studies suggest that detection of high-risk HPV types is more sensitive for detecting CIN than conventional cytology, although the rates of false positives are also higher (Cuzick, 1999). The inclusion of HPV detection as part of the normal cytological screening programme may, therefore, lead to better disease detection, increased intervals between screens and early cessation of screening in women without HPV infection (IARC, 1992). Large trials are needed to determine the value for prevention and cost effectiveness of such an approach.

Prophylactic HPV vaccines are based on the induction of neutralizing antibodies able to prevent infection and associated lesions. Their development has been slowed by the inability to propagate HPV in tissue culture and the consequent lack of a source of antigens, in particular, the structural proteins L1 and L2. These problems have largely been solved with the generation of virus-like particles (VLPs), obtained by expression of the major capsid proteins, which because they are indistinguishable from authentic virions except that they lack a viral genome, can induce neutralizing antibodies. Such vaccines have proven efficacy in protecting against infection in many animal models. Several VLP-based vaccines have been produced to protect against human HPV types (6, 11, 16, 18, 31 and 33) and are being tested in phase I and II trials (Coursaget and Muñoz, 1999). It will be many years before large-scale phase III clinical trials yield results, but such vaccines hold great promise for the ultimate prevention of HPV-associated cancers.

Hepatitis B Virus

The role of chronic infection with hepatitis B virus (HBV) in the aetiology of primary liver cancer, specifically hepatocellular carcinoma, is well established (IARC, 1994a). Hepatitis B is one of the most common infectious

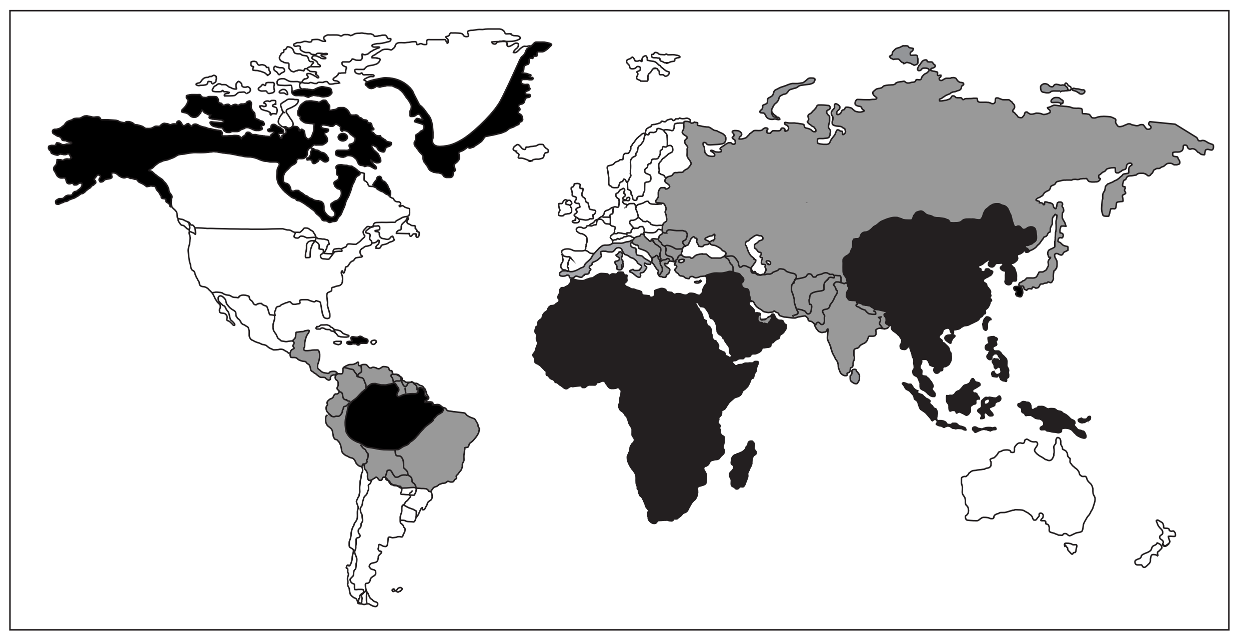


Figure 1 The geographical distribution of hepatitis B infection. Black, $\geq$8%, high; grey, 2–7.9%, intermediate; white, $<$2%, low (from WHO).

diseases worldwide, with between 300 and 350 million chronic carriers – up to 10% of the population in high prevalence areas – two-thirds of whom will develop chronic hepatitis. Of those, 20–25% will eventually die from primary liver cancer or cirrhosis. As a result, primary liver cancer is one of the most common cancers, particularly in parts of sub-Saharan Africa, China and South-east Asia, where the virus is prevalent (**Figure 1**). The hepatitis B virus is responsible for causing up to a quarter of a million cancers per year (**Table 1**), about 50–55% of the total number of liver cancers, and is the second most important oncogenic virus (Parkin *et al*., 1999).

Hepatitis B virus belongs to a group of hepatotropic double-stranded DNA viruses (with a single-stranded region of variable length), called hepadnaviruses and is characterized by the fact that it replicates through an RNA intermediate, via reverse transcription. The genome contains four genes, S, C, pol (the largest gene, coding for reverse transcriptase) and X. The virion, or 'Dane particle,' has an outer protein coat encoded by the S gene (hepatitis B surface antigen (HBsAg), the so-called 'Australia antigen,' discovered by Blumberg in 1963, in the serum of an aboriginal man), and an inner protein core (hepatitis B c antigen (HbcAg)) coded by the C gene. The terminal region of the C gene encodes the hepatitis B e antigen (HbeAg) and the X gene encodes a protein that upregulates transcription from all the viral and some cellular promoters.

Transmission of HBV occurs from a person with acute infection or carrier status (i.e. HBsAg positive on at least two occasions, 6 months apart) who has circulating virus. Individuals who are HBeAg positive are particularly infectious since the presence of this antigen correlates with high serum levels of HBV-DNA. The mode of transmission is not completely clear but varies with age, occurring primarily at three stages in life. First, transmission can occur from a mother to her child. Neonates born to HBeAg-positive mothers have about a 90% chance of becoming infected, whereas children of HBeAg-negative mothers have a 30% probability of infection. This is a particularly important means of transmission in China, but is probably less important in Africa where, for reasons that are not known, fewer infected mothers are HBeAg positive. Transmission in childhood is probably the most important age for infection globally and is associated with residence with infected siblings; the exact means of virus spread within households is unknown. In adult life, the major routes of transmission are by needles and sexual intercourse. Since the introduction of screening for HBV in blood products, the predominant mode of needle transmission is by the sharing of needles amongst intravenous drug users. Sexual transmission has been most clearly documented in countries of Europe and North America, where the prevalence of HBV is low, but may also have a limited role in endemic countries (IARC, 1994a).

The principle risk factor for HBV-induced liver cancer (and cirrhosis) is persistence of replicative infection (as it is with human papillomavirus infection and cancer of the cervix) and a key determinant of persistence is the age of infection. About 90% of children infected in the perinatal period become chronic carriers; up to 5 years old, around 20–30% of those infected become chronic carriers and above this age, the proportion is probably less than 5%. Presumably, the relationship between persistence of

infection and age is due to changes in the ability of the immune system to contain the virus, but the determinants of this are unclear (Wild and Hall, 1999). At all ages, males are more likely than females to become chronic carriers. Other factors are known to contribute to the risk of liver cancer in those chronically infected with HBV, including coinfection with the hepatitis C virus. Similarly, exposure to aflatoxins, common in parts of Africa and Asia, leads to a roughly 10-fold increase in the risk of malignancy above the excess associated with HBV infection.

The mechanisms whereby HBV may induce primary liver cancer are not fully understood. In most instances, the tumour is preceded by several decades of chronic hepatitis or cirrhosis and the associated rapid cell turnover may render host DNA more susceptible to malignant change. Mutagenesis may be induced by the associated inflammatory response or exogenous carcinogens, or may result from genetic instability following integration of the viral genome. Another possible mechanism involves the activation of cellular genes by the hepatitis B virus X gene product: it encodes a protein that may interact with the TP53 and deregulate the cell cycle.

Prospects for prevention of HBV-induced liver cancer are good. Screening of blood and organ donors has reduced the spread of infection among adults in developed countries. Similarly, screening for infection among high-risk pregnant women, such as those from endemic countries, is of value. There is ample evidence that the use of hepatitis B immunoglobulin within hours of birth, followed by hepatitis B vaccine on multiple occasions, is effective in reducing the infection rate in infants born to infected mothers by more than 80% and the carrier rate by more than 90% (reviewed by Cuzick *et al*., 1999). The best hope for prevention, however, lies with mass vaccination against HBV, particularly in populations where infection is most prevalent. Furthermore, there is also some interest in the development of therapeutic vaccines; these might eventually be used to treat liver cancer patients, by inducing an effective immune response to virus-infected cells.

The safety and effectiveness of prophylactic HBV vaccines have been clearly demonstrated in numerous studies, since they became widely available in 1982. Early vaccines were based on purified HbsAg from patients with HBV-induced liver cancer, but these were superseded in 1987 by vaccines produced using recombinant technology. Such vaccines are highly immunogenic and induce protective antibodies in 90–100% of healthy recipients (reviewed by Coursaget and Muñoz, 1999). Some studies have demonstrated protection from infection for up to 12 years. In a study in the Gambia with 10 years of follow-up, hepatitis B vaccine prevented 83% of infections and 94% of chronic carriage (Gambia Hepatitis Study Group, 1987). Although it will be many years before an effect on the incidence of liver cancer is noted in clinical trials, the introduction of mass vaccination in Taiwan has been associated with a sharp decline in the incidence of liver cancer in children (Chang *et al*., 1997). Currently, 80 countries, mostly in high-risk endemic areas, have routine hepatitis B vaccination programmes. Despite this, it is estimated that only a third of infants who might benefit from vaccine worldwide actually receive it. (See also chapter on *Human DNA Tumour Viruses*.)

Epstein–Barr Virus (EBV)

The Epstein–Barr virus is a ubiquitous human herpesvirus of the family Gammaherpesviridae that infects more than 90% of the world's population. The virion consists of a core wrapped with DNA and is contained within a capsid, which is surrounded by a membranous envelope with glycoprotein spikes. The genome is a relatively large 172-kb double-stranded DNA molecule encoding about 100 genes. There are two major subtypes of EBV which differ with respect to some of the genes coding for nuclear proteins in latently infected cells, but there is no evidence that one subtype is more strongly associated with disease than the other.

The EBV is transmitted orally, either by exchange of virus particles or infected cells in buccal fluid and, like other herpesviruses, it establishes a latent infection, with life-long persistence in the infected host. Asymptomatic primary infection usually occurs in childhood, although in more developed countries infection may be delayed. If primary infection occurs in adolescence, about 50% of cases may develop the clinical syndrome of infectious mononucleosis and most of what is known of the events occurring at primary infection are extrapolations from the study of such individuals.

It is not known whether the initial target cell for orally transmitted virus is a B lymphocyte, made accessible by damage to the oral mucosa, or an epithelial cell of the mucosa itself. However, following primary infection, foci of productive (lytic) infection are established in the oro-pharyngeal mucosa and a large pool of latently infected B cells can be found both in the blood and in lymphoid tissues. The overgrowth of virally transformed B cells is controlled by specific cytotoxic T cell responses, the absence of which (in allograft recipients and others with impaired T cell function) can result in EBV-driven lympho-proliferation and even lymphoma.

A recent review of the role of the EBV in the aetiology of cancer concluded that there is sufficient evidence for the carcinogenicity of EBV in the causation of several types of lymphoma (including Burkitt lymphoma, Hodgkin disease and immunosuppression-related lymphomas) and undifferentiated nasopharyngeal carcinoma (IARC, 1997). Other cancer types have also been linked to infection with EBV, although the evidence for causality is less clear. In total, it is estimated that the EBV may be responsible for up to about 100 000 cancers per year worldwide (**Table 1**; Parkin *et al*., 1999).

Burkitt lymphoma occurs throughout the world with varying frequency, but everywhere it represents a

significant proportion of malignant lymphomas in children. There are two broad types of Burkitt lymphoma, which although histologically indistinguishable, differ in several other important ways. The 'endemic' form of the disease was first described in 1958 in parts of sub-Saharan Africa, South-east Asia and Papua New Guinea. It is particularly frequent where malarial infection is heavy and widespread. In these areas it is the most common childhood malignancy, with a peak incidence between 5 and 8 years old. It often presents with facial lesions and is associated with infection with EBV. Viral DNA is almost invariably identified in tumour tissue in monoclonal form, indicating that the original malignant clone must have arisen from a single virus infected cell. Furthermore, high antibody titres to EBV infection in young children are predictive of the subsequent development of the tumour (de-Thé *et al*., 1978). This is in contrast to the sporadic form of Burkitt lymphoma seen throughout the rest of the world, which usually presents with abdominal lesions, occurs at older ages (particularly in teenagers) and is only associated with EBV infection in about a third of cases. In Western populations, sporadic Burkitt lymphoma is about 1000 times more common in HIV-infected individuals than in those uninfected with HIV and about half of these tumours are also associated with EBV infection. Both the endemic and sporadic Burkitt lymphoma involve c-*MYC* gene translocations, but the breakpoints in each type are different: in the endemic disease, the breakpoints tend to be far upstream of the oncogene, whereas in the sporadic form, the breakpoints tend to be adjacent to or within the oncogene. It is therefore possible that the endemic and sporadic forms of Burkitt lymphoma have different aetiologies, with only a proportion (mostly the endemic cases) being caused by EBV infection.

Nasopharyngeal carcinoma of undifferentiated or poorly differentiated type is an epithelial EBV-associated tumour, which, like Burkitt lymphoma, is characterized by marked geographic and population differences in incidence. It is common in southern China and South-east Asia, where it may represent up to 20% of all cancer cases. It also occurs relatively frequently in Eskimo populations and in Mediterranean Africa. Despite this geographical restriction, more cases of nasopharyngeal cancer are attributed to infection with the EBV, than any other cancer (Parkin *et al*., 1999).

The identification of an association between nasopharyngeal cancer and EBV occurred by chance, when patients with the tumour were chosen as controls in a case-control study of Burkitt lymphoma and were found to have high antibody titres to the virus. Specifically, elevated immunoglobulin A (IgA) antibodies to the EBV viral capsid antigens (VCA) predate the development of nasopharyngeal cancer by several years and are also correlated with tumour burden and recurrence. Measurement of IgA levels has formed the basis of a screening programme in southern China, where about 5% of individuals aged 30 years or more are positive for IgA/VCA and 5–12% of these have nasopharyngeal cancer (IARC, 1997). In addition, clonal EBV-DNA is consistently detected in tumour tissue (but not in normal nasopharyngeal epithelium), suggesting that the tumour develops from a single EBV-infected cell.

Because the EBV is ubiquitous, the geographical variation in the incidence of nasopharyngeal cancer may be explained by the combined influences of genetic predisposition in particular racial groups and of local environmental or dietary factors, although the exact nature of these effects is not clear. Salted and preserved food in the diet and a lack of fresh fruits and vegetables are currently the most clearly established cofactors. Another possibility is that particular strains of EBV, which carry a higher risk for nasopharyngeal carcinoma, are common where the disease occurs, although evidence for this remains scant.

Hodgkin's disease is a malignant lymphoma characterized by the loss of lymph node architecture, with the majority of infiltrating cells being benign. Indeed, the malignant cells – the Reed–Sternberg cells – constitute only about 2% of tumour mass. The disease occurs worldwide and in Western populations it has a distinctive age distribution with two peaks of incidence, at ages 25–30 and over 45 years. Sero-epidemiological studies indicate that high antibody titres to EBV precede the development of Hodgkin's disease and a history of infectious mononucleosis is a strong risk factor. Furthermore, about half of cases have evidence of clonal EBV-DNA in tumour tissue. Parkin *et al*. (1999) estimated that about half of the 60 000 new cases annually are related to infection with the EBV.

Some other lymphomas may be related to EBV, but the overall number of cases is likely to be relatively small. Non-Hodgkin lymphomas associated with HIV, or immunosuppressive therapy, are characterized by a number of features, including an aggressive clinical course. High-grade disease is common and extra-nodal sites are often involved, with lesions in the central nervous system being virtually unknown, except in the immunosuppressed. The most common subtype of non-Hodgkin lymphoma both in those with HIV infection and in immunosuppressed transplant recipients, is B cell immunoblastic lymphoma. Post-transplant immunoblastic lymphomas are nearly always associated with EBV and probably represent the end result of an EBV-driven lymphoproliferation in the absence of effective T cell immunity. EBV sequences are detectable in about 50% of HIV-associated immunoblastic lymphomas (although 100% of primary cerebral lymphomas), suggesting that other factors may also be important (Newton *et al*., 1999a). Several nonspecific host factors have been suggested to play a role in lymphomagenesis in immunosuppressed individuals, such as disrupted immunosurveillance, chronic antigenic stimulation and cytokine dysregulation, all of which might be responsible for expanding the B cell population from which a lymphoma subsequently develops.

Infection with the EBV has also been associated with a number of very rare T cell lymphoma subtypes and with

lymphoepithelial carcinomas of the stomach, lung and salivary gland. In addition, smooth muscle tumours (leiomyosarcomas) in immunosuppressed individuals uniformly contain EBV. The number of such cancers, however, is small.

The development of vaccines to control the diseases associated with EBV infection are at an early stage of development, but may be used in a number of ways. First, prophylactic vaccines may modify or prevent primary infection. Normally, infection occurs during the first few years of life, but such vaccines would be particularly useful for the prevention of infectious mononucleosis, which results in 50% of individuals in whom infection is delayed until adolescence. Second, postinfection vaccination may be used to modify the existing immune status of an infected individual and could have some value in preventing the development of nasopharyngeal carcinoma in those who have high IgA antibodies against EBV. Finally, therapeutic vaccination might be selectively targeted against viral antigens expressed in tumour cells. However, for the time being, the effect of any form of vaccination against the EBV remains speculative.

Human Herpesvirus-8 (HHV-8/KSHV)

Before the HIV epidemic, Kaposi sarcoma showed a greater geographical variation in incidence than almost any other cancer. It was as common in parts of sub-Saharan Africa, such as Uganda and eastern Zaire, as colon cancer is in Europe and the USA, representing up to 9% of all cancers in men. Kaposi sarcoma was also endemic, although much rarer, in countries around the Mediterranean, particularly Italy, Greece and the Middle East, but was almost nonexistent elsewhere in the world, except in immigrants from those endemic countries. In all of these areas, Kaposi sarcoma was more common in men than in women (Newton *et al*., 1999a).

It was the appearance of aggressive forms of Kaposi sarcoma in the USA in the early 1980s that heralded the onset of the HIV epidemic in Western countries. Although the incidence of Kaposi sarcoma has increased in populations at high risk of HIV in northern Europe and the USA, it existed at such a low level before the onset of the epidemic that it remains a relatively rare tumour. However, parts of Africa with a high prevalence of HIV and where Kaposi sarcoma was relatively common even before the era of acquired immunodeficiency syndrome (AIDS), have seen an explosion in the incidence of the disease. In the last 10–15 years the incidence of Kaposi sarcoma has increased about 20-fold in Uganda and Zimbabwe, such that it is now the most common cancer in men and the second most common in women (IARC, 1997).

In 1994, Chang, Moore and colleagues identified sequences of a new herpesvirus in a biopsy specimen of Kaposi sarcoma from an HIV-infected homosexual man, using representational difference analysis. The virus – human herpesvirus-8 (HHV-8) or Kaposi sarcoma-associated herpesvirus (KSHV) – has been consistently associated with Kaposi sarcoma and is now considered to be the principal cause of the disease. Genomic sequences of HHV-8 are present in tumour cells of Kaposi sarcoma lesions (specifically in the spindle cells, which constitute the bulk of the tumour) in virtually all subjects, but are not found in other tissues (with the exception of blood). The presence of HHV-8, detected by polymerase chain reaction (PCR) or serology, in peripheral blood, predicts the subsequent development of Kaposi sarcoma, particularly in individuals with high anti-HHV-8 antibody titres.

HHV-8 is not a ubiquitous virus, but is most prevalent in groups or populations at highest risk of developing Kaposi sarcoma, such as HIV-infected homosexual men in the USA and in African populations where the tumour has long been endemic (Boshoff, 1999). The proportion of adults in the general population with antibodies against HHV-8 ranges from fewer than 5% in northern America and northern Europe to around 10% in southern Europe and more than 30% in black Africans. In the United States and Europe, more than 30% of HIV-infected homosexual men have been found to have antibodies against HHV-8.

The modes of transmission of HHV-8 are yet to be fully elucidated. In the USA, sex between men may be an important route of transmission since this is the main behavioural risk factor for Kaposi sarcoma and indeed there is now some evidence that this is so. In some African countries, where HHV-8 is relatively common, the seroprevalence does not vary by sex and has been found to increase with age, from birth, through childhood and into adult life. This suggests that some transmission from a mother to her child and from child to child is also likely.

HHV-8 is a gamma-herpesvirus, closely related to the EBV, and infects CD19+ B cells as well as the endothelial derived spindle cells of a Kaposi sarcoma lesion. In addition to its role in Kaposi sarcoma, the virus also causes a rare type of lymphoma (primary effusion lymphoma) and a lymphoproliferative B cell disorder (a subtype of Castleman disease). The mechanism by which HHV-8 causes disease is controversial, although the genome encodes several putatively transforming genes in addition to genes encoding a number of regulatory cytokines and angio-proliferative factors that may facilitate tumour growth. HHV-8 infection alone, however, may not be enough to induce disease. Cofactors such as immunosuppression, as a result of advancing age, HIV/AIDS or therapy following tissue transplantation are presumably required.

Human T Cell Leukaemia Virus Type 1 (HTLV-1)

HTLV-1 is the main causal agent of adult T cell leukaemia/lymphoma, a disease characterized by malignant proliferation of CD4-positive T lymphocytes. Clinical features

include hypercalcaemia, lymphadenopathy, skin lesions due to leukaemic cell infiltration, involvement of the spleen and liver and immunodeficiency. The prognosis of patients with acute adult T cell leukaemia/lymphoma is poor, and few survive more than a few months following diagnosis. HTLV-1 also causes slowly progressive myelopathy (tropical spastic paraparesis) and uveitis.

HTLV-1 is an enveloped retrovirus, of the family Oncornavirinae, containing two covalently bound genomic RNA strands, which are combined with several viral enzymes, including reverse transcriptase. The prevalence of infection with HTLV-1 varies widely worldwide, with high levels in diverse geographic areas. Antibodies to HTLV-1 are found in 5–15% of indigenous adult populations in southern Japan, the Caribbean, South America, central Africa, Papua New Guinea and the Solomon Islands. Within endemic areas, clusters of especially high prevalence can occur. Carriers can be found elsewhere in the world, but are mostly individuals who moved from endemic areas. It is estimated that there are between 15 and 20 million infected individuals in the world (IARC, 1996).

Three modes of transmission of HTLV-1 have been identified: mother-to-child transmission, mainly due to breast-feeding beyond 6 months, sexual transmission, predominantly from men to women, and transmission by transfusion of cellular blood products and through intravenous drug use. Control and prevention of infection depends on reducing transmission by these three major routes. Perinatal transmission has been greatly reduced in Japan by avoidance of prolonged breast feeding and several countries have introduced universal screening of blood donors. Passive and active immunization is effective in animal models but no preventive vaccine is yet available for humans.

Adult T cell leukaemia/lymphoma (ATLL) occurs almost exclusively in areas where HTLV-1 is endemic, such as Japan, the Caribbean and West Africa, and cases described elsewhere have generally been in immigrants from those endemic regions, or their offspring. Early studies showed that infection with HTLV-1 is so closely associated with adult T cell leukaemia that it is now part of the diagnostic criteria used for defining the disease. All antibody-positive cases of adult T cell leukaemia have monoclonally integrated HTLV provirus in the malignant cells, suggesting that the tumour is an outgrowth of an individual T cell clone. The virus is able to immortalize human T lymphocytes, a property that has been related to a specific viral gene *tax*, which has been identified as a transforming factor. ATLL develops in about 2–5% of HTLV-1-infected individuals and is especially frequent among those infected early in life. No other environmental cofactors for disease have so far been identified. (See chapter on *RNA Viruses*.)

Hepatitis C Virus (HCV)

The identification of HCV in 1989 by Choo and colleagues arose from an investigation of the causes of post-transfusion non-A, non-B hepatitis. It is a single-stranded RNA virus assigned to a separate genus within the family Flaviviruses (which includes yellow fever virus and dengue virus) and is completely unrelated to the HBV. To date, six major subtypes of HCV have been identified, which have different geographical distributions. Acute infection often causes only mild illness, but it is becoming increasingly clear that HCV is responsible for substantial morbidity and mortality, particularly from chronic liver disease and hepatocellular cancer (IARC, 1994a). It may also play a role in the aetiology of other malignancies, such as non-Hodgkin lymphoma and cancers of the oral cavity (Tanaka and Tsukuma, 1999).

The prevalence of infection with HCV varies around the world and is estimated to be about 1–1.5% in Europe and the USA, about 3% in Japan and Oceania (excluding Australia and New Zealand) and up to 3.6% in Africa (Parkin *et al.*, 1999). In most countries, the prevalence of infection is the same in men as it is in women and increases steadily with age. Transmission is primarily by the parenteral route and, before the introduction of screening for hepatitis C, blood transfusions were a major source of infection. Intravenous drug users comprise a substantial proportion of identified cases in western populations and health-care workers, renal dialysis patients and those with clotting disorders are also at an increased risk. Both sexual and perinatal transmission occur and household contact with an infected family member may also account for a proportion of cases. However, almost half of all hepatitis C-infected individuals have no identifiable risk factors.

In contrast to hepatitis B, as many as 85% of hepatitis C virus infections become persistent and at least two-thirds of those individuals go on to develop chronic liver disease, including hepatocellular carcinoma. The risk of liver cancer in chronically infected individuals is around 20-fold higher than in the general population (more in people coinfected with HBV) and it has been estimated that HCV causes around 110 000 cancers per year (Parkin *et al.*, 1999; **Table 1**). This represents about 25% of all liver cancers, with particularly high proportions in Africa (41%), Japan (36%) and Oceania (33%). Although a similar proportion of men and women become chronic carriers of HCV, the development of cancer is more frequent in men. Alcohol and tobacco have been implicated as cofactors and may account for this discrepancy.

The mechanism by which HCV causes cancer is not clear. The virus can replicate in hepatocellular carcinoma cells, but there is no evidence that DNA sequences are integrated into the host genome. Nearly all cases of liver cancer associated with hepatitis C occur in the presence of cirrhosis or severe chronic hepatitis. Indeed, progression from chronic active hepatitis to cirrhosis to hepatocellular carcinoma has been documented in prospective studies (Tanaka and Tsukuma, 1999). In the absence of genomic integration, it is possible that the emergence of a malignant clone reflects a multifactorial process of inflammation,

necrosis and cellular regeneration. Whether HCV contributes more directly to carcinogenesis is not known.

There are currently no practical strategies for the prevention of HCV infections, with the exception of blood-screening programmes (which have greatly reduced post-transfusion hepatitis). Although evidence remains scant, safe sexual practices and distribution of clean needles to intravenous drug users might reduce a small proportion of cases and perinatal transmission might be reduced with carefully managed deliveries and limitation of breast feeding. The efficacy of postexposure prophylaxis with immunoglobulin has yet to be confirmed. Efforts to develop a vaccine have been hampered by the inability to produce large amounts of immunogen *in vitro*, the fact that correlates of immunity are ill-defined and by the lack of an animal model. Although offering the greatest hope of prevention, vaccination against hepatitis C remains only a theoretical possibility.

Human Immunodeficiency Virus (HIV)

There is little evidence that HIV has a direct oncogenic effect in relation to the development of a specific cancer. Instead, it appears to facilitate the development, via its effects on the immune system, of a number of cancers, all of which are known (or thought) to be caused by other infectious agents. The abbreviation HIV is used throughout this chapter and refers specifically to HIV-1; reports on the association of cancers with HIV-2 are infrequent.

The human immunodeficiency virus was discovered in 1983 (Barré-Sinoussi *et al*., 1983) and firmly associated with AIDS in 1984 (Gallo *et al*., 1984). It is a human retrovirus, belonging to the lentivirus sub-family and is distinguished by its single-stranded RNA genome, which replicates via a DNA intermediate through the action of the enzyme ‘reverse transcriptase’ and integrates into the host chromosomal DNA. Although the first cases of HIV disease were reported as recently as 1981, the epidemic continues to escalate and the World Health Organisation (WHO) estimates that over 38 million young adults have been infected, the majority in sub-Saharan Africa. The three primary routes of transmission – sexual intercourse, blood contact and from mother to infant – were proposed on the basis of the epidemiology of AIDS, even before the identification of HIV. Of those, heterosexual transmission accounts for over 80% of new infections worldwide, while the importance of contact with infected blood products is declining since the introduction of routine screening procedures in blood banks.

The immunosuppression resulting from infection with HIV is causally associated with Kaposi sarcoma and non-Hodgkin lymphoma and, in the light of data emerging from sub-Saharan Africa, with squamous cell carcinoma of the conjunctiva. Recent evidence for two other cancers, Hodgkin disease and leiomyosarcoma in children, also suggests a definite increase in risk associated with HIV infection. The scale of the excess risk of these cancers in HIV-infected compared with uninfected individuals tends to be very large, 10-fold or more. However, people with HIV infection do not experience large excess risks of most cancers, including cancer of the uterine cervix and hepatocellular carcinoma, neither of which appear to be increased markedly in people with AIDS (IARC, 1996; Newton *et al*., 1999a).

The cancers identified as being HIV-associated have been linked (with varying degrees of certainty) to other infectious agents. Kaposi sarcoma is caused by the newly discovered human herpesvirus 8. Certain types of non-Hodgkin lymphoma have been linked to infection with the EBV and HHV-8, and conjunctival carcinoma has been linked to human papillomavirus (HPV) infection in some studies, but not others (IARC, 1995; Newton, 1999a). Thus, infection with the HIV appears to facilitate the development of certain cancers with an infectious aetiology. It is not clear why other cancers which are caused by infections, such as hepatocellular carcinoma, are not also AIDS associated.

Many of the increases in cancer risk found in people with HIV disease are similar to the findings in immunodeficient children and in transplant recipients, suggesting that it is the impairment of immune function that is the major factor leading to the appearance of these tumours (Beral and Newton, 1998). In addition, the risk of Kaposi sarcoma and non-Hodgkin lymphoma increases with increasing severity of immunosuppression, suggesting that this is the principle mechanism favouring their development. Also, both tumour types have been shown to regress following treatment with highly active antiretroviral therapy, which leads to improvements in immune function (IARC, 1996).

It has been estimated that, in 1990, there were about 52 000 additional cases of cancer that were a consequence of infection with the HIV (**Table 1**, Parkin *et al*., 1999). This conservative estimate is based on the known impact of infection with the virus on just two cancers with which it has been most clearly linked: Kaposi sarcoma and non-Hodgkin lymphoma. In populations with a high prevalence of HIV infection, the impact of the epidemic is clearly reflected in cancer registry statistics, although if Kaposi sarcoma and non-Hodgkin lymphoma are excluded, there is little evidence of an increase in the incidence of all other cancers combined. In Uganda and Zimbabwe, for example, the incidence of Kaposi sarcoma has increased between 10- and 20-fold in the era of AIDS, such that it is now the most common cancer in males in both countries and amongst the most common in females. There is evidence of a reduction in risk of both Kaposi sarcoma and non-Hodgkin lymphoma in those on antiretroviral therapy. However, in the developing world, where such treatment is prohibitively expensive and therefore not widely available, the incidence of HIV-associated cancers is likely to increase with the spread of the HIV epidemic.

In the absence of an effective vaccine, behavioural change is still the most important method of controlling the spread of HIV. Transmission of the virus in blood and blood products has been largely halted in developed countries, with the introduction of screening and education in combination with needle exchange programmes, which have been shown to be effective in reducing the spread of HIV among intravenous drug users. The bulk of transmission of HIV is sexual, however, and preventive activities include reducing the number of sexual partners, modifying the types of sexual contact and the use of condoms. Several behavioural interventions in high-risk populations have been tried, with variable results, but continued education remains a high priority.

BACTERIA

Helicobacter pylori

Helicobacter pylori is a spiral, flagellated, Gram-negative bacteria that colonizes the human gastrointestinal tract and lives beneath the mucus overlaying gastric epithelium. It causes gastritis in all infected people and although many cases remain asymptomatic, some result in gastric or duodenal ulceration. In a very small proportion of infected individuals, *H. pylori* may be involved in the aetiology of gastric adenocarcinomas and the much rarer primary gastric non-Hodgkin lymphoma – about a third of all gastric cancer cases have been attributed to *H. pylori* infection (IARC, 1994b; Danesh, 1999; Parkin *et al*., 1999).

It is estimated that about 50% of the world's population are chronically infected with *H. pylori*. The prevalence of infection is highest in developing countries and increases rapidly during the first two decades of life, such that 80–90% of the population may be infected by early adulthood. In most developed countries, the prevalence of infection is substantially lower at all ages, particularly in childhood. Everywhere, the prevalence of *H. pylori* is strongly correlated with markers of poverty and, indeed, has been decreasing in developing countries for decades, presumably because of improvements in living conditions. Transmission occurs from person to person, probably from mouth to mouth, faecal–orally or both.

In 1990, there were about 750 000 deaths attributable to gastric cancer, making it the fourteenth leading cause of death in the world and the second leading cause of cancer death (Murray and Lopez, 1997). Despite rapidly declining incidence rates in developed countries, gastric cancer is set to remain a major cause of death for many years, a result of population ageing, population growth in developing countries and a poor prognosis. Data from prospective sero-epidemiological studies suggest that infection with *H. pylori* results in a 2–4-fold increase in the risk of gastric cancer. With average seroprevalence rates of about 80% in developing countries and 50% in the developed world, it is estimated that about 340 000 new cases of gastric cancer each year are attributable to *H. pylori* infection – about 40% of the world total of gastric cancers (Parkin *et al*., 1999; **Table 1**). Similarly, about 4000 cases per year of gastric non-Hodgkin lymphoma (which represent about 3% of all gastric cancers) have been attributed to *H. pylori* infection.

The mechanisms by which *H. pylori* might increase the risk of gastric cancer are unclear. The bacteria cause lifelong inflammation, possibly leading to the production of mutagenic compounds, as well as loss of gastric acidity and epithelial cell proliferation. Any or all of the above might contribute to carcinogenesis. There is evidence that the development of cancer is preceded by progressive changes to the stomach mucosa, from inflammation, to atrophy and cellular proliferation, a process that is thought to be related to infection with *H. pylori*. This may be particularly true of the proinflammatory strains of *H. pylori* – those that possess the cytotoxin-associated gene A – because they markedly affect gastric cytokine levels and promote cell turnover, without a corresponding increase in apoptosis. Therefore, *cagA* strains of *H. pylori* might be expected to show a stronger association with gastric cancer, but the evidence for this remains scant.

Drug therapy consisting of two antibiotics in combination with either a bismuth preparation or an acid inhibitor for 14 days is effective in eradicating *H. pylori* in about 80% of cases. Given the high incidence of stomach cancer, the availability of screening tests and eradication regimens, but the relatively low progression rates to cancer in people with *H. pylori* infection, very large randomized trials are needed to establish the value of eradication for the prevention of gastric cancer. In order to achieve statistically reliable results, up to 100 000 people aged 60 years may have to be randomized to eradication therapy or placebo and followed for at least two decades (Danesh, 1999). However, the large-scale use of antimicrobial treatment is problematic – eradication has proved to be difficult in some developing countries and reinfection is common. Furthermore, extensive use of antibiotics may lead to the development of resistant strains (Coursaget and Muñoz, 1999).

In the future, immunization may be a better strategy for the prevention of *H. pylori*-associated diseases, particularly in developing countries. It has been demonstrated in mouse models that *H. pylori* vaccines not only can protect against infection, but may also induce regression of associated lesions. However, in models more relevant to humans, such as monkeys, the results have been disappointing. In phase I trials of a recombinant vaccine in humans, no adverse events were observed, but neither were there any changes in gastric bacterial density. Further work is required to identify appropriate target antigens and delivery systems and to understand the mechanism of protective immunity.

HELMINTHS

Schistosomes

Schistosomiasis (or 'bilharzia') is the generic term given to disease caused by the parasitic blood flukes of the genus *Schistosoma*, class Trematoda of the phylum Platyhelminthes (or flatworms). The genus contains 19 species of which three (*S. mansoni, S. haematobium* and *S. japonicum*) are of major importance to humans. Most infections are subclinical, but in those who develop severe disease the clinical features vary, depending on the species of schistosome. Cancer is an important but relatively rare outcome of infection; the bulk of morbidity and mortality is caused by nonmalignant conditions, such as renal or hepatic failure.

Infection occurs via exposure to water containing the larvae (cercariae). The worms mature in the veins that drain the bladder (*S. haematobium*) or intestine (other species). The adults can survive in the body for several years producing eggs, some of which leave the body in urine or faeces and hatch in water, freeing the miracidium larva. This stage infects certain types of fresh water snail, within which the parasites multiply asexually to produce free-swimming cercariae. These infect humans via skin penetration. Retained eggs elicit hypersensitivity reactions and cause disease of the urogenital system (*S. haematobium*) or of the liver and intestines (other species).

It is estimated that 200 million people in 74 countries are infected with schistosomes and over 600 million are at risk of infection. The geographical distribution of schistosomiasis corresponds to the distribution of susceptible snail hosts, which are present in many tropical and subtropical regions. *S. mansoni* is the most widespread species and is found in 54 countries in Africa, the eastern Mediterranean, South America and the Caribbean. *S. haematobium* has a similar distribution to *S. mansoni* in Africa and the eastern Mediterranean, where coinfection is relatively common, but does not occur in the Americas. *S. japonicum* is endemic in China, the Philippines and Indonesia (WHO, 1993; **Figure 2**). Within endemic areas, however, transmission tends to be highly focal (depending as it does on exposure to contaminated fresh water) and the prevalence and intensity of infection may vary between different communities, or even between households.

Contact with contaminated freshwater is the most important risk factor for infection with schistosomes. The level of contamination can depend on the size and distribution of the intermediate snail population, human population density and behaviour in relation to bodies of fresh water and, climatic and hydrological features (WHO, 1993). Infection is acquired cumulatively over a period of years and the severity is strongly related to the worm burden, or intensity of infection. The age distribution of all schistosome infections is similar, with a characteristic peak in both prevalence and intensity (as measured by active egg excretion) in the second decade of life and a gradual decline thereafter, probably resulting from changes in behaviour in relation to water exposure (WHO, 1993; IARC, 1994b).

Once inside a human host, worm pairs of the species *S. haematobium* reside primarily in the small venules that drain the bladder and ureters and are associated with disease of the urinary system. A causal association between infection with *S. haematobium* and bladder cancer was first postulated near the beginning of the twentieth century and has since been supported by clinical and experimental observations. Bladder cancers associated with schistosomiasis are primarily of the squamous cell type, rather than the transitional cell carcinomas that predominate in nonendemic parts of the world.

The link between infection with *S. haematobium* and bladder cancer is established first from the observation of elevated incidence rates in places where infection is endemic. Furthermore, the sex ratio of squamous cell carcinoma of the bladder in different countries correlates with the relative involvement of men and women in agricultural work (a risk factor for infection). Numerous case series and case-control studies confirm that *S. haematobium* is an important cause of bladder cancer in endemic countries. Infection may play a role in the aetiology of other cancers, in particular cancer of the uterine cervix, although this has yet to be established.

Worm pairs of the species *S. japonicum* reside in the venules that drain the gastrointestinal tract or in the liver, and chronic infection is associated with diseases of these organs. Data from ecological studies, case series and case-control and cohort studies indicate that in endemic areas, *S. japonicum* is probably an important cause of primary liver cancer and colorectal cancer (Chen *et al*., 1990; IARC, 1994b; Newton *et al*., 1999b). There are case reports of *S. japonicum* occurring in conjunction with a range of other cancers, but whether it plays an aetiological role is not clear.

The precise mechanism(s) by which schistosomiasis induces cancer is not known, although several possible explanations have been proposed (IARC, 1994b). These can be broadly categorized as involving (1) exogenous or endogenous agents which either induce DNA damage or have a tumour-promoting activity, (2) altered host metabolism, (3) pathological changes leading to increased cell proliferation and (4) altered immune responses. In relation to *S. haematobium* and bladder cancer, the first suggestions involved the effects of chronic inflammation and urinary retention. More generally, alteration of liver function by hepatic schistosomiasis leads to the production and excretion of potentially carcinogenic tryptophan metabolites, although the importance of these *in vivo* is not clear. In addition, chronic bacterial infection of the bladder can complicate schistosomal infestation, leading to the production of carcinogenic nitrosamines from precursors in urine.

S. mansoni
S. intercalatum

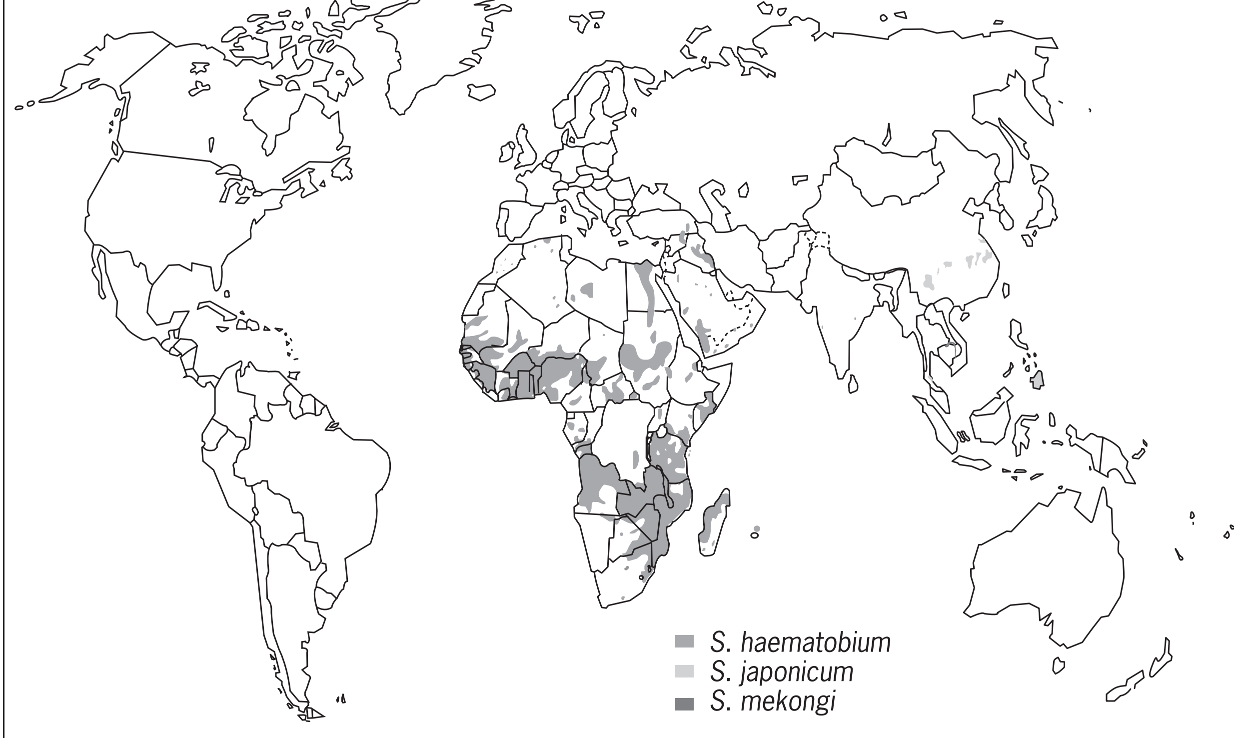


Figure 2 The geographical distribution of schistosomiasis (from WHO).

Safe, effective chemotherapy against all the schistosomes that infect humans has been available for more than two decades (WHO, 1993). The most versatile drug currently available is praziquantel, which is effective in a single oral dose, although it is relatively expensive if used extensively (approximately US $0.35 per treatment). However, treatment can result in resolution of infection, prevention or arrest of disease in heavily infected people and reversal of some manifestations of infection, such as haematuria.

Avoidance of contaminated water would prevent infection with schistosomes and is a relatively simple thing for occasional visitors to endemic areas to achieve. Control and prevention of infection in the community, however, where residents do not always have the luxury of avoiding contact with contaminated water, are complex. Many countries have initiated control programs involving a many pronged approach, including (1) the use of chemotherapy to remove adult worms, (2) elimination of the snail intermediate hosts by habitat modification or chemical attack, (3) changing human behaviour through health education and (4) providing safe water supplies and sanitation.

Use of these integrated control measures over many decades has led to the recent eradication of schistosomiasis in Japan, Tunisia and Monserrat (WHO, 1993). In China, 40 years of unremitting control measures have reduced the prevalence of infection by about 90%. Elsewhere in the World, including Brazil, Egypt, Iran, the Philippines and Venezuela, significant reductions in disease prevalence have been achieved. Even in places where the prevalence of infection has remained high, serious manifestations of disease are becoming less common with the use of effective chemotherapy, although declines in cancer incidence are not yet apparent (WHO, 1993). Despite this, the number of cases of schistosomiasis worldwide was estimated to be the same in 1993 as it was in 1984 (WHO, 1993). In endemic areas, populations (and hence the number of susceptible hosts) continue to grow. In addition, developments in water resource management, land use and irrigation have led to a spread of schistosomiasis to new areas. There is currently no vaccine available, although intensive efforts are being made to develop one, and so the use of complex, integrated control measures remains paramount.

Liver Flukes

Three species of food-borne liver flukes, *Opisthorchis viverrini*, *O. felineus* and *Clonorchis sinensis*, of the class Trematoda, are pathologically important to humans. They establish chronic infection within the intrahepatic bile ducts and occasionally in the pancreas and gall-bladder and have been associated with diseases of these organs. Infection is acquired by eating raw or undercooked freshwater fish, which contain the infective stage (metacercaria). The flukes migrate to the biliary tree via the ampulla of Vater and mature in the intrahepatic bile ducts, producing eggs, which are excreted in faeces. If the eggs reach fresh water and are consumed by an appropriate species of snail, they hatch, undergo asexual reproduction to produce free-living larvae, which can infect freshwater fish and become encysted metacercariae (IARC, 1994b; Vatanasapt *et al*., 1999).

O. viverrini is common in north-east Thailand and Laos, where it is estimated that up to 9 million people are infected (about one-third of the population). About 1.5 million people are infected with *O. felineus*, mainly in central Russia, and 7 million are infected with *C. sinensis*, in Korea, China, Macao and Vietnam (**Figure 3**). The distribution of human infection depends not just on that of the flukes, but also on the habit of eating raw freshwater fish. Infection usually occurs in the first decade of life and men are often more affected than women.

Cholangiocarcinoma has long been recognized as a serious complication of liver fluke infection. This tumour is very rare throughout most of the world, but in endemic areas the number of cases of cholangiocarcinoma is usually higher than that of hepatocellular carcinoma. Evidence from ecological studies relating the prevalence of antibodies to liver fluke infection (or of eggs in faeces) to the incidence of cholangiocarcinoma suggest a causal link. Case-control studies indicate that infected individuals have about a fivefold increase in the risk of cancer compared with the uninfected, a finding that is supported by data from animal studies.

The pathological changes associated with acute liver fluke infection include oedema, desquamation and acute inflammation in the bile ducts. Chronic infection is characterized by goblet cell metaplasia, adenomatous hyperplasia and thickening of the walls. It is thought that cholangiocarcinoma arises from this pre-existing damage, as part of a progressive process. The mechanisms by which this might occur are poorly understood.

The antihelminthic drug praziquantel is currently the drug of choice for the treatment of liver fluke infection, both for the individual patient and for community-based treatment programmes. The best way to reduce the incidence of cholangiocarcinoma is to control liver flukes. Eradication programmes need continuous and intensive health education, together with drug treatment for existing infections. Mass treatment without health education is unlikely to be effective. Health education should focus on the need to cook fish and fish products properly. As yet, eradication programmes have had little effect on the incidence of cholangiocarcinoma and a vaccine is not available.

THE FUTURE

Given that the discovery of an infectious cause of cancer (or any other chronic disease) has such important implications for prevention, how does one identify and confirm a causal association? First, this requires an insight into which diseases may be linked to infection. Certain epidemiological features of a cancer may offer clues as to an infectious aetiology. These include a high incidence of disease in people who are prone to infections (such as the immunosuppressed), large variations in incidence by geographical region or other patterns of clustering, apparent improvements with antimicrobial treatments and a strong correlation with markers of poverty. So what

Figure 3 The geographical distribution of liver fluke infections (from WHO).

cancers might next be linked to infections? It is dangerous to prophesy, but there is considerable speculation about leukaemias and lymphomas, squamous cell skin cancers and even breast cancer. Second, the relevant infection must be identified and isolated from tumour tissue or diseased individuals. Finally, causality must be established. The relative risk of a specific cancer in an infected individual is usually very high, but it is harder to identify large excess risks in infected individuals when the cancer is a rare outcome of a common, or even ubiquitous, infection. In such instances, establishing causality may depend in part, for example, on identifying abnormal immunological responses to infection in diseased individuals, compared with those who are without disease (e.g. differences in antibody titre to infection).

As for prevention of cancers caused by infections, it is clear that vaccination programmes, although ultimately cost effective, require a long-term commitment, involving substantial investment of time, money and effort. In the meantime, an understanding of the biology and epidemiology of the relevant infections is essential to reduce the associated cancer burden in other ways. In certain circumstances, many infections can be avoided by behaviour modification or screening of blood products, for example. In addition, the risk of transmission of an infection from a pregnant mother to her child may be reduced with changes in breast-feeding behaviour or the use of anti-viral therapies. Perhaps one of the most exciting possibilities involves the introduction of HPV testing within the framework of the cervical screening programme, as a means of identifying women at high risk of disease, for intensive follow-up.

All of the above makes a strong case for a coordinated effort by clinicians, molecular biologists and epidemiologists to search for new – as well as known – infections in human tissues and to study their possible associations with disease. Research into a problem, however, is not the whole story. Understanding the causes of cancer is often difficult – prevention requires something more: a commitment to invest in public health.

REFERENCES

Barré-Sinoussi, F., *et al*. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for aquired immunodeficiency syndrome (AIDS). *Science*, **220**, 868–871.

Beral, V. and Newton, R. (1998). Overview of the epidemiology of immunodeficiency-associated cancers. *Monographs of the National Cancer Institute*, **23**, 1–6.

Boshoff, C. (1999). Kaposi's sarcoma associated herpesvirus. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 157–190.

Chang, M. H., *et al*. (1997). Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children. *New England Journal of Medicine*, **336**, 1855–1859.

Chen, J., *et al*. (1990). *Diet, Life-style and Mortality in China: A Study of the Characteristics of 65 Chinese Counties* (Oxford University Press, Oxford, Cornell University Press, Ithaca, NY and People's Medical Publishing House, Beijing).

Coursaget, P. and Muñoz, N. (1999). Vaccination against infectious agents associated with human cancer. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 355–382.

Cuzick, J. (1999). Screening. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 345–354.

Danesh, J. (1999). Is *Helicobacter pylori* infection a cause of gastric neoplasia? In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 263–272.

de Thé, G., *et al*. (1978). Epidemiological evidence for a causal relationship between the Epstein–Barr virus and Burkitt's lymphoma from a Ugandan prospective study. *Nature*, **274**, 756–761.

Gallo, R. C., *et al*. (1984). Frequent detection and isolation of cytopathic retrovirus (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, **224**, 500–503.

Gambia Hepatitis Study Group (1987). The Gambia Hepatitis Intervention Study. *Cancer Research*, **47**, 5782–5787.

Herrero, R. and Muñoz, N. (1999). Human papillomavirus and cancer. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 75–98.

IARC (1992). *The Epidemiology of Cervical Cancer and Human Papillomavirus*. IARC Scientific Publications, Vol. 119 (International Agency for Research on Cancer, Lyon).

IARC (1994a). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 59, Hepatitis Viruses* (International Agency for Research on Cancer, Lyon).

IARC (1994b). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 61, Schistosomes, Liver Flukes and Helicobacter pylori* (International Agency for Research on Cancer, Lyon).

IARC (1995). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 64, Human Papillomaviruses* (International Agency for Research on Cancer, Lyon).

IARC (1996). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 67, Human Immunodeficiency Viruses and Human T-cell Lymphotropic Viruses* (International Agency for Research on Cancer, Lyon).

IARC (1997). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 67, Epstein–Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8* (International Agency for Research on Cancer, Lyon).

McGregor, J. M. and Proby, C. M. (1996). The role of papillomaviruses in human non-melanoma skin cancer. In: Leigh, I. M.,

et al. (eds), *Cancer Surveys, Vol. 26, Skin Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 219–236.

Murray, C. J. L. and Lopez, A. D. (1997). Mortality by cause for eight regions of the world: global burden of disease study. *Lancet*, **349**, 1269–1276.

Newton, R., *et al*. (1999a). Human immunodeficiency virus infection and cancer. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 237–262.

Newton, R., *et al*. (1999b). Schistosomes and human cancer. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 291–312.

Parkin, D. M., *et al*. (1999). The global health burden of infection associated cancers. In: Newton, R., *et al*. (eds), *Cancer Surveys, Volume 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 5–34.

Phillips, A. C. and Vousden, K. H. (1999). Human papillomavirus and cancer: the viral transforming genes. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 55–74.

Tanaka, H. and Tsukuma, H. (1999). Hepatitis C virus. In: Newton, R., *et al*. (eds), Cancer Surveys, Vol. 33, Infections and Human Cancer (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 55–74.

Vatanasapt, V., *et al*. (1999). Liver flukes and liver cancer. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 313–344.

Walbloomers, J. M., *et al*. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *Journal of Pathology*, **189**, 12–19.

WHO (1993). *The Control of Schistosomiasis*. Second Report of the WHO Expert Committee. WHO Technical Report Series, No. 830 (World Health Organization, Geneva).

Wild, C. P. and Hall, A. J. (1999). Hepatitis B virus and liver cancer: unanswered questions. In: Newton, R., *et al*. (eds), *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp. 35–54.

FURTHER READING

Newton, R., *et al*. (eds) (1999). *Cancer Surveys, Vol. 33, Infections and Human Cancer* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Parsonnet, J. (ed.) (1999). *Microbes and Malignancy: Infection as a Cause of Human Cancer* (Oxford University Press, Oxford).

Chapter 25

Short-term Testing for Genotoxicity

David Gatehouse
GlaxoWellcome Research and Development Limited, Ware, UK (Present address: Concave Laboratories Ltd, Harrogate, UK)

CONTENTS

- Introduction
- Primary Test Systems
- Supplementary Test Systems
- Interpretation of Results
- Future Developments

INTRODUCTION

It has been known for several hundred years that exposure to particular chemicals or complex mixtures can lead to cancer in later life. There has been accumulating evidence that many cancers can arise from damage to DNA and the resulting mutations. This has been discussed in detail in the preceding chapters. As a consequence of this, it has become necessary to determine whether widely used chemicals or potentially useful new chemicals possess the ability to damage DNA. Data concerning the genotoxicity of a new chemical have become part of the basic toxicological information package. They are needed for decision making and to reduce risks that might otherwise be unforeseen.

The field of genetic toxicology began in the 1960s when several seminal conferences were held focusing on chemical mutagens and in particular their effects on germ cells and the risk to future generations. Although germ cell risk was the initial concern, this was broadened in the 1970s when evidence relating genotoxicity and carcinogenicity began to accumulate. This was further supported by the use of *in vitro* metabolic activation systems capable of producing electrophilic metabolites, and the fact that early analysis of rodent carcinogens and noncarcinogens suggested that almost all carcinogens were also genotoxic. This view has now been modified, since it is clear that nongenotoxic carcinogens also exist, as discussed in an earlier chapter (see the chapter *Non-Genotoxic Causes of Cancer*). From this time onwards, various national expert committees were formed to advise governments on the type of approach that should be taken to screen new chemicals for carcinogenic risk (and any potential heritable effects). Consequently, numerous guidelines have been prepared over the past 20 years describing the tests which should be used to investigate the genotoxicity of chemicals. It is not the intention of this chapter to give an exhaustive list of these guidelines, except to mention that a harmonized approach to the genotoxicity testing of drugs has recently been introduced as a result of the International Conference on Harmonisation (ICH) programme. In this process, areas of disharmony were identified and differences in regulation with respect to genotoxicity were discussed, resulting in the creation of two guidelines (referenced at the end of this chapter). At the same time as the ICH process was occurring, the Organisation for Economic Cooperation and Development (OECD) also updated a number of its genotoxicity guidelines. Both processes influenced each other resulting in similar recommendations. A summary of the testing strategy recommended by the ICH is given in **Figure 1**.

In this chapter, the individual test systems that are required for genotoxicity screening will be described together with a discussion on how the results obtained should be interpreted. As well as the primary test systems, there will also be a description of the supplementary assays that may be required when investigating positive effects. It is not intended to give detailed guidance on the performance of these tests. For this information, the reader is directed to the Further Reading list at the end of the chapter. The chapter concludes with a brief overview of some of the new developments in the field.

1. A test for gene mutation in bacteria
2. An *in vitro* test with cytogenetic evaluation of chromosomal damage with mammalian cells or an *in vitro* mouse lymphoma tk assay
3. An *in vivo* test for chromosomal damage using rodent haematopoietic cells

Figure 1 Recommended ICH genotoxicity test battery.

PRIMARY TEST SYSTEMS

In Vitro Metabolic Activation

Before describing the individual *in vitro* test systems, it is necessary to mention briefly a factor of critical importance in genotoxicity screening, namely the need to include some form of *in vitro* metabolizing system. This is because most of the indicator cells (bacteria and mammalian cells) possess a very limited capacity for endogenous metabolism of xenobiotics. Many carcinogens and mutagens are unable to interact with DNA unless they have undergone some degree of metabolism (see also the chapter *Mechanisms of Chemical Carcinogenesis*). To improve the ability of the test systems to detect as many authentic *in vivo* mutagens and carcinogens as possible, extracts of mammalian liver (usually rat) are incorporated. The liver is a rich source of mixed-function oxygenases capable of converting carcinogens to reactive electrophiles. Crude homogenate such as the 9000 *g* supernatant (S9 fraction) is used, which is composed of free endoplasmic reticulum, microsomes, soluble enzymes and some cofactors. The oxygenases require the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH), which is normally generated *in situ* by the action of glucose-6-phosphate dehydrogenase on glucose-6-phosphate and reducing NADP, both of which are normally supplied as cofactors. Normal uninduced S9 preparations are of limited value for screening as they are deficient in particular enzyme activities. In addition, species and tissue differences are most divergent in such preparations. These problems are reduced when enzyme inducers are used, and most commonly preparations are made from rat livers after enzyme induction with Aroclor 1254, which is a mixture of polychlorinated biphenyls. Concern about the toxicity, carcinogenicity and persistence of this material in the environment has led to the introduction of alternatives, such as a combination of phenobarbitone and β-naphthoflavone. This combination induces a similar range of mono-oxygenases and has been recommended as a safer alternative to Aroclor (Elliott *et al*., 1992).

It should be noted that this system is only a first approximation to the complex metabolic processes that occur *in vivo*, and in particular there is little account taken of the phase II detoxification reactions. Such factors should be considered when interpreting positive *in vitro* results which are only seen in the presence of S9 mix.

In Vitro Tests for Gene Mutation in Bacteria

The most widely used assays for detecting chemically induced gene mutations are those employing bacteria. These assays feature in all test batteries for genotoxicity as it is relatively straight forward to use them as a sensitive indirect indicator of DNA damage. Bacteria can be grown in large numbers overnight, permitting the detection of rare mutational events. The extensive knowledge of bacterial genetics that was obtained during the twentieth century allowed the construction of special strains of bacteria with exquisite sensitivity to a variety of genotoxins. An offshoot of the studies on genes concerned with amino acid biosynthesis led to the development of *Escherichia coli* and *Salmonella typhimurium* strains with relatively well defined mutations in known genes. The most commonly used bacteria are the *S. typhimurium* strains which contain defined mutations in the histidine operon. These were developed by Bruce Ames, and form the basis of the 'reverse' mutation assays (Ames *et al*., 1971). In these assays, bacteria which are already mutant at the histidine locus are treated with a range of concentrations of test chemical to determine whether the compound can induce a second mutation that directly reverses or suppresses the original mutations. Thus, for the *S. typhimurium* strains which are histidine auxotrophs, the original mutation resulted in the loss of ability to grow in the absence of histidine. The second mutation (induced by the chemical) restores prototrophy, i.e. the affected cell is now able to grow in the absence of histidine, if provided with inorganic salts and a carbon source. This simple concept underlines the great strength of these assays for it provides enormous selective power which can identify a small number of the chosen mutants from a population of millions of unmutated cells and cells mutated in other genes. Each of the *S. typhimurium* strains contains one of a number of possible mutations in the histidine operon, and each can be reverted by either base-change or frameshift mutations. The genotype of the most commonly used strains is shown in **Table 1**, together with the types of reversion events that each strain detects.

In order to make the bacteria more sensitive to mutation by chemical agents, several additional traits have been introduced. Ames and colleagues realized that many carcinogens (or their metabolites) are large molecules that are often unable to cross the protective cell wall of the bacteria. Wild-type cells produce a lipopolysaccharide that acts as a barrier to bulky hydrophobic molecules. Consequently, an *rfa* mutation was introduced into the *Salmonella* strains, which resulted in defective lipopolysaccharide and increased permeability.

Bacteria possess several major DNA repair pathways that appear to be error-free. The test strains were constructed, therefore, with a deletion removing the *uvrB* gene. This codes for the first enzyme in the error-free excision repair pathway, and so gene deletion renders the strains excision repair deficient, thus increasing their sensitivity to many genotoxins by several orders of magnitude. Lastly, some of the bacterial strains do not appear to possess classical error-prone repair as found in other members of the Enterobacteria such as *E. coli*. This results from a deficiency in *umuD* activity. This deficiency is overcome by insertion of a plasmid containing *umuDC* genes. Plasmid

Table 1 Genotype of commonly used strains of *S. typhimurium* LT2 and reversion events detected by these strains

Strain	Histidine mutation	Full genotype[a]	Reversion events
TA1535	hisG46	rfa Dgal chlD bio *uvr*B	Subset of base pair substitution events,
TA100	hisG46	rfa Dgal chlD bio *uvr*B (pKM101)	extragenic suppressors
TA1538	hisD3052	rfa Dgal chlD bio *uvr*B	Frameshifts
TA98	hisD3052	rfa Dgal chlD bio *uvr*B (pKM101)	Frameshifts
TA1537	hisC3076	rfa Dgal chlD bio *uvr*B	Frameshifts
TA97	hisD6610	his O $_{1242}$rfa Dgal chlD bio *uvr*B (pKM101)	Frameshifts
TA102	hisG428	his D (G) $_{8476}$rfa galE (pAQ1) (pKM101)	All possible transitions and transversions small deletions and extragenic suppressors

[a]rfa = Deep rough; galE = UDP galactose 4-epimerase; ChlD = nitrate reductase (resistance to chlorate); bio = biotin; *uvr*B = UV endonuclease component B; D = deletion of genes following this symbol; pAQ1 = a plasmid containing the his G_{428} gene; pKM101 = a plasmid carrying the *uvr*A and B genes that enhance error-prone repair.

pKm101 is the most useful (Mortelmans and Dousman, 1986) conferring on the bacteria sensitivity to mutation without a concomitant increase in sensitivity to the lethal effects of test compounds. Further sensitivity is gained by the fact that the initial mutation responsible for the histidine growth requirement is situated at a site within the gene that is particularly sensitive to reversion by specific classes of genotoxin (i.e. hotspots). The incorporation of strain TA102 into the test battery has been proposed, as the target mutation has an AT base pair at the critical site. This allows the detection of genotoxins not detected by the usual battery of *S. typhimurium* strains that possess mutations exclusively at GC base pairs. As an alternative many guidelines recommend the use of the *E. coli* WP2 trpE strains which contain a terminating ochre mutation in the *trpE* gene. The ochre mutation involves an AT base pair, and so reverse mutation can take place at the original site of mutation or in the relevant tRNA loci. A combination of *E. coli* WP2 *trpE* (pKm101) and *E. coli* WP2 *trpE uvrA* (pKm101) can be used as alternatives to *S. typhimurium* TA102 for the detection of point mutations at AT sites.

Consequently in the recently published ICH Guidelines (ICH, 1998), the following base set of bacterial test strains were recommended:

S. typhimurium: TA98, TA100, TA1535
S. typhimurium: TA1537, or TA97, or TA97a
S. typhimurium: TA102, or *E. coli* WP2 *uvrA* or *E. coli* WP2 *uvrA* (pKm101)

The use of the repair-proficient *E. coli* strain WP2 (pKm101) allows the detection of cross-linking agents that require an intact excision repair pathway to generate mutations and this strain may also be selected.

It is not the intention of this chapter to give detailed instruction on the performance of these assays. For this information, the reader is directed to other detailed references (Gatehouse *et al*., 1994; Tweats and Gatehouse, 1999). In general, the most widely used protocol is the 'plate incorporation assay.' In this method, the bacterial strain, test material and an *in vitro* metabolic activation system (S9 mix) are added to a small volume of molten agar containing a trace of histidine and biotin. The mixture is poured across the surface of a basal agar plate and allowed to set prior to incubation at 37 °C for 48–72 h. The trace of histidine allows the growth of the auxotrophic bacteria in the presence of the test compound and/or any *in vitro* metabolites. The period of several cell divisions is essential to allow fixation of any premutagenic lesions that have been induced in the bacterial DNA. Exhaustion of the histidine halts growth of the auxotrophic cells. Only those cells that have been reverted to histidine independence will continue to grow and form discrete visible colonies. The growth of nonreverted cells forms a visible background lawn, the thinning of which can be used as a non-quantitative indication of drug related toxicity. Revertant colonies can be counted manually or by use of an Image analyser. Untreated (vehicle) and suitable positive controls are included. The test concentration range is determined by performance of a preliminary toxicity test, whilst for nontoxic compounds a maximum concentration of 5 mg per plate is recommended. Some genotoxins are poorly detected using the plate incorporation procedure, particularly those that are metabolized to short-lived reactive electrophiles (e.g. aliphatic nitrosamines). In these cases, a preincubation procedure should be used in which bacteria, test compound and S9 mix are incubated together in a small volume at 37 °C for 30–60 min prior to agar addition. This maximizes exposure to the reactive species and limits nonspecific binding to agar.

Several statistical approaches have been applied to the results of these assays. In general, positive results should be statistically significant, dose related and reproducible. Bacterial mutation tests have been subjected to several large-scale trials over the years (e.g. Tennant *et al*., 1987). These studies were primarily concerned with assessing the correlation between results obtained in the assays and the

carcinogenic activity of chemicals. Most of the studies suggest that there is a good qualitative relationship between genotoxicity in the *Salmonella* assay and carcinogenicity for many, although not all, chemical classes. This figure varies between a sensitivity of 60 and 90% dependent upon chemical class. The bacterial assays seem to be particularly efficient in detecting trans-species, multiorgan animal carcinogens (Ashby and Tennant, 1988).

In Vitro Tests for Gene Mutation in Mammalian Cells

Although the prokaryotic systems described above are extremely versatile, rapid and mostly accurate in detecting genotoxins, the intrinsic differences between prokaryotic and eukaryotic cells in terms of genome structure and organization necessitate the use of mammalian test systems within any screening battery designed to detect the widest spectrum of genotoxins. A variety of *in vivo* mutation systems have been described in the literature but only a few have been defined adequately for quantitative studies (Cole *et al*., 1990). Unlike the bacterial reverse mutation systems, these tests are based upon the detection of forward mutations. A defined large number of cells are treated with the test agent and then, after a set interval (termed the expression period), the cells are exposed to a selective toxic agent, so only those cells that have mutated can survive. Usually mutation is measured in genes located on the X chromosome in male cells (XY) where only one copy of the target gene is present, or in cells (called heterozygotes) where two copies of the gene are present but only one copy is active, as the other has been inactivated through mutation or deletion. The most common systems make use of genes which are not essential for cell survival but allow the cell to salvage nucleic acid breakdown products (nucleotides) from the culture medium for reuse in metabolism. Toxic nucleotides placed in the culture medium will be transported into normal (nonmutated) cells that consequently die. However, loss of the salvage enzyme through genotoxic damage (mutation, chromosome deletions or rearrangements) of the corresponding gene will render the cell resistant to the toxic nucleotide and so it will survive. The surviving mutant cells can be detected by colony formation in tissue culture plates. The two most popular genes for measuring mutation *in vitro* are those coding for hypoxanthine guanine phosphoribosyl transferase (*Hgprt*) and thymidine kinase (*tk*). The former is located on the X chromosome in both human and Chinese hamster cells and loss of activity in this gene can be measured by resistance to the antimetabolite 6-thioguanine. The TK gene is located on chromosome 11 in mouse cells and on chromosome 17 in humans. Loss of activity in this gene can be measured by resistance to the toxic chemical trifluorothymidine (Clive *et al*., 1987). Three cell lines have been used most extensively for routine *in vitro* mutation assays. Two are Chinese hamster cell lines (V79 and CHO) and one is a heterozygous mouse line, mouse lymphoma (L5178Y). The Chinese hamster cells have been used extensively over the past 15–20 years, but more recently there has been a significant move towards the use of the heterozygous mouse lymphoma L5178Y cell line because of the variety of genetic events detected in this system. In fact, recently published ICH guidelines recommend this system as the method of choice when evaluating the potential genotoxicity of new drugs.

The theoretical basis for the mouse lymphoma assay is shown in **Figure 2**. Two main protocols have been devised for performing the assay, plating the cells in soft agar or a fluctuation test approach. Mitchell *et al*. (1997) published a full evaluation of this assay. In these tests, ideally the highest test concentration should produce at least 80% toxicity (i.e. no more than 20% survival).

At least two types of colony are obtained when mutations at the *tk* locus are selected, large colonies which grow at the normal rate and slow-growing small colonies. Initial molecular analysis of these colonies indicated that a high percentage of small colony mutants have a wide variety of visible chromosome 11b aberrations, whereas large colonies do not. However, recent chromosome painting analysis of colonies indicates that this initial generalized premise on colony size may be oversimplistic. What is clear is that a wide range of mutations and genetic events can be detected by the L5178Y system, including both point mutations and chromosomal damage.

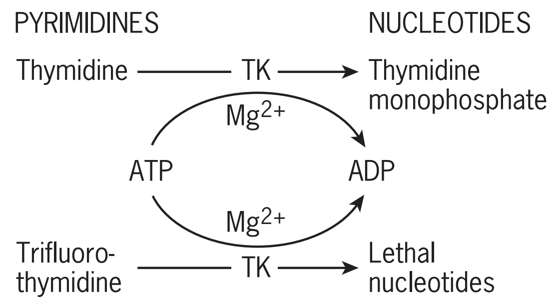
Recently, the *p53* status of a number of L5178Y cell lines has been investigated, and it has been found that the cells contain two mutant *p53* alleles resulting in a

The gene coding for thymidine kinase (TK) is on mouse chromosomes number 11

The mouse lymphoma L5178Y line is $TK^{+/-}$,
i.e. one gene copy is already inactivated by mutation

∴ Each cell contains only one functional gene

∴ One mutational event can render the cell $TK^{-/-}$



$TK^{+/-}$ cells are killed by trifluorothymidine
$TK^{-/-}$ cells survive (thymidine monophosphate can be synthesized *de novo* via thymidylate synthetase, etc.)

Figure 2 The theoretical basis of the L5178Y $TK^{+/-}$ mouse lymphoma assay.

dysfunctional p53 protein (Storer *et al*., 1997). As this tumour-suppressor protein is so important in regulating cellular responses to DNA damage, this may account for the sensitivity of these cells to genotoxins. It has been suggested that the *p53* status of the cells makes them even more appropriate for use in screens for genotoxic carcinogens, as the development of cancer is often associated with mutant p53 protein. A system that contains a component evaluating a chemicals' ability to induce additional mutations in *p53*-deficient cells may provide a more appropriate model for the human situation.

In Vitro Tests for Chromosome Damage in Mammalian Cells

An alternative method to measuring mutation induction within mammalian genes involves the examination of mammalian chromosomes microscopically for the presence of visible damage.

In simplest terms these tests generally involve exposure of cultured cells to the test material in the presence and absence of a metabolic activation system. The cells are then harvested at one or more time intervals after treatment. Before fixation the cells are exposed to a mitotic poison to arrest them in metaphase and then to hypotonic solution to swell them. They are fixed, metaphase spreads are prepared and after staining (usually with Giemsa) they are analysed microscopically. Gross damage to the chromosomes such as terminal deletions, breaks and exchanges are recorded. Much of this damage is lethal to the cell during the cell cycle following induction of the damage. However, such changes are used as indicators of the presence of nonlethal more subtle changes (e.g. reciprocal translocations and small deletions) which are difficult to observe microscopically, but which may have important consequences in both somatic and germ cells.

Established cells (e.g. Chinese hamster fibroblasts) or primary cells (e.g. human peripheral lymphocytes) can be used. Both have advantages and disadvantages. The Chinese hamster cells contain a small number of chromosomes facilitating rapid analysis, but can contain karyotype abnormalities. The human lymphocytes, on the other hand, are karyotypically stable, but donor variation in responsiveness to some mutagens may be an issue.

A detailed discussion of the test procedures employed was given by Scott *et al*. (1990). Detectable levels of chromosome aberrations are often found only at doses which induce some evidence of cytotoxicity, and consequently the recommended protocols require that the maximum test concentration should induce $>50\%$ reduction in cell number or culture confluency for cell lines, or an inhibition of mitotic index by $>50\%$ for lymphocytes. However, there are growing concerns as to the relevance of genotoxic effects which are found only at highly cytotoxic concentrations.

In Vivo Tests for Chromosome Damage in Rodents

Although there are a number of methods available for measuring chromosome damage *in vivo*, by far the most popular and extensively validated technique is the micronucleus test. In this assay damage is induced in the immature erythroblast population in the bone marrow. Chromosome breakage is detected in the form of centric and acentric fragments which form micronuclei observable microscopically within these cells after the main nucleus has been expelled. This test is also sensitive to numerical changes, thus allowing the detection of chemicals that cause whole chromosome loss (aneuploidy) in the absence of clastogenic activity. The basis for the assay is given in **Figure 3**. When the immature erythroblast matures into a polychromatic erythrocyte, the micronucleus, the formation of which results from either chromosome loss during cell division or from chromosome breakage, is not extruded with the nucleus. Young polychromatic erythrocytes (usually 2000) are examined in the bone marrow of rodents that have been previously exposed to the test material by an appropriate route. Analysis is carried out at two time points (usually 24 and 48 h) after a single exposure. As an alternative, multiple dosing can be used, usually three or more daily doses, in which case the bone marrow is analysed 24 h after the last dose. In addition to micronucleus incidence, the ratio of polychromatic to normochromatic erythrocytes is also assessed as an indication of any compound-related cytotoxic or cytostatic effect within the marrow. Male animals are sufficient for use in the test, unless there are obvious differences in toxicity or metabolism between male and female rodents. In these studies the test substance is evaluated at the maximum tolerated dose (i.e. one which produces some signs of toxicity such as hypoactivity, ataxia, ptosis, etc.). The highest dose to use when toxicity is not evident ranges from 2 to 5 $g\,kg^{-1}$ according to different regulatory guidelines. A more detailed discussion of the study design and test procedure was given by Hayashi *et al*. (2000).

Micronuclei can also be detected in reticulocytes circulating within peripheral blood, and in future it is likely that such studies, using either bone marrow or blood as the target tissue, will be incorporated into standard rodent toxicology studies, rather than be performed as stand-alone experiments. Also, it is possible that the use of flow cytometric techniques will replace microscopic analysis as a means of generating results more quickly and accurately (Dertinger *et al*., 1997).

In Vivo Test for Unscheduled DNA Synthesis in Rat Liver

It is recognized that there is a need for an assay that can complement the micronucleus assay, by using tissue other

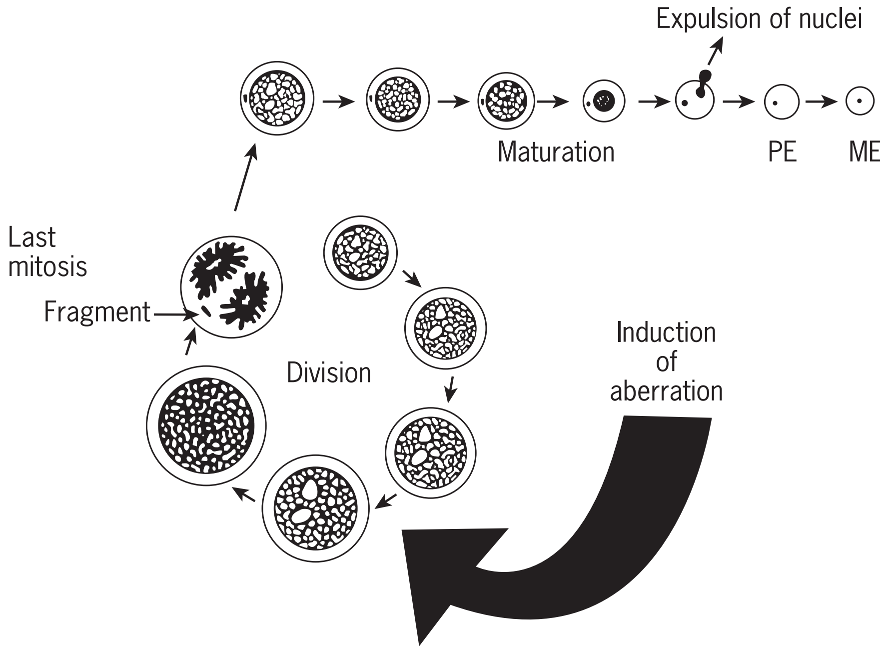


Figure 3 The theoretical basis for the formation of micronuclei in rodent bone marrow. PE = polychromatic erythrocyte; ME = mature erythrocyte.

than the bone marrow. This is because in some instances genotoxins will fail to be detected in the bone marrow owing either to poor drug distribution or other organospecific effects. It is important to note that such an assay is only likely to be needed if there is clear, unequivocal evidence of genotoxic effects in one or more of the *in vitro* tests and the *in vivo* micronucleus test is negative. At the present time the preferred approach is to investigate unscheduled DNA synthesis (UDS) in rat liver *in vivo*. The liver, as the major organ of xenobiotic metabolism, is an appropriate target and this assay fulfils a useful complementary role when used in conjunction with the micronucleus test, allowing the detection of a wider spectrum of genotoxins (**Table 2**). However, new test systems that allow the investigation of genotoxic effects in other tissues are now becoming increasingly available and will be discussed in the next section.

Table 2 *In vivo* genotoxicity results for 12 *in vitro* genotoxins and *in vivo* carcinogens

Chemical	Mouse micronucleus test	Rat liver UDS
Dimethylnitrosamine	−	+
Diethylnitrosamine	−	+
2,4-Dinitrotoluene	−	+
3′-Methyldimethylaminoazobenzene (DAB)	−	+
6BT (DAB analogue)	−	+
Dimethylhydrazine	−	+
Cyclophosphamide	+	−
Benzo[*a*]pyrene	+	−
Hexamethylphosphoramide	+	−
2-Acetylaminofluorene	+	+
Benzidine	+	+
N-Methyl-*N*-nitronitrosoguanidine	−	−

UDS assays quantify the resultant excision repair of DNA following a permanent change such as covalent binding of an activated mutagen or a reactive chemical species generated intracellularly. Cells undergoing such repair synthesize DNA at stages of the cell cycle other than S-phase, where normal replicative (scheduled) DNA synthesis takes place, hence the term 'unscheduled DNA synthesis.' This technique is potentially highly sensitive because the whole genome is theoretically a target for chemical reaction.

A detailed description of the methodology was given by Madle *et al*. (1994). Briefly, rodents, usually rats, are treated with the suspect chemical by an appropriate route. For nontoxic compounds an upper limit of $2\,g\,kg^{-1}$ is recommended, whereas for compounds eliciting toxicity a dose approaching the maximum tolerated dose should be used. A single-dose regimen is the most common procedure, and two sampling times are employed (12–16 and 2–4 h after treatment). At these times, viable hepatocyte populations are prepared by perfusing the livers with collagenase. Quantification of UDS is achieved by determining the uptake of the radiolabelled nucleotide [^{3}H]thymidine during DNA repair. A microautoradiographic procedure is the most favoured method in which slide preparations are made, on which the cells are fixed, developed autoradiographically

and then stained prior to analysis with the aid of an image analyser. Cells undergoing repair are identified by increases in the number of silver grains overlying the nuclei. S-phase cells exhibit extremely high numbers of nuclear silver grains and are excluded from analysis. Hepatocytes from untreated control animals are included, and the net nuclear grain counts are compared. For a positive response there should be a net nuclear grain count for at least one dose and sampling time which exceeds a laboratory-specific threshold justified on the basis of historical control data.

SUPPLEMENTARY TEST SYSTEMS

The assays outlined in the previous section form the basis of most test batteries within a range of regulatory guidelines concerned with the testing of drugs, chemicals, food additives, etc., for genotoxicity. However, the current testing strategies are still deficient in a number of respects. Regulatory guidelines for genotoxicity screening are a compromise. Although there are accepted tests for measuring chromosome damage *in vitro* and *in vivo*, at present there is no validated test in widespread use to detect point mutation *in vivo*. Studies have shown that many carcinogens induce tumours in specific tissues, hence there is also a need for an *in vivo* genotoxicity test that could be applied to any tissue. In addition, the need for a specific assay to detect genome mutation (i.e. chromosome loss/gain) is currently under debate within some regulatory authorities. All of these areas are the subject of active research and at the present time new assay procedures are becoming available to meet these needs.

Transgenic Models

With the advent of the new DNA technologies, several promising transgenic mouse models are now being validated for use as *in vivo* genotoxicity screens. If suitable, these assays would enable somatic mutation to be measured within most tissues of the exposed animal, and more importantly enable mutation induction and specificity to be measured in the actual target tissue for carcinogenesis. At present the two best validated mouse transgenic models are based on the insertion of the bacterial genes *LacI* or *LacZ* into the mouse genome. In the first, called the 'MutaMouse,' an *E. coli lacZ* (β-galactosidase) gene was cloned into a lambda-bacteriophage vector, and then microinjected into the pronucleus of (BALB/c $\times$ DBA/2) CD2 F1 eggs. In the second, called the 'Big Blue,' a *lacI* (lac repressor) gene cloned into a lambda shuttle vector was integrated into the genome of an inbred C57B/6 mouse. In both cases the transgenic mouse is treated with the test compound by the likely route of exposure, and genomic DNA is isolated from the desired tissues. The lambda vector is rescued from the genomic DNA by mixing with an *in vitro* 'Lambda-packaging' extract that excises the target lambda sector and packages the DNA into a lambda phage head. These phages are then used to infect host *E. coli* strain. For the *lacZ* system a *lacZ*$^-$ *E. coli* C indicator strain is used. For the *lacI* system the *E. coli* SCS-8 indicator strain is used which has a deleted *lac* region and contains a *LacZ*$^-$ gene on a phi 80 insert.

For the *lacZ* system the bacteria and the lambda phage containing the rescued DNA are incubated in soft agar with a chromogenic agent (X-gal). After 16 h of incubation the lysis plaques produced by the phage are scored. If a phage contains a normal *LacZ* gene, the X-gal is metabolized to a blue product. If a mutated *LacZ* gene is carried by the phage, the X-gal remains intact and no colour is produced. Thus scoring consists of counting clear plaques against a background of nonmutated blue plaques. For the *lacI* system a similar plating routine is carried out. In this case mutants are detected as blue mutant plaques against a background of nonmutant colourless plaques. These can arise due to mutations that inactivate the *lac* repressor protein or *lacI* promoter or mutations within the *lac* repressor binding domain which block repressor binding to the *lacZ* operator. The density of plaques is limited to 1500 pfu per plate to ensure accuracy in detection of plaques with a mutant phenotype.

However, for both systems the detection of mutants by the colour selection method is laborious and expensive. For this reason, the 'MutaMouse' model has been adapted to a positive selection method using *galE*$^-$ mutants of *E. coli*. This permits high plating densities and minimizes the need for multiple platings. In this system the phage containing the rescued vectors are used to infect *E. coli lac*$^-$*galE*$^-$ cells, which are then plated in the presence and absence of phenylgalactosidase (P-gal). In the presence of P-gal only *lacZ*$^-$ bacteria will grow and produce plaques. Those containing *lacZ*$^+$ (i.e. nonmutated gene) produce the enzyme β-galactosidase which converts P-gal to galactose. This is further converted to a toxic intermediate UDP-galactose which accumulates in the *galE*$^-$*E. coli* and kills the cells. **Figure 4** gives a summary of the procedure for the 'MutaMouse' system.

The optimal protocol for this type of assay has yet to be finalized. The role of cell proliferation and its effects on mutation fixation (expression), the need for single or multiple doses and the effects of age and DNA packaging efficiency on assay sensitivity remain to be optimized. However, there has been considerable progress in the validation of these techniques in recent years and they are now being used as *in vivo* screens to detect point mutation induction in a variety of tissues. The strategy for using these assays has been discussed by Gorelick (1995).

The systems described here are only the first generation of tests arising from the application of transgenic technology, and many of the basic parameters of assay conduct and data interpretation are still underdeveloped. It remains to be shown that mutational analysis of these transgenics is a true reflection of mutation of resident endogenous genes. The *lacI/lacZ* models are unsuited for the detection of

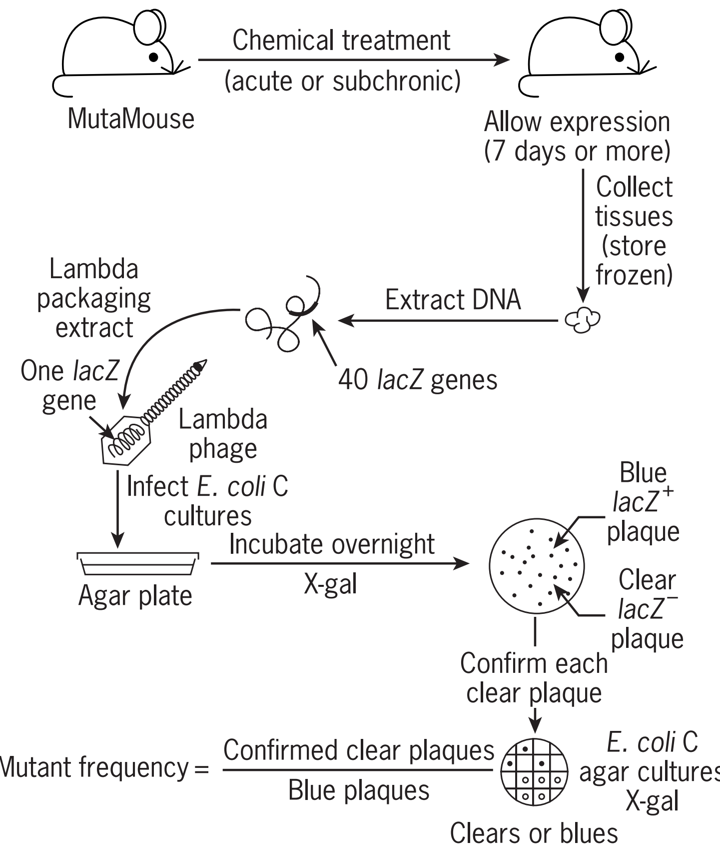


Figure 4 The transgenic MutaMouse *lacZ* system.

deletion mutations induced by clastogens, probably owing to the difficulty in packaging lambda phage vectors <42 or >52 kb. A new method using plasmid based transgenic animals may overcome these restrictions since recovery of plasmids is not so size dependent. It is believed that this system is potentially 10–20 times more efficient at rescuing *lac* genes (Dolle *et al*., 1996). (See also chapter on *Transgenic Technology in the Study of Oncogenes and Tumour Suppressor Genes*.)

Comet Assay

The use of transgenic mouse assays is one of several promising approaches that could provide an *in vivo* assay capable of detecting genotoxic damage in any tissue. Another possible method utilizes microgel electrophoresis and is called the single-cell gel electrophoresis assay or 'comet' assay. This is a rapid and simple system for measuring alkali labile sites and overt strand breaks in the DNA of mammalian cells (Fairbairn *et al*., 1995). Damaged (fragmented) DNA penetrates further than undamaged DNA into the agar following electrophoresis. The basis for this assay is represented in **Figure 5**. The technique can be applied to virtually any cell population or tissue type from which a single-cell suspension can be prepared. After treatment, the cells are suspended in agar and exposed to strong alkali, which denatures the proteins and permits DNA unfolding. Electrophoresis is then performed during which time the supercoiled DNA relaxes and fragmented DNA is pulled towards the anode. After electrophoresis, the slides are neutralized and stained with a DNA-specific stain such as propidium iodide, when the cell ghosts with damaged DNA are visible as comets of various sizes (hence the name), whereas those with undamaged DNA are visible as round images. The most common method of measurement is the tail moment, which takes into account the degree of DNA migration (comet length), and the DNA content within the comet. The technique has been adapted so that primary somatic cells of a variety of tissues can be studied, including cells from the GI tract, nasal mucosa and lung. The method requires only extremely small numbers of cells (from 1000 to 10 000) and results can be generated very quickly within a couple of days. It still requires further validation as the *in vivo* protocols are poorly defined, but the method holds great promise as an important tool for detecting DNA damage in virtually any mammalian cell population.

Tests for Aneuploidy (*In Vitro* Micronucleus Test)

Aneuploidy is considered to be a condition in which the chromosome number of a cell or individual deviates from

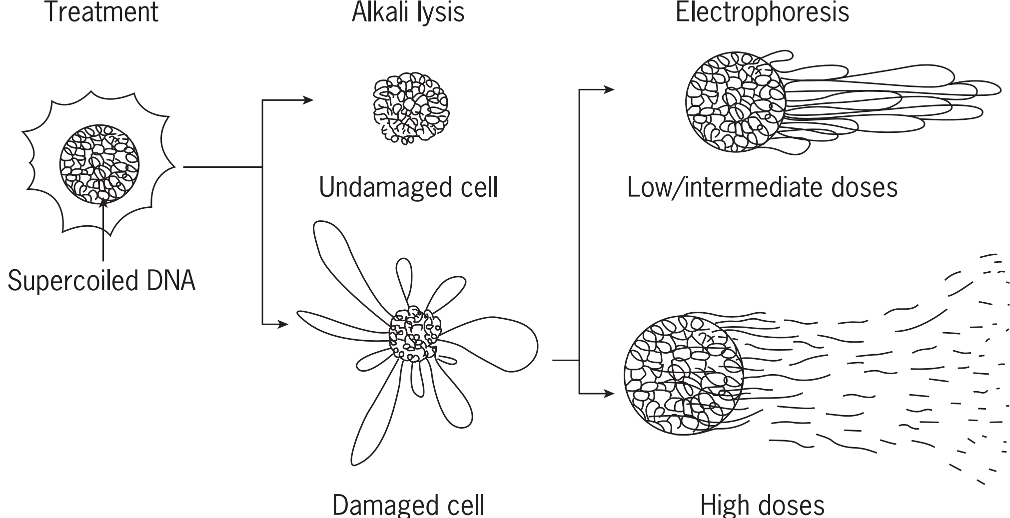


Figure 5 The theoretical basis for the formation of DNA comets.

a multiple of the haploid set. The maintenance of karyotype during cell division depends upon the fidelity of chromosome replication and the accurate segregation of chromosomes to daughter cells. In turn, these events depend upon different cell organelles functioning correctly and a number of metabolic activities related specifically to cell division (e.g. synthesis of nuclear spindle, proteins, etc.). Aneuploidy can occur through errors of many types, hence there are numerous cellular targets that can lead to chromosome gain or loss. Briefly, the mechanisms by which aneuploidy can occur fall into several categories, including damage to the mitotic spindle and associated elements, damage to chromosomal substructures, chromosome rearrangements, alterations to cellular physiology and mechanical disruption. The importance of aneuploidy to adverse human health is well accepted and the effects of aneuploidy include birth defects, spontaneous abortions and infertility. Tumour cells frequently have alterations in chromosome number and several specific aneuploidies have been associated with tumour development, although whether this is the cause or the effect of tumorigenesis is not clear.

When chromosomes fail to segregate correctly, this process of nondisjunction can lead to the production of both monosomic and trisomic progeny cells. If chromosomes are lost from the dividing nucleus they produce monosomic progeny without the reciprocal trisomic cell and the expelled chromosomes become membrane bound and are detected as micronuclei outside the main progeny nuclei. Consequently, chromosome loss can be measured by monitoring micronucleus formation and the *in vitro* micronucleus assay using mammalian cells provides such a technique (Fenech and Morley, 1985). This methodology has also been proposed as a simpler method for detecting chromosome breakage as micronuclei may also arise from acentric chromosome fragments (lacking a centromere) which are unable to migrate with the rest of the chromosomes during the anaphase of cell division. Because micronuclei in interphase cells can be assessed much more objectively than chromosomal aberrations in metaphase cells, there is not such rigorous a requirement for training personnel and slides can be scored more quickly. This makes it practical to score thousands instead of hundreds of cells per treatment and thus imparts greater statistical power to the assay. Micronuclei formed by aneuploidy induction can be distinguished from those produced by clastogenic activity by the presence of centromeric DNA or kinetochore proteins in the micronuclei. Fluorescent *in situ* hybridization (FISH) with pancentromeric DNA probes can be used to detect the former, whereas specific antibodies can be used to detect the presence of kinetochores.

The *in vitro* micronucleus assay may employ cultures of established cell lines, cell strains or primary cultures, including human and Chinese hamster fibroblasts, mouse lymphoma cells and human lymphocytes. To analyse the induction of micronuclei it is essential that nuclear division has occurred in both treated and untreated cultures. It is therefore important to provide evidence that cell proliferation has occurred after test chemical exposure. Analysis of the induction of micronuclei in human lymphocytes has indicated that the most convenient stage to score micronuclei in this cell system is the binucleate interphase stage. Such cells have completed one mitotic division after chemical treatment and are thus capable of expressing micronuclei. Treatment of the cells with the inhibitor of actin polymerization cytochalasin B inhibits microfilament assembly and cytokinesis, thus preventing the separation of daughter cells after mitosis and trapping them at the binucleate stage. A schematic for this method is shown in **Figure 6**.

The principle of the method is to expose cell cultures to the test substance in both the presence and absence of an

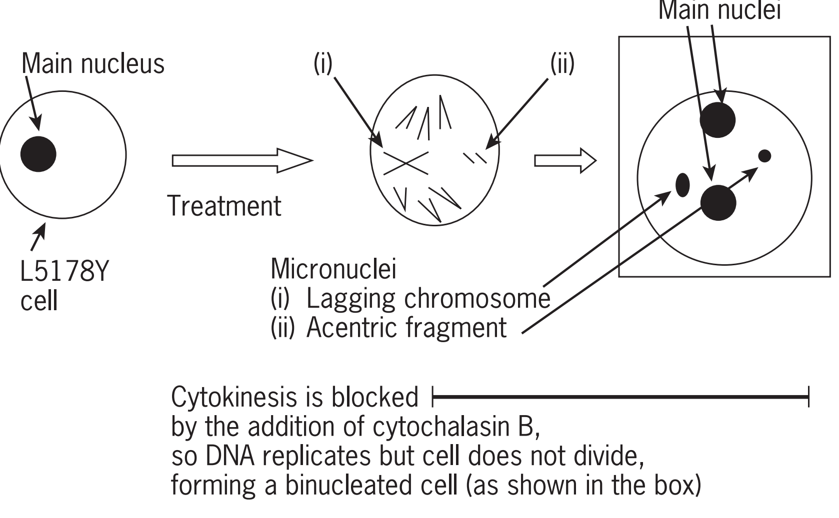


Figure 6 The basis for the *in vitro* micronucleus assay using cytochalasin B.

in vitro metabolizing system. After exposure, the cultures are grown for a period sufficient to allow chromosome damage or chromosome loss to lead to the formation of micronuclei in interphase cells. Harvested and stained interphase cells are then analysed microscopically for the presence of micronuclei. Micronuclei should only be scored in those cells which have completed nuclear division following exposure to the test chemical, normally for a period equivalent to approximately 1.5 normal cell cycles. Slides can be stained using various methods but fluorescent DNA-specific dyes are preferred as they will facilitate the detection of even very small micronuclei. At least 1000 cells per duplicate culture should be scored to assess the frequency of cells with one, two or more micronuclei. If the cytokinesis-block technique is used, micronucleus analysis is restricted to binucleate cells, and at least 1000 lymphocytes per duplicate culture should additionally be classified as mononucleates, binucleates or multinucleates to estimate the cytokinesis-block proliferation index, which is a measure of cell cycle delay.

At present, regulatory guidelines mainly focus on gene mutations and structural chromosome damage and do not address the induction of numerical abnormalities. However, this is currently under review, and it is recommended that such studies should be carried out, using a method such as the *in vitro* micronucleus assay, in the following situations: in the standard *in vitro* cytogenetics assay when there is evidence of an increased mitotic index, an increase in polyploidy or some evidence of increases in chromosome number on the slides (loss is not a useful measurement as this is a natural phenomenon due to slide artefacts), if there is a structural relationship with a known aneugen or if the mechanism of action involves modification of cell division.

It will continue to be of considerable importance to establish a specific role for chromosome loss in tumour development. The analysis of aneuploidy in interphase cells of solid tumours using FISH will be greatly advantageous in this respect. For cancer risk assessment purposes, results from aneuploidy assays can be considered particularly useful when the mode of action of a chemical is known to result in chromosome loss or nondisjunction.

INTERPRETATION OF RESULTS

Comparative trials have shown that false-positive and false-negative results in relation to cancer predictivity can be generated by each genotoxicity test. Experimental conditions such as the limited capacity of the *in vitro* metabolic activation system can lead to false-negative results in *in vitro* tests, whilst culture conditions such as changes in pH and high osmolality are known to cause false-positive results in *in vitro* mammalian assays. Guidelines for testing new chemical entities require a battery of tests measuring effects on a variety of genetic endpoints designed to detect the widest spectrum of genotoxins. This reduces the risk of false-negative results, whilst a positive result in any one *in vitro* assay does not necessarily mean that the chemical poses a genotoxic/carcinogenic hazard to humans. Further investigation in relevant *in vivo* assays is required to put the results into perspective. The *in vivo* tests have advantages in terms of relevant metabolism, etc., and also allow the influence of detoxification mechanisms to be assessed. The final assessment of the genotoxic potential of a chemical should take into account the totality of the findings and chemical class information when available. A risk–benefit approach

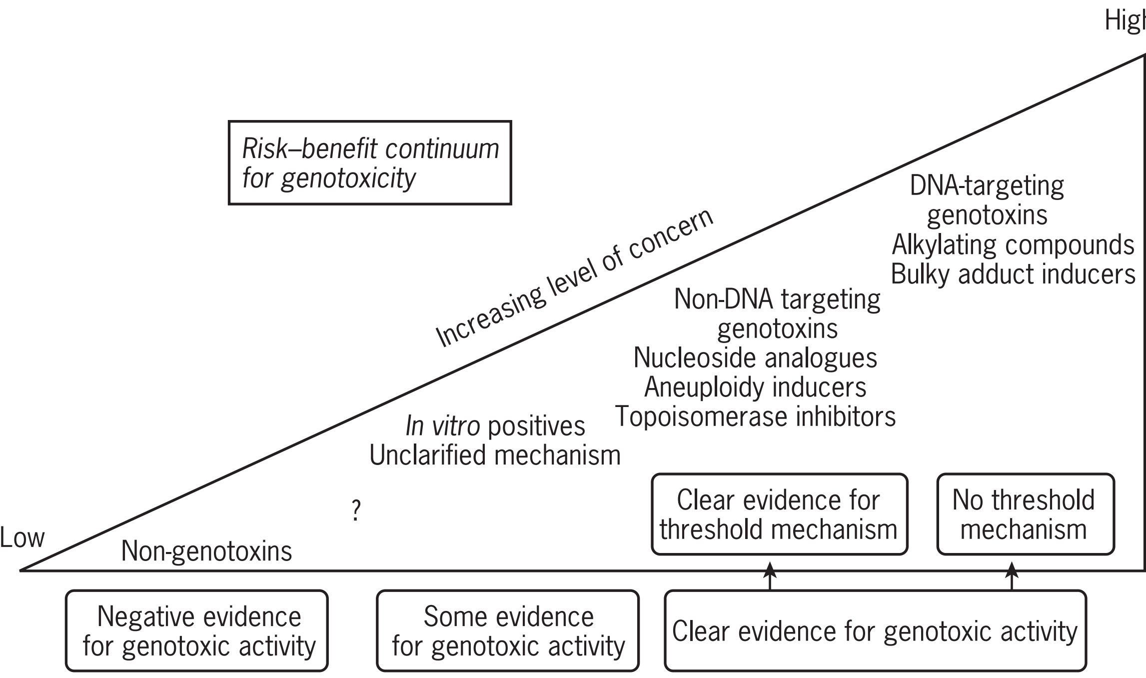


Figure 7 Interpretation of results from genotoxicity tests.

can be applied, such as that illustrated in **Figure 7**, which shows increasing levels of concern as the amount of positive data accumulates.

Unequivocally genotoxic chemicals are presumed to be trans-species carcinogens. Such chemicals should show 'clear evidence of genotoxicity,' for example, those which are positive in *in vivo* tests with supporting evidence from *in vitro* tests would be placed in this category. At the other end of the continuum are the nongenotoxic chemicals for which there is negative evidence from genotoxicity tests. In this case the chemicals should be negative in the standard battery of tests (**Figure 1**) with appropriate consideration of structural alerts, metabolism and exposure. It should be accepted that there is a low level of concern for these nongenotoxins, although it should be remembered that some chemicals could be carcinogenic by nongenotoxic mechanisms of action (see also the chapter *Non-genotoxic Causes of Cancer*). A number of chemicals have to be classified somewhere between these two extremes, that is, they show 'some evidence for genotoxic activity.' Certainly chemicals exist which show significant genotoxicity in *in vitro* tests without supporting evidence from *in vivo* tests. It is also known that due to the inherent limitations of currently employed *in vivo* genotoxicity tests, some of these chemicals may be carcinogenic through genotoxic mechanisms.

In the past, the existence of acceptable thresholds for exposure has not been considered in the field of genetic toxicology and in this respect risk–benefit consideration differed from other fields of toxicology. Chemicals representing a high level of concern and for which thresholds would not exist are probably those which target DNA directly, e.g. alkylating agents and bulky adduct inducers, although for some the risk might be low due to effective DNA repair. However, in recent years there has been a growing awareness of the existence of thresholds for those chemicals that act through non-DNA targets. Examples of such chemicals include nucleoside analogues, topoisomerase inhibitors and aneugens. In the last case, this can be illustrated by an example. For accurate segregation of chromosomes at mitosis, the spindle apparatus plays a significant role. A component of this is the attachment of multiple microtubules to the kinetochore and a smaller number to the chromosome arms. The number of microtubules that are attached varies among chromosomes, and so it is reasonable to assume that the average is above the minimum necessary. The effect of a chemical on microtubule formation or attachment to the chromosome might be to reduce the number attached but produce no effect until the number falls below the minimum necessary for segregation. This mechanism would be predicted to have a threshold type of dose response. Similar arguments can be made for perturbations of other components of a mitotic apparatus.

A more contentious area to which threshold arguments have been applied is the association between a number of cytotoxicity parameters and *in vitro* chromosomal damage. It is possible that a number of false-positive results arise *in vitro* from the evaluation of chemicals at concentrations that are overtly cytotoxic. However, the existence of cytotoxic clastogens and cytotoxic nonclastogens in the

same chemical class indicates that this is an oversimplistic view. Consequently, this area requires further research before an association between severe cytotoxicity and clastogenesis *in vitro* can be considered sufficient to suggest some type of threshold, and allow the conclusion that such results are less relevant to the *in vivo* situation.

FUTURE DEVELOPMENTS

Genetic toxicology is still an evolving science and there are a number of new techniques under development that could impact on the way in which screening is performed in the future. In conclusion to this chapter, a few examples will be briefly described.

High-throughput Screening

Chemical synthesis techniques are creating millions of new compounds and to cope with the vast numbers of tests required an approach termed high-throughput screening is being developed. A new branch of chemistry, called combinatorial chemistry, has emerged which allows chemists to take a number of relatively simple molecules and to combine them in every single possible permutation and in turn combine the products of these combinations. The result is a huge number new compounds which must be screened quickly and cheaply. However, as only tiny quantities are available, screening needs to be miniaturized. The essential elements of ultrahigh-throughput screening are easily measurable tests to determine whether the compound interacts with the target coupled with a high degree of automation and miniaturization. This presents a new challenge for genotoxicity screen development, and there are currently a number of new bacterial assays which allow such screening to take place using microwell technology (plates containing from 96 up to 1536 wells). One possible technique involves the measurement of SOS repair in bacteria. Many carcinogens and mutagens generate DNA lesions which block replication, often because of their bulky character, and result in the induction of the SOS response. This response in turn increases mutagenesis and survival by allowing replicative bypass of the DNA lesions. As the induction of the SOS response strongly correlates with DNA damage, spectrophotometric assays based on the expression of a β-galactosidase reporter gene transcribed from an SOS promoter have been developed. The SOS chromotest and the umu test are such examples. The attraction of these assays is that they have a simple spectrophotometric endpoint that is amenable to rapid automated scoring. An improvement to these assays has been the replacement of the chromogenic reporter system by expression of luciferase, an enzyme which catalyses ATP-dependent light emission from the substrate luciferin (Rettberg *et al*., 1999). It can be seen that although the standard *S. typhimurium* and *E. coli* strains are still essential for routine genotoxicity screening, the introduction of newly engineered varieties and assay miniaturization will provide much more rapid assays in the future.

Expert Computer Systems for Predicting Genotoxicity

An alternative way of screening large numbers of chemicals for genotoxicity is to develop software programs which can predict the likely outcome of the *in vitro* and *in vivo* tests, so-called '*in silico*' screens. There are now several commercially available systems that allow genotoxicity (and carcinogenicity) predictions to be made (**Table 3**).

The reliability of any expert system for predicting genotoxicity is crucially dependent on the quality of the database and the rulebase. The database and rulebase associated with an expert system enable the user to rationalize individual model predictions. Individual rules within the rulebase are generally of two main types. Some rules are based on mathematical induction, that is, by the extraction of correlations from a particular data set, whereas other rules are based on existing knowledge and expert judgement. Rules of the former type, 'induced rules,' offer the advantage of extending existing knowledge

Table 3 Commercially available expert systems for predicting genotoxicity and carcinogenicity

System	Acronym	Source	Type
Deductive Estimation of Risk from Existing Knowledge	DEREK	LHASA UK (Leeds University, UK)	Knowledge-based
Computer-assisted Structure Evaluation	CASE	Multicase Inc. (Cleveland, OH, USA)	Correlative
Toxicity Prediction by Computer-assisted Technology	TOPKAT	Health Design Inc. (Rochester, NY, USA)	Correlative
Hazard Expert	–	CompuDrug Chemistry Ltd (Budapest, Hungary)	Correlative
Oncologic	–	LogiChem Inc. (Boyertown, PA, USA)	Knowledge-based

without being biased toward particular mechanisms of toxic action. Their disadvantage, however, is that they may be nothing more than empirical relationships devoid of biological meaning. In contrast, rules of the latter type, 'expert rules' or 'knowledge-based rules,' are likely to have a strong mechanistic basis, but they are expressions of existing knowledge rather than of new knowledge. Typically, induced rules are (quantitative) structure–activity relationships, whereas expert rules are often based on reactive chemistry. A system based on induced rules is called a 'correlative system' whereas a system based on expert rules is referred to as a 'knowledge-based system.'

Many structural factors affect the mutagenicity and carcinogenicity of chemicals, including: (1) the intrinsic reactivity (electrophilicity); (2) the electron density in and near reactive centres; (3) substituent effects, e.g. steric hindrance; (4) susceptibility to metabolic activation and detoxification; (5) the stability of reactive forms of chemicals; (6) the ability of chemicals and their metabolites to traverse biological membranes; (7) the size and shape of molecules which control access to target sites on DNA; (8) the type and conformation of adducts formed between the chemical and DNA; and (9) the susceptibility of the adduct to DNA repair. Ideally, expert systems should take all of these factors into account when assessing the activities of mutagens and carcinogens.

At the present time, the available systems are still undergoing development and validation. Each system has particular strengths and weaknesses, and it seem unlikely that any one system will come to be regarded as the best and only choice for chemical risk assessment. The ideal situation might be an integrated approach that exploits the strengths of all of the available systems. In view of the possible presence of unknown contaminants with potent genotoxicities in a sample, expert systems can be used for screening purposes but not for providing a complete assurance of safety. They may be especially useful during the development of pharmaceuticals as a means of providing a high throughput screen which is potentially useful for early compound prioritization.

Metabolically Competent Cell Lines

One of the weakest elements within current *in vitro* genotoxicity screens is in the provision of an exogenous metabolizing system for the activation of pro-mutagens to DNA-reactive metabolites. First, such systems (usually rat liver S9 fraction with appropriate cofactors, see earlier) can be poor models for the likely human metabolism of a novel chemical. Second, they have the disadvantage that reactive metabolites formed exogenously may be unable to penetrate the cell membrane, or have short half-lives. The generation and detection of a genotoxic metabolite within the same cell therefore has obvious advantages. These problems may be alleviated by the introduction of human cDNAs expressing the appropriate metabolizing enzymes into cell lines that also possess those features required for the detection of genotoxic effects. A number of these so-called 'metabolically competent' cell lines have been developed, such as those derived from hepatocytes (e.g. HepG2 cells) or human B-lymphoblastoid cells (e.g. MCL5 cells). In the latter case, a panel of human cell lines have been developed expressing either individual or multiple human cytochrome P450s (Crespi *et al*., 1991). Although the use of these cell lines for genotoxicity screening has yet to be fully exploited, they do seem to provide an elegant system to overcome some of the failings inherent within current *in vitro* assays. Furthermore, the use of human cytochromes might reveal the formation of unique genotoxic human metabolites that could remain undetected with standard assays.

REFERENCES

Ames B. N. (1971). The detection of chemical mutagens with enteric bacteria. In: Hollander, A. (ed.), *Chemical Mutagens, Principles and Methods for Their Detection*, Vol. 1. 267–282 (Plenum Press, New York).

Ashby, J. and Tennant, R. W. (1988). Chemical structure, *Salmonella* mutagenicity and extent of carcinogenicity as indices of genotoxic carcinogens among 222 chemicals tested in rodents by the US NCI/NTP. *Mutation Research*, **204**, 17–115.

Clive, D., *et al*. (1987). Guide for performing the mouse lymphoma assay for mammalian cell mutagenicity. *Mutation Research* **189**, 145–156.

Cole, J., *et al*. (1990). Gene mutation assays in cultured mammalian cells. In: Kirkland, D. J. (ed.), *UKEMS Subcommittee on Guidelines for Mutagenicity Testing Report Part 1 Revised* 87–114 (Cambridge University Press, Cambridge).

Crespi, C., *et al*. (1991). A metabolically competent human cell line expressing five cDNAs encoding procarcinogen activation enyzmes: application to mutagenicity testing. *Chemical Research and Toxicology*, **4**, 566–572.

Dertinger, S. D., *et al*. (1997). Flow cytometric analysis of micronucleated reticulocytes in mouse bone marrow. *Mutation Research*, **390**, 257–262.

Dolle, M. E. T., *et al*. (1996). Evaluation of a plasmid-based transgenic mouse model for detecting *in vivo* mutations. *Mutagenesis*, **11**, 111–118.

Eliott, B., *et al*. (1992). Report of UKEMS Working Party: Alternatives to Aroclor1254 induced S9 in *in vitro* genotoxicity assays. *Mutagenesis*, **7**, 175–177.

Fairbairn, D. W., *et al*. (1995). The comet assay: a comprehensive review. *Mutation Research*, **339**, 37–59.

Fenech, M. and Morley, A. (1985). Measurement of micronuclei in lymphocytes. *Mutation Research*, **147**, 29–36.

Gatehouse, D., *et al*. (1994). Report from the Working Group on Bacterial Mutation Assays. *Mutation Research*, **312**, 217–233.

Gorelick, N. (1995). Overview of mutation assays in transgenic mice for routine testing. *Environmental and Molecular Mutagenesis*, **25**, 218–230.

Hayashi, M., *et al*. (2000). *In vivo* rodent erythrocyte micronucleus assay: II. The method for continuous repeat treatment and review of some aspects on the applications of the micronucleus assay. *Environmental and Molecular Mutagenesis*, **35**, 234–252.

ICH (1998). D'Arcy, P. F. and Harron, D. W. G. (eds), *Proceedings of the Fourth International Conference on Harmonisation, Brussels, 1997*. (Queens University of Belfast, Belfast).

Madle, S., *et al*. (1994). Recommendations for the performance of UDS tests *in vitro* and *in vivo*. *Mutation Research*, **312**, 263–285.

Mitchell, A. D., *et al*. (1997). The L5178Y tk+/− mouse lymphoma specific gene and chromosomal mutation assay. A Phase III Report of the US EPA GeneTox Programme. *Mutation Research*, **394**, 177–303.

Mortelmans, K. E. and Dousman, L. (1986). Mutagenesis and plasmids. In: de Serres, F. J. (ed.), *Chemical Mutagens, Principles and Methods for their Detection*, Vol. 10. 469–508 (Plenum Press, New York).

Rettberg, P., *et al*. (1999). Microscale application of the SOS-LUX Test as a biosensor for genotoxic agents. *Analytica Chimica Acta*, **387**, 289–296.

Scott, D., *et al*. (1990). Metaphase chromosome aberration assays *in vitro*. In: Kirkland, D. J. (ed.), *UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report Part 1. Revised Basic Mutagenicity Tests. UKEMS Recommended Procedures*. 3–86 (Cambridge University Press, Cambridge).

Storer, R., *et al*. (1997). The mouse lymphoma L5178Y tk+/− cell line is heterozygous for a codon 170 mutation in the p53 tumour suppressor gene. *Mutation Research*, **373**, 157–165.

Tennant, R. W., *et al*. (1987). Prediction of chemical carcinogenicity in rodents from *in vitro* genetic toxicity assays. *Science*, **236**, 933–941.

Tweats, D. and Gatehouse, D. (1999). Mutagenicity. In: Ballantyne, B., *et al*. (eds), *General and Applied Toxicology*, 1017–1078 Vol. 2 (Macmillan Reference, London).

FURTHER READING

Ballantyne, B., *et al*. (eds) (1999). *General and Applied Toxicology*, 2nd edn, Vol. 2. Part 5: *Genetic Toxicology, Carcinogenicity and Reproductive Toxicology*. (Macmillan, London).

Cartwright, A. C. and Matthews, B. R. (eds) (1997). *International Pharmaceutical Product Registration, Aspects of Quality, Safety and Efficacy*. Ch. 14, Mutagenicity (Ellis Horwood, Chichester).

EU (1998). *Testing of Medicinal Products for their Mutagenic Potential. The Rules Governing Medicinal Products in the European Union*, Vol. 3B, 45–50 (European Union, Brussels).

ICH S2A (1996). *Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals*. http://www.ifpma.org/ich1.html.

ICH S2B (1997). A Standard Battery for Genotoxicity Testing of Pharmaceuticals. http://www.ifpma.org/ich1.html.

Kirkland, D. (ed.) (1990). Basic Mutagenicity Tests: *UKEMS Recommended Procedures.* (Cambridge University Press, Cambridge).

Kirkland, D. J. and Fox, M. (eds) (1993). Supplementary Mutagenicity Tests: *UKEMS Recommended Procedures.* (Cambridge University Press, Cambridge).

Muller, L. (1998). The significance of positive results in genotoxicity testing. In: D'Arcy, P. F. and Harron, D. W. G. (eds), *Proceedings of the Fourth International Conference on Harmonisation* 253–259 (Greystone Books, Antrim).

Muller, L., *et al*. (1999). ICH – Harmonised guidance on genotoxicity testing of pharmaceuticals: evolution reasoning and impact. *Mutation Research*, **436**, 195–225.

OCED (1998). *Ninth Addendum to the OECD Guidelines for the Testing of Chemicals*. (OECD, Paris).

Sofuni, T. (1998). The role of the mouse lymphoma assay in genotoxicity testing. In: D'Arcy, P. F. and Harron, D. W. G. (eds), *Proceedings of the Fourth International Conference on Harmonisation*. 241–245 (Greystone Books, Antrim).

Chapter 26

Cancer Bioassays for Pharmaceuticals: a Regulatory Perspective

Joseph F. Contrera and Joseph J. DeGeorge
US Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Rockville, MD, USA

CONTENTS

- Introduction
- Background
- Synopsis of Carcinogenicity Guidances
- Discussion
- Conclusions
- Acknowledgement

INTRODUCTION

In the development of human pharmaceuticals, the lifetime 2-year mouse and rat carcinogenicity studies are the most costly toxicology studies in both time and resources. The outcome of these studies can also seriously impact the marketability of a product. The rodent carcinogenicity bioassay is a pivotal component regulatory policy and, until recently, the design of the studies had changed little over the last 20 years. The acceptability of the current 2-year rodent carcinogenicity bioassay was based upon the reasonable expectation of a relationship between a biological outcome in animals (neoplasia) and the desired regulatory endpoint (the assessment of potential human carcinogenic risk). The carcinogenicity bioassay was never formally validated, but evolved over years. The study gained acceptance by the scientific community after the accumulation of sufficient experience and the ability to identify compounds known or reasonably expected to be human carcinogens. Study protocols evolved with use, influenced greatly by practical considerations of statistical power, assay sensitivity and economic considerations. A similar process is now under way for the development and regulatory application of *in vivo* alternatives to the lifetime rodent carcinogenicity study such as transgenic models. Information derived from *in vivo* transgenic carcinogenicity models can contribute additional insights (Goldsworthy *et al*., 1994; Harris, 1995) into the mechanism of carcinogenesis and potential human risk that may be of greater value and potential relevance to humans than information from a conventional 2-year rodent study. New cancer bioassays and new approaches to carcinogenicity assessment that accompany them will significantly improve our ability to evaluate the potential carcinogenic risk of pharmaceuticals. (See section on *Treatment of Human Cancer*.)

BACKGROUND

Study Design and Analysis

The design of rodent carcinogenicity studies for pharmaceuticals is essentially the same as the design employed for industrial and environmental chemicals and US National Toxicology Program (NTP) rodent carcinogenicity studies. Male and female rats and mice are divided randomly into one or two control and three treatment groups of 50–70 animals per group per species. Historically, the highest dose in the studies analysed generally approximates the maximum tolerated dose (MTD) in the test species, and is administered daily usually in the feed or by oral gavage for 2 years. The rodent strains most often used in NTP studies is the inbred Fisher 344 rat and the hybrid B6C3F1 (C3H × C57B16) mouse. In pharmaceutical studies submitted to the Food and Drug Administration (FDA), the predominant rodent strains are the Sprague–Dawley-derived CD rat and the CD-1 Swiss–Webster-derived mouse. Despite our long experience with these assays, the significance of tumours from lifetime exposure at the maximum tolerated dose, dose–response extrapolation and the relevance of rodent tumours to humans continue to be highly controversial issues.

In studies reviewed by the FDA Center for Drug Evaluation and Research (CDER), tumour findings are usually classified as positive if either benign and/or malignant findings are statistically significant in pairwise comparison with concurrent controls by Fisher's exact or equivalent analysis, or by trend analysis. An adjustment for rare and common events that was recommended by Haseman (1983) is also applied to tumour findings. Tumours are considered significant in pairwise comparisons if they attained a level of $p \leq 0.01$ for common tumours and $p \leq 0.05$ for

rare tumours (≤1% background incidences rate). The incidences of benign and malignant tumours (adenomas and carcinomas) are combined where appropriate and evaluated statistically (McConnell *et al.*, 1986). Until recently, the results of such studies were generally viewed as either positive or negative since there was little supporting mechanistic, pharmacokinetic or comparative metabolism information to support decisions on the relevancy of the findings for humans. Discussions regarding study results primarily focused on whether an adequate dose was achieved without exceeding the MTD, and whether statistically significant increases in tumours were of biological significance. If biological significance was concluded, the predicted risk to humans was then considered in relation to the potential benefit of the drug in considering product market availability.

It has now become clear within the scientific and regulatory communities that there are important distinctions between pharmaceuticals and nonpharmaceuticals that influence study design and the evaluation of study results. Five of the more critical, interrelated distinctions of pharmaceuticals are as follows:

1. they are designed to have biological activity;
2. they are extensively evaluated in test species and in humans for their pharmacodynamic properties, including toxic effects;
3. they are tested under conditions where systemic exposure to drug and metabolites can be qualitatively and quantitatively compared between species;
4. they are used in humans in well controlled settings where benefits are provided to the user; and
5. they are developed only in limited circumstances when frank genotoxicity is observed.

Advances in analytical technology have greatly expanded the quantity and quality of supporting mechanistic, pharmacokinetic and comparative metabolism information in recent years. In conjunction with a large pharmaceutical database of carcinogenicity test results (Contrera *et al.*, 1995, 1997), these advances have supported a redefinition of dose selection criteria, carcinogenicity test methods and study evaluation procedures by the International Conference on Harmonization (ICH) of Technical Requirements for the Registration of Pharmaceuticals for Human Use. ICH guidances defining multiple acceptable endpoints for dose selection, acceptable test methods and an integrative 'weight of evidence' approach to evaluating the relevance of study results are now internationally adopted for human pharmaceuticals and available on the FDA CDER website (http://www.fda.gov/cder/guidance).

Study Reproducibility

With over 40 different organ and tissue sites examined in an average 2-year rodent carcinogenicity study, there is a high probability that statistically significant treatment-related tumour findings will arise by chance. It has been estimated that there is a 50% chance of identifying at least one false-positive tumour type in a treated group by chance alone when evaluating a mouse or rat carcinogenicity study (Haseman *et al.*, 1986). This is particularly true for common spontaneous tumours observed in older animals. In assessing the results of rodent carcinogenicity studies, the magnitude of the treatment-related carcinogenic response is an important factor when evaluating possible false-positive findings. The greater the treatment-related tumour incidence rate compared with concurrent controls, the more likely it is that the finding may be real (Haseman and Clark, 1991). Likewise, supporting evidence such as concordant findings across species and/or gender, tumour-related toxicity or pharmacodynamic target organ responses also help to substantiate tumour findings and rule out statistical false positives.

The significant potential for false-positive findings in 2-year rodent carcinogenicity studies also raises the question of study reproducibility. Unfortunately, there are only limited data addressing the repeatability or reproducibility of the results of carcinogenicity studies. Two-year rodent carcinogenicity studies are rarely repeated because of their considerable cost in time and resources. When they are repeated, it is generally due to failure of the original study or protocol deficiencies and therefore the results of repeated studies are rarely comparable owing to protocol and other differences between studies (e.g. dose, strain differences). A retrospective analysis of 18 NTP rodent carcinogenicity studies of colour additives in which two identical concurrent control groups were used demonstrated that significant variability can occur in the tumour incidence between two apparently identical control groups (Haseman *et al.*, 1986). The difference in tumour incidence between identical controls in these studies was often sufficiently large enough to alter the statistical significance of treatment-related tumours. These results suggest that variability inherent in 2-year rodent carcinogenicity studies may place a limit on the sensitivity value of these studies for single-gender, single-species tumour findings. An increasing number of carcinogenicity studies for pharmaceuticals now incorporate two identical but separately housed and analysed control groups (C1 and C2) to assess the variation in the spontaneous control tumour rate and reduce false-positive findings. In analysing dual control studies it is important to compare the tumour incidence of treated groups with each control group separately (C1 vs treated; C2 vs treated), in addition to a comparison against combined controls (C1 + C2 vs treated). Tumours are considered biologically significant when statistical significance is achieved compared with each of the control groups. Tumour findings are generally considered to lack biological significance when a treated group fails to achieve statistical significance compared with either control group. In cases where only one comparison shows significance, consideration of the tumour incidence rates

from contemporary historical controls may play a role in assessing the results of the study.

Test Species, Strains and Genetic Drift

Genetically inbred rodent strains are employed in the 2-year rodent carcinogenicity bioassay to minimize variability in tumour responses due to genetic differences. The use of rodents and the rodent strains currently used in carcinogenicity studies were not selected on the basis of their suitability as human surrogates but for pragmatic reasons such as reasonable sensitivity to carcinogens, an acceptable spontaneous tumour rate and reasonable lifespan and animal size. The large accumulated historical record and experience with currently used rodent strains is a major reason for their continued use, although this has also retarded the development of improved rodent strains. Although inbred strains are employed to reduce genetic diversity and variability in carcinogenicity studies, genetic drift is still a concern. For many rodent strains this has resulted in progressively larger, more obese animals with higher spontaneous tumour rates and reduced lifespan. This genetic drift has compromised assay sensitivity and the usefulness of some rodent strains in carcinogenicity testing and has diminished the value of the historical tumour record for these strains (Rao *et al*., 1990; Keenan *et al*., 1992, 1995). The relative sensitivity of most rodent strains commonly used in 2-year carcinogenicity studies has not been fully evaluated employing identical study protocols and a reference set of known human carcinogens. Positive controls of known carcinogens could address this, but they are rarely used in the standard 2-year rodent carcinogenicity study protocols (although they are usually used in the design of transgenic animal carcinogenicity studies). Evidence suggests that there are similarities and significant variations among different rodent strains in their sensitivity to various carcinogens and differences in the nature and incidence of spontaneous tumours (Drew *et al*., 1983; Haseman, 1983; Dragani *et al*., 1995). Many of these differences may be due to species and strain-related (genetic) differences in metabolism and pharmacokinetics. A wide variability in spontaneous tumour incidence can also occur within a rodent strain (Ward and Vlahakis, 1978; Barrett and Wiseman, 1992). In addition, it is now apparent that *ad libitum* feeding and body weight in rodent carcinogenicity studies have been an uncontrolled variable that significantly influenced the sensitivity of the bioassay to carcinogens (Tarone *et al*., 1981; Keenan *et al*., 1992; 1995). The tumorigenic response to a potent carcinogen such as aflatoxin was demonstrated to be significantly altered by varying the caloric intake and body weight of rodents (Everett, 1984).

Tumour Site Concordance

Tumour site concordance of rats and mice is relatively poor. In the FDA, NTP and CPD databases only 20–30% of all compounds with positive findings produced tumour findings in at least one common site in the rat and mouse (Gold *et al*., 1989; Huff and Haseman 1991; Contrera *et al*., 1997). The lack of site specificity has been cited as part of the rationale for a reduced 2-year study protocol employing only male rats and female mice (Haseman and Lockhart, 1993; Lai *et al*., 1994). Since tumour site cannot be reliably predicted between rodent species, extrapolation of rodent tumour sites to humans has also been questioned (Freedman *et al*., 1996). There appears to be better trans-species tumour site concordance for a select subset of International Agency for Research on Cancer (IARC) 1, 2A and 2B classified pharmaceuticals (Marselos and Vainio, 1991). This apparent concordance may be related to the genotoxic nature of a majority of the nonhormonal compounds in this group (Shelby and Zeiger, 1990).

SYNOPSIS OF CARCINOGENICITY GUIDANCES

The International Conference on the Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human use (ICH)

ICH Guidance S1A: The Need for Long-term Rodent Carcinogenicity Studies of Pharmaceuticals

1. Carcinogenicity studies are performed for any pharmaceutical expected to be used clinically for 6 months on a daily basis or those that will be used repeatedly in an intermittent manner to treat a chronic or recurrent condition.
2. Carcinogenicity studies are performed when there is cause for concern. This can include evidence of genotoxicity, a similarity to known carcinogens (e.g. SAR), the presence of preneoplastic findings in toxicity studies or long-term tissue retention of drug or drug-related products associated with pathophysiological responses.

For human pharmaceuticals, clinical considerations primarily determine the need for carcinogenicity studies and influence the assessment of benefit–risk. Clinical considerations include the expected duration of treatment, the severity of the disease or disorder, the nature of the patient population (e.g. children; the elderly) and the availability of other therapies and their toxicity profile.

ICH Guidance S1B: Testing for the Carcinogenic Potential of Pharmaceuticals

1. Carcinogenicity testing is generally done in two rodent species.

2. Use of an appropriate *in vivo* alternative to a second 2-year rodent carcinogenicity study is an option; however, the scientific justification for the use of any alternatives with a specific pharmaceutical is considered for acceptance of the assay.

The Basis for Carcinogenicity Studies in Two Rodent Species

The rodent carcinogenicity databases of the USA, the European Union and Japan were evaluated by the ICH safety expert working group. This analysis supported a more flexible approach to the then existing requirement for 2-year carcinogenicity studies in two rodent species. In the FDA database there is a significant (74%) concordance in the results of rat and mouse carcinogenicity studies and this is in accord with the results from the databases of the other ICH regions and the rodent carcinogenicity database of National Toxicology Program (NTP). Approximately 78% of all positive compounds identified in rat or mouse carcinogenicity studies would be identified in a rat study alone. Mouse-only carcinogenicity studies would be relatively less effective in this regard, identifying only 64% of all compounds with positive tumour findings identified in rat or mouse carcinogenicity studies (Contrera *et al*., 1997).

In considering the contribution of the second rodent species carcinogenicity study, there was recognition that the identification of compounds with trans-species tumorigenic effects is an important component of the weight of evidence for the assessment of human carcinogenic potential. Compounds that produce trans-species tumours are considered to pose a relatively higher risk to humans than single species-positive compounds (Tennant, 1993; Gray *et al*., 1995). It is postulated that the capacity of a drug to induce neoplasia across species suggests that the mechanism(s) involved in the induction of the neoplasia are conserved and, therefore, may have more significance to humans. Thus, compounds that produce rodent trans-species tumours are generally considered more potentially hazardous than compounds with single-species, single-site tumour findings. A major regulatory concern in relying on the results of a single carcinogenicity study in a rodent species is that, although all trans-species carcinogens would be detected employing a single species study, it would not be possible to identify trans-species carcinogens. In the FDA database, 52 of 125 drugs (42%) with tumour findings are trans-species positive (Contrera *et al*., 1997). Within this group there is a relatively high proportion of unmarketed drugs, older drugs and drugs that are marketed with restricted clinical indications related to carcinogenicity findings. Trans-species rodent tumour findings have also contributed significantly to regulatory decisions that prevented the marketing of a drug or resulted in the removal of a drug from the market. For example, trans-species rodent tumour findings contributed to the withdrawal of iodinated glycerol from the US market.

There are circumstances where tumour findings in a single rodent species can influence the weight of evidence when assessing potential cancer risk. Benzene is classified a human carcinogen by the IARC, on the basis of some human evidence and carcinogenic findings only in the mouse. Methapyrilene, an antihistamine, was removed from the market in the USA primarily owing to hepatocarcinogenicity with short latency in rats (Lijinsky *et al*., 1980). The IARC classified the tranquillizer oxazepam as a possible (2B) human carcinogen mainly as a result of the strength of mouse liver findings (IARC, 1996). Thus, the elimination of carcinogenicity testing in a second rodent species, likely the mouse, would significantly reduce the evidence available on which to base regulatory decisions regarding potential human risk. In the absence of a second species study, regulatory decisions would by necessity be based primarily upon the results of a single conventional carcinogenicity study, usually conducted in the rat. In this situation, more reliance may be placed on positive tumour findings in rats and these findings would effectively be regarded as equivalent to positive findings in two rodent species. As importantly, the ICH concluded that elimination of the second test species would not contribute to improving the assessment of potential human carcinogenic risk and would not advance the state of regulatory science.

The Optional Use of an Alternative to the Second 2-Year Rodent Carcinogenicity Study

The ICH evaluation of rodent carcinogenicity studies supported greater flexibility in the requirement for 2-year carcinogenicity studies in two rodent species. It was also acknowledged that compounds that produce tumours in two species may pose a higher risk to humans than those that produce tumours in only one species. Although it was concluded that studies in two species were usually necessary, the additional information from a second species does not need to be derived from a standard 2-year rodent carcinogenicity study. Some short-term *in vivo* transgenic mouse carcinogenicity models were considered sufficiently characterized to be used as alternatives to a conventional 2-year mouse study (Contrera *et al*., 1997). Advances in molecular biology have identified a growing number of genes such as proto-oncogenes and tumour-suppressor genes that are highly conserved across species and are associated with a wide variety of human and animal cancers. *In vivo* transgenic rodent models incorporating such mechanisms have application in identifying mechanisms involved in tumour formation and as selective tests for carcinogens. The generation of transgenic rodent models can be considered as a further extension of genetic manipulation by selective breeding that has long been employed in science and agriculture.

The ICH guidance contains provisions for the use of an *in vivo* transgenic or other suitable alternative to a second rodent study. One option is a single 2-year rodent carcinogenicity study in the most appropriate species (usually the

rat), and an additional short-term *in vivo* carcinogenicity study such as a transgenic mouse model in place of a 2-year mouse study. Transgenic models have been developed that contain regulated transgenes, unexpressed reporter genes or knocked out alleles of tumour-suppressor genes. (See section on *In Vitro and Animal Models for Human Cancer*.) Promising *in vivo* transgenic rodent models include the *TG.AC* v-Ha-*ras* oncogene-based mouse model (Tennant *et al*., 1996; Spalding *et al*., 1999), the human c-Ha-*ras* transgenic mouse model (Yamamoto *et al*., 1996), the *p53*-deficient mouse (Tennant *et al*., 1995, 1996) and the *XPA*-deficient mouse model (De Vries *et al*., 1997). Although transgenic models incorporating these relatively specific genetic mechanisms may not be responsive to all compounds that tested positive in the 2-year mouse or rat study, transgenic models can and have been developed that incorporate carcinogenic response elements that are known to be present and functioning similarly in humans. Positive findings in such transgenic models can offer valuable insights into the potential relevance and applicability of tumour findings to humans that are not readily available from standard 2-year rodent studies. In addition to providing carcinogenicity information from a second species, transgenic studies have contributed mechanistic information to the weight of evidence assessment of the potential carcinogenic risk of pharmaceuticals. The nature of the carcinogenic mechanisms underlying a tumour finding in a 2-year rodent study and the potential implications to humans can be addressed by an appropriate transgenic study. A tumour finding in a 2-year rodent study for a compound that was found to be clastogenic in a nonstandard genotoxicity study may have greater regulatory significance if it is demonstrated that the compound also produces tumours in *p53*-deficient mice (Dunnick *et al*., 1997). Since the *p53*-deficient mouse model is considered to be sensitive mainly to genotoxic carcinogens, the finding in the *p53* model suggests that the carcinogenic effect of the compound is derived from the genotoxicity observed. The specific molecular basis of the carcinogenic effect may also be further evaluated and provide further insights regarding the relevance of the finding to humans.

The choice of an appropriate species for a 2-year study should be supported by pharmacokinetic and metabolism information in the appropriate rodent strain and in humans. The use of one 2-year rodent study supplemented by a short-term *in vivo* carcinogenicity study in a second rodent species would be sufficient to identify most compounds that produce tumour findings in two rodent species. Together with comparative systemic exposure information in rodents and humans, the additional mechanistic information improves and strengthens the assessment of the potential carcinogenic risk of a compound to humans.

The treatment duration for transgenic mouse studies is generally 6–9 months with 15–20 male and female animals per group and three dose groups. The relatively small number of animals required for transgenic studies compared with 50 or more in 2-year studies is due largely to the low spontaneous background incidence of tumours and mortality in 6-month studies. A major difference in the design of transgenic studies compared with 2-year studies is the use of a positive control treated with a known carcinogen. The use of a positive control is considered essential for evaluating assay sensitivity. Dose selection studies for transgenic animals have been performed in either the transgenic animals themselves or the appropriate related wild-type strain. For example, in the case of the *p53*-deficient heterozygous mouse, the C57BL/6 wild-type mouse has been employed for dose-ranging studies. For compounds tested in both 2-year mouse carcinogenicity studies and transgenic models where the route of administration allows for a direct comparison, other than an earlier onset of tumours, transgenic models do not appear to be more sensitive than wild-type animals (i.e. doses showing positive responses are similar in both tests). (See also chapter *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*.)

The Issue of Assay Validation

For regulatory applications it is important to use bioassays which have been sufficiently characterized. In a discussion of the application of transgenic rodents in carcinogenicity testing it is necessary to consider objectively the relative strengths and deficiencies of transgenic models in the context of the strengths and deficiencies of the standard 2-year rodent carcinogenicity study. Study protocols for the 2-year bioassays have evolved over time until a relatively standard protocol developed, influenced by practical considerations of statistical power, assay sensitivity and economic/resource considerations. The current study gained acceptance by the scientific and regulatory community after an accumulation of a sufficient body of experience and the demonstrated ability to identify compounds reasonably expected to be carcinogens based on human and other data. Regulatory agencies began to apply the results of rodent carcinogenicity studies on the basis of relatively limited experience with these assays. No effort has been undertaken to evaluate or validate the existing design systematically. In contrast, there is an effort under way to evaluate systematically the application of transgenic models for carcinogenicity assessment, both at the NIEHS and as a collaborative effort by academia, government and industry coordinated by International Institute for Life Sciences (ILSI) Alternatives to Carcinogenicity Testing Committee. To date, the results of this effort to characterize systematically alternative methods are encouraging. Concern has been raised, however, about the application of transgenic models for quantitative risk assessment. The data available suggest that it is feasible, provided that there is adequate consideration of the response characteristics of the model being used. At present these assays may more readily be employed as qualitative assessments, and quantitative assessments could

be generated only in conjunction with standard assay data. It is also worth noting that the 2-year rodent carcinogenicity study was originally intended to be a qualitative screen for potential carcinogens and may itself be poorly suited for accurate quantitative risk assessment (Huff, 1993).

ICH Guidance S1C: Dose Selection for Carcinogenicity Studies of Pharmaceuticals

Several methods of dose selection are now in use:

1. maximum tolerated dose;
2. a pharmacokinetic endpoint (25-fold rat/human AUC ratio);
3. a dose-limiting pharmacodynamic endpoint;
4. saturation of absorption;
5. maximum feasible dose;
6. limit dose.

Maximum Tolerated Dose (MTD)

The MTD has been defined by the ICH for pharmaceuticals as 'the top dose or maximum tolerated dose is that which is predicted to produce a minimum toxic effect over the course of a carcinogenicity study.' Such an effect may be predicted from a 90-day dose range-finding study in which sufficient toxicity is observed. Toxicological factors to consider are alterations in physiological function which would be predicted to alter the animal's normal lifespan or interfere with interpretation of the study. These factors include not more than a 10% decrease in body weight gain relative to controls, target organ toxicity and significant alterations in clinical pathological parameters. This definition is considered in the ICH guidance to be equivalent to definitions of the MTD used earlier by the EU, Japan, the US FDA and the US National Toxicology Program.

A Pharmacokinetic Endpoint for Dose Selection

It has been claimed that the high doses required to achieve the MTD are likely to produce tissue damage and associated increased cellular proliferation which may increase the probability of positive tumour findings in carcinogenicity studies. It has also been stated that tumours generated at doses well above the clinical range may be of little of no relevance to human risk (McConnell, 1989; Ames and Gold, 1990; Perera, 1990; Carr and Kolbe, 1991; Cogliano *et al*., 1991; Rall, 1991). This issue is especially relevant to nongenotoxic drugs with low rodent toxicity which may require the administration of many multiples of the clinical dose to achieve the MTD.

As part of an ICH effort to develop an exposure (pharmacokinetic) based alternative to the MTD for high dose selection in carcinogenicity studies for drugs, the FDA examined carcinogenicity studies in FDA files (Contrera *et al*., 1995). Rodent and human pharmacokinetic exposure data were analysed to evaluate the relationship between the human maximum systemic exposure of drugs at the maximum therapeutic daily dose and rat systemic exposure achieved in carcinogenicity studies carried out at the MTD. The systemic exposure is equal to the area under the plasma concentration-time curve (AUC) for drug and or metabolites. Although carcinogenicity studies of drugs are carried out at the MTD, that often represents high multiples of the maximum human therapeutic dose, in a large proportion of these studies high multiples of the maximum human clinical systemic exposure were not attained. In about one third of the drugs tested at the MTD in rodents, systemic exposure in rodents was less than the human systemic exposure and 61% (20/33) attained a systemic exposure of less than 10-fold that in humans at the maximum therapeutic dose. A similar distribution was apparent when the rat/human body surface area (milligrams per square metre) dose ratio or mouse/human body surface area dose ratio was employed. When compared with the body surface area dose ratio, the milligrams per kilogram dose ratio usually overestimates the relative systemic exposure, and this has contributed to the false conclusion that excessively high and biologically irrelevant doses are routinely used in the carcinogenicity testing of pharmaceuticals. The MTD for pharmaceuticals in rodent carcinogenicity studies generally does not represent high multiples of the estimated human daily systemic exposure in part because drugs represent a class of compounds selected for high biological and pharmacological activity, which can produce adverse dose-limiting effects.

For use as a dose selection criterion, the magnitude of the relative systemic rodent/human exposure ratio (plasma AUC (rodent)/plasma AUC (human)) should be sufficiently high to detect known or probable IARC 1 and 2A carcinogenic pharmaceuticals and compounds with positive carcinogenicity findings in the FDA database. In order to be useful as a dose selection endpoint, an appropriate systemic exposure ratio should also be attainable by a reasonable proportion of compounds tested. On the basis of the current database, a minimum rodent/human systemic exposure ratio of 25 is sufficiently high to detect all IARC 1 and 2A carcinogenic pharmaceuticals and chemicals with positive carcinogenicity studies in the FDA database. On the basis of this data set a rodent/human systemic exposure ratio of 25 would be attained by approximately one in four nongenotoxic pharmaceuticals. A systemic exposure ratio of 25 thus represents an adequate margin of safety for rodent carcinogenicity studies, exceeding the systemic exposure currently attained by 75% of carcinogenicity studies carried out at the MTD in the FDA database. This ratio approach is considerd for all carcinogenicity studies irrespective of route of administration that meet the other criteria for use.

This approach to high dose selection in carcinogenicity studies can be used for nongenotoxic drugs with relatively low rodent toxicity that are similarly (at least qualitatively) metabolized in rodents and humans. In this context, genotoxicity is defined on the basis of a positive response in a standard battery of *in vitro* and *in vivo* tests used

for pharmaceutical testing. Genotoxic compounds are excluded from the systemic exposure approach because of the presumption that members of this class may produce cancer after a single dose and tumours caused at greater than 25-fold the human exposure could be relevant. Ideally, all human metabolites identified by radioisotopic or equivalent methods should be present in the rodent strain used in a carcinogenicity study. High dose selection can be based on either the parent drug, parent plus major metabolite(s) or solely on a metabolite(s), depending on the degree of biotransformation and the degree of similarity to that which occurs in humans. A large difference in the pattern of biotransformation of a drug in a rodent test strain compared with humans would suggest that the compound is a poor candidate for high dose selection based on systemic exposure. Compounds with relatively high rodent toxicity would likely reach the MTD well before a relative systemic exposure ratio of 25 could be attained and available data suggests that this occurs in approximately 75% of drugs tested.

A Dose-limiting Pharmacodynamic Endpoint

For pharmaceuticals, dose may be limited by an extension of the primary pharmacodynamic properties of the drug. Hypotension, anorexia or inhibition of blood clotting are examples of dose-limiting pharmacodynamic effects that would be associated with antihypertensive drugs, appetite suppressants or anticoagulant drugs.

Saturation of Absorption

High dose selection for systemically active drugs may be based on saturation of absorption as measured by the systemic availability (e.g. plasma concentration) of drug-related substances. At saturation, further increases in administered dose produce no significant increase in plasma concentration and are considered to have no biological significance relevant to assessing carcinogenic potential. The guidance recommends that mid and low doses selected for the carcinogenicity study, however, should also take into account saturation of metabolic and elimination pathways, regardless of the method used for high dose selection.

Maximum Feasible Dose

Currently, the maximum feasible dose by dietary administration is 5% of the diet. When other routes of administration are appropriate, the high dose will be limited based on practical considerations such as dosing volume 10–20 ml/kg per day of a viscous suspension and local intolerance.

Limit Dose

For nongenotoxic compounds where the maximum human recommended daily dose does not exceed 500 mg/kg it may not be necessary to exceed a dose of 1500 mg/kg per day in a carcinogenicity study. This only applies provided there is an absence of evidence of genotoxicity in the comprehensive testing used for pharmaceuticals and assurance that the systemic exposure in the rodent exceeds that in human by greater than an order of magnitude. From available experience, this endpoint is rarely applicable either as a result of genotoxicity findings or because of an inability to ensure that the exposure multiple has been achieved.

The ‘ Weight of Evidence’ Assessment of Carcinogenic Potential

An underlying principle in the application of ICH guidance’s on carcinogenicity testing is the need to evaluate and consider all information on the pharmaceutical that may be relevant to the interpretation of tumour findings and their significance to humans. Clearly, the biological relevance and clinical implications of carcinogenic findings should not be based solely on the results of the rodent bioassay but must include consideration of other information about the pharmaceutical in the context of its clinical use. Even when it is concluded that potential carcinogenic risk to human is associated with a pharmaceutical, information on the intended use and patient population, anticipated therapeutic benefit and margin of safety considerations can play a major role in any regulatory actions. The strength and nature of the tumour findings (e.g. rare tumours, especially those histologically similar to human tumours), evidence of genotoxicity and the degree of similarity in drug metabolism and pharmacokinetics compared with humans all contribute to the weight of evidence approach for assessing potential human carcinogenic risk in the risk–benefit balance (**Figure 1**). The sources of this information and the types of information are many, but can generally be divided into that derived from the carcinogenicity study itself and that available from external information sources.

The severity and morbidity associated with the clinical indication and the availability and toxicity of alternative therapies are also important considerations for human pharmaceuticals that influence the overall assessment of acceptable risk.

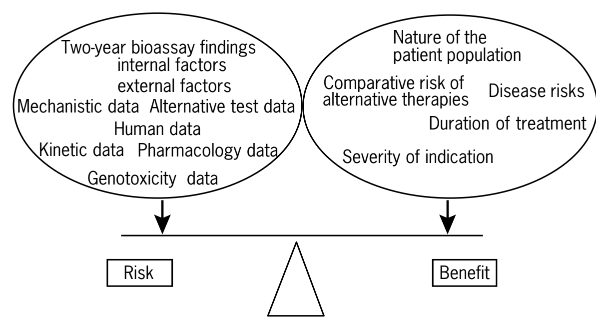


Figure 1 Summary assessment of carcinogenic risk and benefit for pharmaceuticals.

Carcinogenicity Study-derived Information

Study Acceptability

Adequacy of dose selection.
Appropriate route of administration relative to clinical route or other considerations.
Use of appropriate dose selection criteria as outlined in ICH guidance.
Adequate duration of dosing and study duration.
Adequate number of dose groups, animals/group and survival.
Appropriate selection and sampling of tissues and organs.
Scientific rationale for special design features (e.g. intermittent dosing, stop-dosing).
Scientific justification and rationale to support any alternative *in vivo* carcinogenicity studies.

Study Results

Organ- and tissue-specific tumour findings.
Magnitude and statistical significance of tumour findings compared with concurrent controls.
Malignant tumour incidence.
Benign tumour incidence.
Combined tumour incidence (e.g. McConnell *et al*., 1986).
Common or rare tumour classification (spontaneous incidence rate 1% or < 1%, respectively).
Dose–tumour response relationship.
Presence or absence of preneoplastic lesions and related tumour site toxicity.
Body weight effects that may confound or mask tumour responses.
Tumour latency.
Tumour multiplicity (presence of multiple tumours at the same site).
Multiple-site tumours (presence of tumours at different tissue and organ sites).

External Supporting Information

Relevant biological and pharmacodynamic mechanisms. Examples of such mechanisms include hormonal alterations, receptor-mediated effects, immune suppression and cell proliferation linked toxicity.
Comparative rodent and human pharmacokinetic and metabolic profiles and consideration of the adequacy of the rodent model.
Relative systemic exposure in humans and animals and margin of safety.
Evidence of genotoxicity.
Evidence of epigenetic mechanisms.
Relevant findings from toxicity studies.
SAR (structure–activity relationship) relationship to known carcinogens.
Comparisons with contemporary historical tumour rate (derived from recent studies from the same laboratory).
Results of studies with similar findings or of related products.
Results of special studies, e.g. non-standard genotoxicity studies, novel alternative *in vivo* tumour assays.
Relationship of rodent tumour genotypes to human cancer genotypes.
Human epidemiology and clinical data, if available.

DISCUSSION

A sponsor must evaluate critical decision points in each of the ICH guidance's and choose appropriate assays and assay parameters, or risk conducting assays that are not considered acceptable assessments of carcinogenic potential by the regulatory authorities. As might be anticipated because of the many recent changes in carcinogenicity testing, there have been some problems with implementation of these new guidances. These arise both from an absence of specific information needed to address questions for a given pharmaceutical and from an incomplete understanding or an incomplete application of the guidances as written. In part, it was the intention in writing the guidances to avoid excess proscription and detail. This allows flexibility to accommodate changes in the state of the art of toxicology. One result is that few of the decision points in the guidances are written with definitive detailed explanatory text. Hence the guidances are subject to much individual interpretation without benefit of the extensive background dialogue that occurred when each guidance was written.

To decrease the potential for studies to fail to achieve regulatory acceptance and improve consistency in the interpretation of study findings, FDA/CDER established and expanded the role of the Carcinogenicity Assessment Committee (CAC) and has adopted a policy of centralized, tertiary review of both study protocols and completed studies. During 1999, approximately 130 protocols and completed carcinogenicity studies were evaluated for adequacy, including over 20 involving alternative carcinogenicity study models. CDER has found that an open dialogue with industry on the scientific merits of the study design and dose selection and interpretation of study results in relation to human risk is extremely beneficial to both agency staff and industry personnel. This dialogue, accomplished through the use of CDER's Carcinogenicity Assessment Committee, is especially useful in cases of divergent viewpoints related to the interpretation and significance of study results. Sponsors are invited to participate in discussions regarding their study proposals or study results. While this process has on occasion been criticized as lengthy, performance goals have been instituted to ensure communication within 45 days of receipt of a study protocol for consultation. This goal has been achieved in

over 95% of the evaluations. This approach fosters greater mutual understanding of testing rationales and more consistent application of ICH guidances but, unfortunately, this process is generally only available for products where an Investigational New Drug (IND) application has been filed.

The ICH guidances on assessing the carcinogenic potential of pharmaceuticals have significantly changed the testing and evaluation process and have fostered research initiatives to improve test methods further. Given the rapid progress that is being made in developing new approaches and the changing scientific landscape upon which the regulatory evaluation process rests, a dialogue between the regulatory authorities and the regulated industry is essential if these science-driven guidances, requiring thoughtful consideration in their implementation, are to remain current. The outcomes of such discussions need to be shared broadly to achieve and maintain a harmonized understanding and consistent application of the guidances within the industry and between regulatory agencies. While this is being attempted within CDER with participation of the regulated industry, a forum for sharing information within the international drug regulatory arena has not yet arisen.

CONCLUSIONS

The 2-year rodent carcinogenicity study represented a pragmatic compromise, balancing factors such as test animal sensitivity, spontaneous tumour rate, lifespan and cost. To encourage the development of new assays that provide better information for evaluating human carcinogenic risk, special consideration is needed to foster progress by not prematurely demanding a higher degree of validation for new methods than has been attained by our current standard. A challenge for regulatory agencies and industry is to replace obsolete or redundant test requirements without compromising existing safety standards and while improving standards for the future. Promising transgenic rodent carcinogenicity models are currently being extensively characterized and many more will be developed in the future. Regulatory authorities can play an important role in fostering this process by demonstrating a willingness to consider and apply new scientifically acceptable methods through more flexible policies that accommodate improved, innovative approaches for assessing potential human carcinogenic risk.

The ICH process, based on mutual understanding and recognition of the available science, has resulted in guidances that have significantly changed and improved the process for assessing carcinogenic risk for humans for pharmaceuticals. A continued effort needs to be made to ensure further improvements in the future.

ACKNOWLEDGEMENT

This article is not an official FDA guidance or policy statement. No official support or endorsement by the Food and Drug Administration is intended or should be inferred.

REFERENCES

Ames, B. N. and Gold, L. S. (1990). Chemical carcinogenesis: too many carcinogens. *Proceedings of the National Academy of Sciences of the USA*, **87**, 7772–7776.

Barrett, J. C. and Wiseman, R. W. (1992). Molecular carcinogenesis in humans and rodents. In: Klein-Szanto, A. J. P., *et al*. (eds), *Comparative Molecular Carcinogenesis.* 1–30. (Wiley, New York).

Carr, C. J. and Kolbye, A. C. (1991). A critique of the maximum tolerated dose in bioassays to assess cancer risks from chemicals. *Regulatory Toxicology and Pharmacology*, **14**, 78–87.

Cogliano, V. J., *et al*. (1991). Carcinogens and human health: Part 3. *Science*, **251**, 606–608.

Contrera, J. F., *et al*. (1995). A systemic exposure based alternative to the maximum tolerated dose for carcinogenicity studies of human therapeutics. *Journal of the American College of Toxicologists*, **14**, 1–10.

Contrera, J. F., *et al*. (1997). Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. *Regulatory Toxicology and Pharmacology*, **25**, 130–145.

De vries, A., *et al*. (1997). Spontaneous liver tumours and benzopyrene-induced lymphomas in XPA-deficient mice. *Molecular Carcinogenesis*, **19**, 46–53.

Dragani, T. A., *et al*. (1995). Genetics of liver tumour susceptibility in mice. *Toxicology Letters*, **82/83**, 613–619.

Drew, R. T., *et al*. (1983). The effect of age and exposure duration on cancer induction by a known carcinogen in rats, mice and hamsters. *Toxicology and Applied Pharmacology*, **68**, 120–130.

Dunnick, J. K., *et al*. (1997). Phenophthalein rapidly induces malignant hematopoietic tumours and loss of heterozygosity in p53 wild type allele in heterozygous p53 deficient mice. *Toxicologcial Pathology*, **25**, 533–540.

Everett, R. (1984). Factors affecting spontaneous tumour incidence rates in mice: a literature review. *CRC Critical Reviews in Toxicology*, **13**, 235–251.

Freedman, D. A., *et al*. (1996). Concordance between rats and mice in bioassays for carcinogenesis. *Regulatory Toxicology and Pharmacology*, **23**, 225–232.

Gold, L. S., *et al*. (1989). Interspecies extrapolation in carcinogenesis: prediction between rats and mice. *Environmental Health Perspectives*, **81**, 211–219.

Goldsworthy, T. L., *et al*. (1994). Symposium overview: transgenic animals in toxicology. *Fundamentals of Applied Toxicology*, **22**, 8–19.

Gray, G. M. *et al*. (1995). An empirical examination of factors influencing prediction of carcinogenic hazard

across species. *Regulatory Toxicology Pharmacology*, **22**, 283–291.

Harris, C. C. (1995). Deichmann Lecture – p53 tumor suppressor gene: at the crossroads of molecular carcinogenesis, molecular epidemiology and cancer risk assessment. *Toxicology Letters*, **82/83**, 1–7.

Haseman, J. K. (1983). A re-examination of false-positive rates for carcinogenicity studies. *Fundamentals of Applied Toxicology*, **3**, 334–339.

Haseman, J. K. and Clark, A. (1990). Carcinogenicity results for 114 laboratory animal studies used to assess the predictivity of four *in vitro* genetic toxicity assays for rodent carcinogenicity. *Environmental and Molecular Mutagenesis*, **16, Suppl. 18**, 15–31.

Haseman, J. K. and Lockhart, A. (1993). Correlations between chemically related site-specific carcinogenic effects in long-term studies in rats and mice. *Environmental Health Perspectives*, **101**, 50–54.

Haseman, J. K., *et al*. (1986). Use of dual control groups to estimate, false positive rates in laboratory animal carcinogenicity studies. *Fundamentals of Applied Toxicology*, **7**, 573–584.

Huff, J. E. (1993). Issues and controversies surrounding qualitative strategies for identifying and forecasting cancer causing agents in the human environment. *Pharmacology and Toxicology*, **72, Suppl. 1**, 12–27.

Huff, J. and Haseman, J. (1991). Long-term chemical carcinogenesis experiments for identifying potential human cancer hazards: collective database of the National Cancer Institute and National Toxicology Program. *Environmental Health Perspectives*, **96**, 23–31.

IARC (1996). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 66. Some Pharmaceutical drugs*. (International Agency of Research on Cancer, Lyon).

Keenan, K. P., *et al*. (1992). The effect of diet and dietary optimization (caloric restriction) on survival in carcinogenicity studies. In: McAuslane, J., *et al*. (eds), *The Carcinogenicity Debate.* 77–102 (Quary Publishing, Lancaster).

Keenan, K. P. and Soper, K. A. (1995). The effects of ad libitum overfeeding and moderate dietary restriction on Sprague-Dawley rat survival, spontaneous carcinogenesis, chronic disease and the toxicologic response to pharmaceuticals. In: Hart, R. W., *et al*. (eds), *Dietary Restriction.* 99–126, (ILSI Press, Washington, DC).

Lai, D. Y., *et al*. (1994). Evaluation of reduced protocols for carcinogenicity testing of chemicals: report of a joint EPA/NIEHS workshop. *Regulatory Toxicology and Pharmacology*, **19**, 183–201.

Lijinsky, W., *et al*. (1980). Liver tumours induced in rats by oral administration of the antihistamine methapyrilene hydrochloride. *Science*, **209**, 817–819.

Marselos, M. and Vainio, H. (1991). Carcinogenic properties of pharmaceutical agents evaluated in the *IARC Monographs* programme. *Carcinogenesis*, **12**, 1751–1766.

McConnell, E. E. (1989). The maximum tolerated dose: the debate. *Journal of the American College of Toxicology*, **8**, 1115–1120.

McConnell, E. E., *et al*. (1986). Guidelines for combining neoplasms for evaluation of rodent carcinogenesis studies. *Journal of the National Cancer Institute*, **76**, 283–394.

Perera, F. P. (1990). Carcinogens and human health: Part 1. *Science*, **250**, 1644–1645.

Rall, D. P. (1991). Carcinogens and human health: Part 2. *Science*, **251**, 10–13.

Rao, G., *et al*. (1990). Growth, body weight, survival and tumor trends in F344/N rats during an eleven-year period. *Toxicological Pathology*, **18**, 61–70.

Shelby, M. D. and Zeiger, E. (1990). The genetic toxicity of human carcinogens and its implications. *Mutation Research*, **234**, 83–115.

Spalding, J. W., *et al*. (1999). Development of a transgenic mouse model for carcinogenesis bioassays: evaluation of chemically induced skin tumours in TG. AC mice. *Toxicological Sciences*, **49**, 241–254.

Tarone, R. E., *et al*. (1981). Variability in the rates of some common naturally occurring tumours in Fischer 344 rats and (C57BL/6N X C3H/HeN)F1 mice. *Journal of the National Cancer Institute*, **66**, 1175–1181.

Tennant, R. W. (1993). Stratification of rodent carcinogenicity bioassay results to reflect relative human hazard. *Mutation Research*, **286**, 111–118.

Tennant, R. W., *et al*. (1995). Identifying chemical carcinogens and assessing potential risk in short-term bioassays using transgenic mouse models. *Environmental Health Perspectives*, **103**, 942–950.

Tennant, R. W., *et al*. (1996). Evaluation of transgenic mouse bioassays for identifying carcinogens and noncarcinogens. *Mutation Research*, **365**, 119–127.

Ward, J. M. and Vlahakis, G. (1978). Evaluation of hepatocellular neoplasms in mice. *Journal of the National Cancer Institute*, **61**, 807–810.

Yamamoto, S., *et al*. (1996). Rapid induction of more malignant tumors by various genotoxic carcinogens in transgenic mice harboring human prototype c-Ha-ras gene than in control nontransgenic mice. *Carcinogenesis*, **17**, 2455–2461.

FURTHER READING

Alison, R. H., *et al*. Neoplastic lesions of questionable significance to humans. *Toxicological Pathology*, **22**, 170–186.

Gold, L. S., *et al*. (1991). Target organs in chronic bioassays of 533 chemical carcinogens. *Environmental Health Perspectives*, **93**, 233–246.

Hart, R., *et al*. (1995). *Dietary Restriction: Implications for the Design and Interpretation of Toxicity and Carcinogenicity Studies*. (ILSI Press, Washington, DC).

Huff, J., *et al*. (1991). Scientific concepts, value, and significance of chemical carcinogenesis studies. *Annual Reviews of Pharmacology and Toxicology*, **31**, 621–652.

Huff, J., *et al*. (1991). Chemicals associated with site-specific neoplasia in 1394 long-term carcinogenesis experiments in

laboratory rodents. *Environmental Health Perspectives*, **93**, 247–270.

ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods) (1997). *Validation and Regulatory Acceptance of Toxicological Test Methods*. NIH Publication No. 97–3981. (National Institute of Environmental Health Sciences, Washington, DC).

McClain, R. M. (1984). Mechanistic considerations in the regulation and classification of chemical carcinogens. In: Kotsomis, F. N., *et al*. (eds), *Nutritional Toxicology*. 273–303. (New York, Raven Press).

Milman, H. A. and Weisburger, E. K. (eds). (1985). *Hand book of Carcinogen Testing*. (Noyes Publications, New Jersey).

National Research Council (1993). *Issues in Risk Assessment*. (National Academy Press, Washington, DC).

Office of Science and Technology Policy (1985). Chemical carcinogens: a review of the science and its associated principles. *Federal Register*, **50**, 10371–10442.

Vaino, H., *et al*. (eds) (1992). *Mechanisms of Carcinogenesis in Risk Identification*: IARC Scientific Publication No. 116. (International Agency of Research on Cancer, Lyon).

Chapter 27

Molecular Epidemiology of Cancer

Paolo Vineis

Institute for Scientific Interchange and University of Turin, Turin, Italy

CONTENTS

- Why Molecular Epidemiology?
- Markers of Internal Dose: The Example of a Haemoglobin and DNA Adducts
- Markers of Early Response: The Example of *p53* Mutations
- Markers of Susceptibility
- Conclusions
- Acknowledgements

WHY MOLECULAR EPIDEMIOLOGY?

Epidemiology is the study of health and disease in populations, and of their determinants. The term molecular epidemiology may sound like an oxymoron, since it encompassess such different entities as molecules and populations. The aim of molecular epidemiology is to overcome some of the limitations of conventional epidemiology by linking research in the laboratory with research in free-living populations. In fact, rather than introducing an opposition between conventional and molecular epidemiology, it is more appropriate to consider the latter as a part of the first, involving the use of laboratory methods in order to overcome some of the methodological problems that are encountered in the study of the aetiology of human diseases.

The tools of molecular epidemiology include an appropriate study design, a specific attention to sources of bias and confounding and the development of markers that can be applied on a population scale. Study design is particularly important, because research that applied laboratory methods to human populations in the past was often based on 'convenience samples' that were affected by bias. Confounding, as we shall see, is a methodological problem that deserves special care.

Figure 1 is a simple representation of the goals of molecular epidemiology of cancer. Markers used in the molecular epidemiology of cancer are usually divided into the three categories of markers of internal dose, markers of early response and markers of susceptibility. In fact, each category includes subcategories. For example, protein adducts and DNA adducts are both markers of internal dose, but their biological meaning is different. 'Adduct' is a word that refers to the binding of an external compound to a molecule such as a protein or DNA. Whereas protein adducts are not repaired, i.e. they reflect external exposure more faithfully, DNA adducts are influenced by the individual repair ability; in fact, if they are not eliminated by the DNA repair machinery, they will induce a mutation. Also markers of early response are a heterogeneous category, that encompasses DNA mutations and gross chromosomal damage. The main advantage of early response markers is that they are more frequent than cancer itself and can be recognized earlier, thus allowing researchers to identify earlier effects of potentially carcinogenic exposures. Finally, markers of susceptibility include several subcategories, in particular a type of genetic susceptibility that is related to the metabolism of carcinogenic substances, and another type that is related to the repair of DNA (see below).

The complexity of the processes that lead to cancer and the ensuing multifactorial nature of epidemiological investigations is depicted in **Figure 2**.

Technical advancements such as high-throughput technologies for the analysis of (single-nucleotide polymorphisms (SNPs)) (see below) will make molecular epidemiology more powerful in the future, but will also bring new scientific and ethical challenges.

In the following a few examples are described and some methodological issues related to molecular epidemiology are raised.

MARKERS OF INTERNAL DOSE: THE EXAMPLE OF HAEMOGLOBIN AND DNA ADDUCTS

Engine exhaust, tobacco smoke and other complex mixtures contain several groups of carcinogenic compounds, including arylamines, polycyclic aromatic hydrocarbons (PAHs) and nitrosamines, many of which are able to

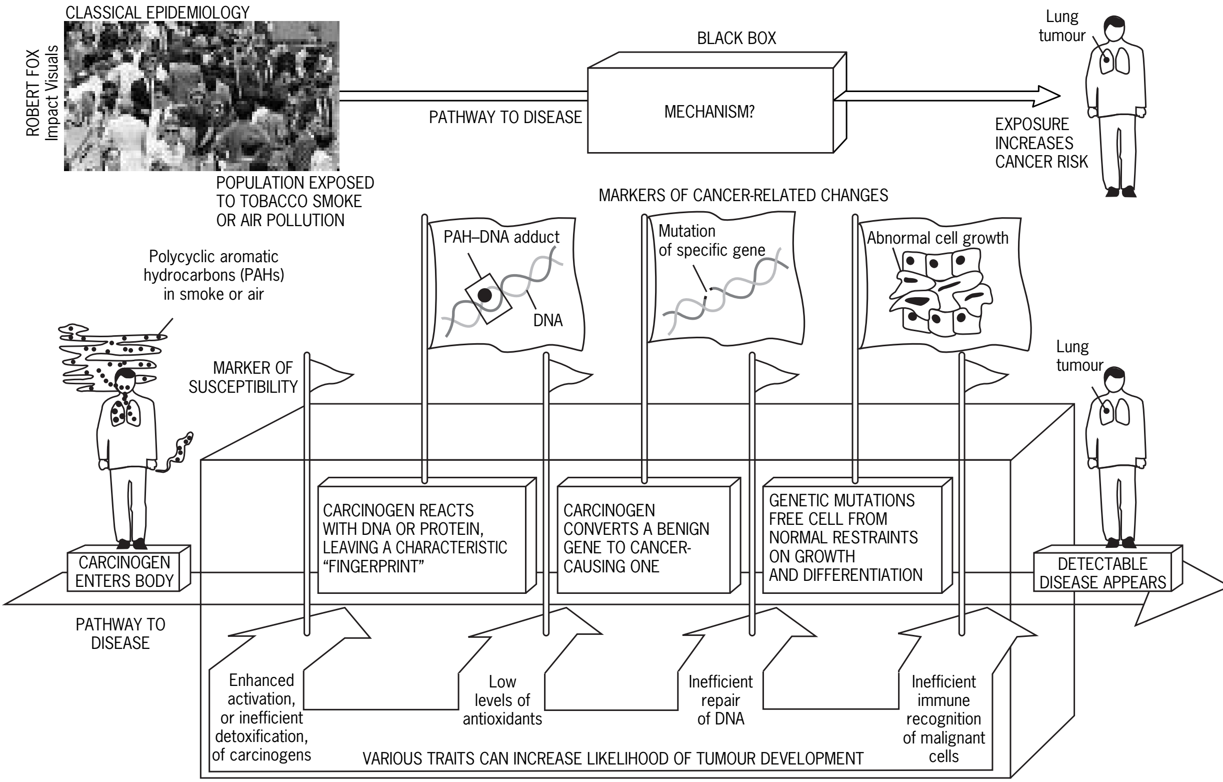


Figure 1 Uncovering new clues to cancer risk.

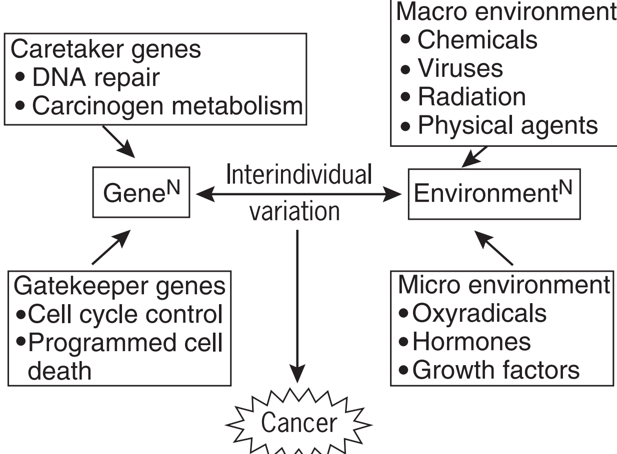


Figure 2 Many genes and environmental exposures contribute to the carcinogenic process. The effects can be additive or multiplicative, which are modifiable by interindividual variation in genetic function. It is proposed to include carcinogen metabolic activity and detoxification genes as caretaker genes involved in maintaining genomic integrity. (From Shields and Harris, 2000, *Journal of Clinical Oncology*, **18**, 2309-2315.)

form protein or DNA adducts after metabolic activation. The measurement of adducts is a method to partially overcome the inaccuracies inherent in traditional exposure assessment in epidemiology Different techniques have different biological meanings. Haemoglobin adducts have the advantage that large amounts of haemoglobin are available and well-established methods can lead to the identification of specific adducts, in particular from arylamines. The latter is a group of carcinogenic substances found in tobacco smoke, fried meat, car exhaust and some occupational environments. Arylamine–haemoglobin adducts have been shown to be associated with active and passive tobacco smoke (McLure *et al*., 1989; Perera, 1996). The disadvantage of haemoglobin adducts is the short half-life (8 weeks), i.e. they reflect recent exposure.

White blood cell (WBC)–DNA adducts measured with the ^{32}P-postlabelling technique give an overall measurement of aromatic compounds, i.e. a large group of environmental contaminants. The half-life is months, i.e. such adducts express a cumulative exposure in the last several months. Different extraction methods give different fractions of compounds. The nuclease P1 technique, alone, is able to detect bulky adducts, such as those formed by the PAHs and by some arylamines bound to the exocyclic position of guanine or adenine, while extraction with butanol is effective for most of the aromatic amines bound to the C8 position of guanine and some low molecular weight alkylating agents.

Previous ^{32}P-postlabelling studies have reported conflicting results of the association between the adduct levels in peripheral WBCs and tobacco smoking. Discrepancies may depend on the marked interindividual variation in the

metabolism of carcinogens, which results in different DNA adduct levels for similar degrees of exposure (Perera, 1996).

In addition to tobacco smoke, other exposures have been considered. Several studies in Europe, for example, have shown that the levels of WBC–DNA adducts were higher among subjects heavily exposed to air pollutants. This observation has been made among police officers (Peluso *et al*., 1998a), newspaper vendors exposed to urban traffic (Pastorelli *et al*., 1996), residents in a highly industrialized area in the United Kingdom (Farmer *et al*., 1996) and bus drivers in Denmark (Nielsen *et al*., 1996). In all these cases the more exposed subjects had significant differences from those less exposed, with WBC–DNA adducts of the order of about 3×10^8 for the relative adduct labelling (RAL) in the former and 1 in the latter. Benzo[*a*]pyrene is frequently used as a model compound for the class of PAHs. Benzo[*a*]pyrene is metabolically activated to benzo[*a*]pyrene diolepoxide (BPDE), the ultimate carcinogenic metabolite known to bind to DNA and blood proteins.

Lewtas *et al*. (1997) observed that human populations exposed to PAHs via air pollution show a non-linear relationship between levels of exposure and WBC–DNA adducts. Among highly exposed subjects, the DNA adduct level per unit of exposure was significantly lower than that measured for environmental exposures. The same exposure–dose non-linearity was observed in lung DNA from rats exposed to PAHs. One interpretation proposed for such an observation (Lutz, 1990; Garte *et al*., 1997) is that saturation of metabolic enzymes or induction of DNA repair processes occurs at high levels of exposure (see below). Also in humans occupationally exposed to PAHs a less than linear relationship between external exposure and WBC–DNA adducts was observed.

Dietary habits may also influence DNA adducts. The relationship between fruit and vegetable consumption and DNA adduct formation has been examined in a case-control study of bladder cancer. The level of WBC–DNA adducts measured by the ^{32}P-postlabelling method was strongly associated with the case/control status. The age-adjusted odds ratio (OR) for a level of adducts greater than the median was 3.7 (95% confidence interval = 2.2–6.3), and a dose–response relationship with quartiles of adducts was apparent. The level of WBC–DNA adducts decreased with increasing levels of fruit and vegetable consumption (Peluso *et al*., 2000). The association between the case/control status and the level of WBC–DNA adducts (below or above the median value) was stronger in the subjects who consumed less than two portions of vegetables per day (OR = 7.80; 95% confidence interval = 3.0–20.3) than in heavy consumers (OR = 4.98 for consumers of two portions per day; OR = 2.0 for consumers of three or more portions per day) (OR measures the extent of the risk increase in the exposed compared with the unexposed subjects).

MARKERS OF EARLY RESPONSE: THE EXAMPLE OF *p53* MUTATIONS

There is growing interest in the study of the relationship among carcinogenic exposures, the risk of cancer at specific sites and mutation spectra in relevant cancer genes (i.e. oncogenes and tumour-suppressor genes). One current theory is that specific carcinogens or carcinogenic mixtures (such as tobacco smoke) would leave a fingerprint (characteristic mutation spectrum) in relevant cancer genes.

Many *in vitro* and *in vivo* experimental studies have been conducted to elucidate carcinogenic mechanisms, with particular reference to mutation spectra in cancer genes. The fact that mutagens do not act randomly was already noted 30 years ago by Benzer, and then in a classical series of papers by Miller on the *lacI* gene, regarding mutational specificity in bacteria (see, for example, Coulondre *et al*., 1977; Miller *et al*., 1979). Denissenko *et al*. (1996, 1997, 1998) published some of the most interesting recent studies. They showed a strong and selective formation of adducts by the PAH 7,8,9,10-tetrahydrobenzo[*a*]pyrene at guanines in CpG sequences of codons 157, 248 and 273 of the *p53* gene, the major mutational hotspots in lung cancer. They concluded that methylated CpG dinucleotides are a target for chemical carcinogens in cancer genes.

An early example of a characteristic mutation spectrum in the human *p53* gene involved hepatocellular carcinoma in South Africa and China, and exposure to aflatoxins, a well-known dietary carcinogen. In early studies it was observed that about 50% of the patients in those areas had a relatively rare mutation, a G to T transversion at codon 249 (Bressac *et al*., 1991; Hsu *et al*., 1991). This mutation was deemed to be rare because it was not previously observed in patients living in areas where food contamination by aflatoxins is not common; furthermore, the same mutation could be induced experimentally by aflatoxin B_1 *in vitro*. On the contrary, *p53* mutational spectra induced by aflatoxin were not observed in rats (Hulla *et al*., 1993; Tokusashi *et al*., 1994). Another well-known example is represented by the characteristic C–T mutations induced by ultraviolet (UV) radiation at dipyrimidine sequences in the *p53* gene (Brash *et al*., 1991).

Some difficulties should be considered in the study of human cancer gene ‘fingerprints,’ such as (1) the multifactorial nature of human cancers, which hampers their attribution to single carcinogenic agents and/or the identification of common pathogenetic pathways; (2) the high genetic instability of cancer cells that may increase the frequency of mutations in certain cancer genes regardless of exposure factors; (3) the importance of DNA repair mechanisms and of the corresponding degree of population variation; (4) tissue selection bias that may affect the results, although its extent is difficult to establish; (5) the

simultaneous presence of clinical (i.e. treatment) and biological factors (i.e. stage, grading or unknown factors) related to the exposure and to the frequency of mutations that may confound its association; and (6) the need for a consideration of temporal sequences in the activation/deactivation of cancer genes.

Tumours are extremely dynamic entities and selection of tumour cell subpopulations is a continuous process. Mutations found in an advanced tumour may not be representative of the type of damage created by an agent in the DNA of the original target cell, since only those genetic changes that confer a growth advantage are selected in the course of the carcinogenic process. Certain mutations may provide sensitivity to the effects of tumour promoters, whereas others may not, thus resulting in the selective expansion of the former rather than the latter; this complicates the interpretation of the association between the genetic end-product and exposure to carcinogens.

Specific characteristics of bias and confounding in studies on mutational spectra should be considered. The size of the biopsies that are selected for investigation provides a prominent example of the types of selection bias that can occur in studies on cancer genes. In bladder cancer, for example, it is likely that early-stage tumours are too small to allow the urologist to obtain a biopsy sufficiently large for both research and clinical purposes. However, large biopsies tend to correspond to more advanced cases, which in turn may show a higher proportion of mutations in certain genes (Yaeger *et al*., 1998).

Another kind of bias that is particularly difficult to characterize and quantify is publication bias. Publication bias refers to the greater probability that positive studies (i.e. those showing an association for example between *p53* spectrum and exposure) become published. A way to identify publication bias is to plot the result of each study (expressed, for example, as an OR) against their size. In the absence of publication bias, the plot is expected to show great variability with small samples and lower variability with large samples, around a central value of the true OR. If publication bias occurs, negative results are not published, particularly if they arise from small investigations, and their results do not appear in the plot (Begg and Berlin, 1988). For example, in the large database available at the International Agency for Research on Cancer on *p53* mutations (Hainaut *et al*., 1998), the distribution of the proportion of mutations from different studies is skewed: all the studies with a proportion greater than 50% had less than 50 cases, while lower proportions were found in both small and large studies. This distribution does not necessarily imply that publication bias occurred; it might also suggest that large studies were based on heterogeneous populations, with a variable prevalence of mutations, while small studies refer to small subgroups with specific exposures to carcinogens and a genuinely high proportion of mutations.

Information bias is related to material mistakes in conducting laboratory or other analyses, or in reporting mutations; for example, a distortion arose from incorrect reporting of the *p53* sequence in an early paper, which influenced subsequent reports (Lamb and Crawford, 1986). However, this causes a genuine bias if the mistaken frequency is nonrandomly distributed in the categories that are compared.

Confounding occurs when a third variable creates a spurious association between the exposure at issue and the mutation spectrum. Several variables may act as confounders, for example if they modify the expression of oncogenes or tumour-suppressor genes. One such variable is chemotherapy: for example, cytostatic treatment for leukaemia induces characteristic cytogenetic abnormalities in chromosomes 5 and 7. Confounding arises if, for some reason, therapy is related to the exposure at issue. Stage is also a potential confounder. Therefore, studies that aim to determine the expression of cancer genes in humans should be restricted to untreated patients or specific stages, or statistical analyses should be stratified according to treatment/stage.

MARKERS OF SUSCEPTIBILITY

Metabolic polymorphisms

The human genome contains approximately 3 500 000 000 base pairs; of these, 10 000 000 are likely to differ among different individuals. Such variants in single base pairs are called single nucleotide polymorphisms (SNPs) and are potentially responsible for susceptibility to disease. It is becoming clear that only a minority of cancers have a frank genetic origin, in the sense that they are due to a highly penetrant gene (e.g. 5–10% of breast cancers occur in women carrying mutations of the *BRCA1* gene, which confers a risk of breast cancer of about 50–60% in the carriers). At the other extreme of the scale, there are diseases that are entirely due to the environment, with no role for genetic predisposition (this is the case with 15 workers who were exposed to β-naphthylamine in the British chemical industry in the 1950s: all of them developed transitional-cell bladder cancer, with no role for individual susceptibility) (Case *et al*., 1954). Most cancers, however, are likely to be in the middle, i.e. to be due to an interaction between external exposure and genetic susceptibility caused by a low penetrance gene. One type of such susceptibility is related to the metabolism of carcinogens. Subjects with an SNP at a particular gene locus have a defect in the enzyme involved in the metabolism of a carcinogen, and therefore develop cancer more easily if they are exposed to the substance.

Enzymes involved in the metabolism of carcinogens belong to two categories, Phase I and Phase II enzymes, the former being involved in activation (usually by

oxidation) and the latter in the inactivation of carcinogens (usually by conjugation). Examples of Phase I enzymes are cytochromes P450 involved in the activation of nitrosamines, and examples of Phase II enzymes are glutathione-*S*-transferases or *N*-acetyltransferase, involved in the deactivation of arylamines. Such enzymes can be less active because of genetic polymorphisms, i.e. common variants due to SNPs. **Figure 3** is an example of a meta-analysis that has been prepared for an IARC Publication on metabolic polymorphisms (Vineis *et al*., 1999) and shows that, overall, the *N*-acetyltransferase slow genotype is associated with a 40% increase in the risk of bladder cancer. The vertical lines correspond to different studies, in which the effect of *N*-acetyltransferase has been measured by an OR and the corresponding confidence interval, while at the extreme right the overall OR (deriving from the meta-analysis) is reported. In fact, the risk was much higher in populations occupationally exposed to arylamines; an interaction with tobacco smoke has also been shown in a different meta-analysis (Marcus *et al*., 2000). One aspect that deserves a comment is the heterogeneity of the effects shown in **Figure 3**, suggesting that the interaction between the genetic trait and environmental exposures is important indeed, depending on the specific populations investigated.

Methodological Issues in the Study of Metabolic Polymorphisms

One issue that is relevant to all epidemiological investigations aimed at interactions between exposures and genetic traits (gene–environment interactions) is statistical power (sample size). Statistical power is usually inadequate in most studies on this subject, and appropriate *a priori* calculations of the size required to detect an interaction are needed. A second, related issue is subgroup analysis: statistically significant associations may arise by chance when multiple comparisons are made within a single study. This problem is not only a statistical issue, i.e. one that can be solved with mathematical tools. Rather, the best way to avoid the pitfalls associated with subgroup analysis, multiple comparisons and publication bias is to define sound scientific hypotheses *a priori*. This goal can be accomplished by ensuring a strong cooperation among all the figures involved, i.e. geneticists, biochemists, molecular biologists, epidemiologists and biostatisticians. A sound *a priori* hypothesis implies (1) that evidence has been provided that a genetic polymorphism is implicated in the metabolism of a given carcinogen, (2) that the polymorphism can be measured with a reasonably small degree of misclassification and (3) that epidemiological tools

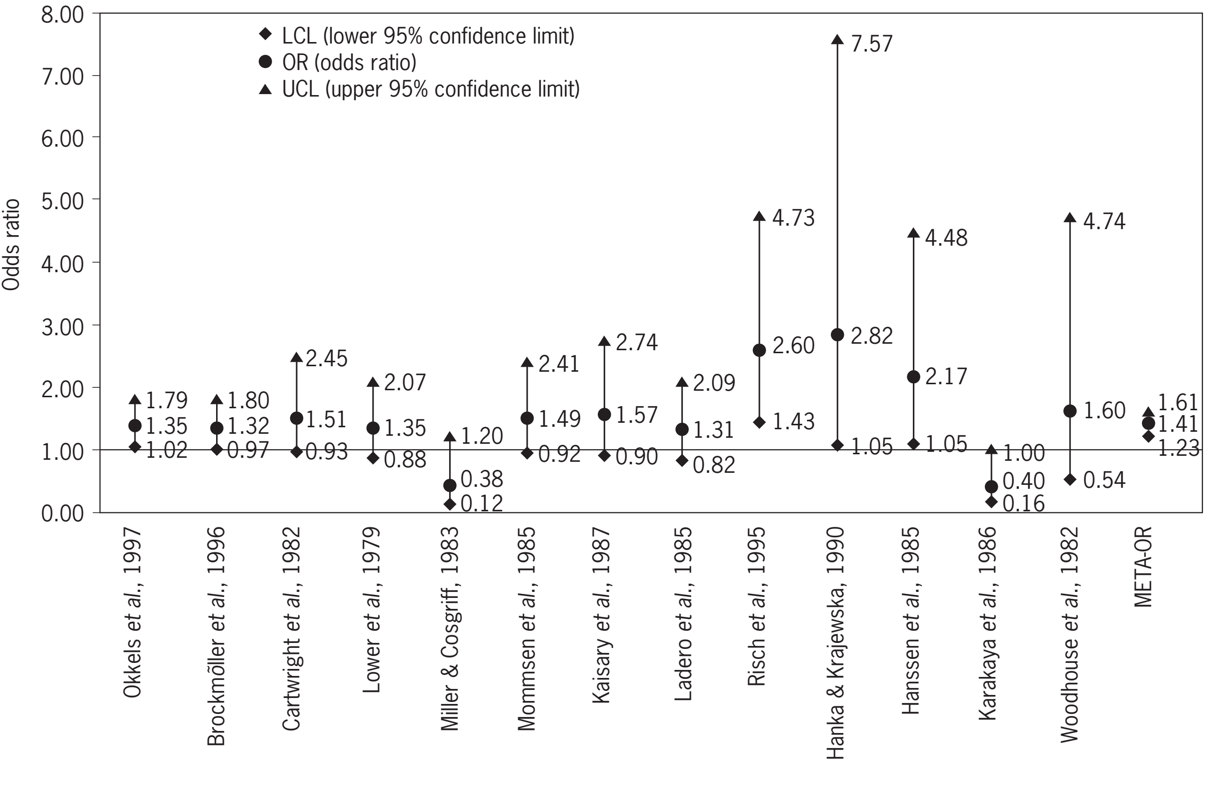


Figure 3 Meta-analysis of studies on NAT2 polymorphism and bladder cancer (caucasians). Note that the authors given are not cited in the present reference list. (From International Agency for Research on Cancer, 1998.)

allow the researcher to identify the exposed subjects with sufficient accuracy (i.e. exposure assessment is sound).

Cause–Effect Relationships

The *a posteriori* evaluation of published studies, in order to assess cause–effect relationships, is a difficult task that can only be accomplished by consensus in the context of Working Groups. I will not try to identify those gene–environment interactions that are more likely to be causal, but only suggest a method to accomplish this goal. Such a method is derived from the guidelines developed by Sir Bradford Hill to evaluate cause–effect relationships in observational studies (Hill, 1965):

- Strength of association (weak associations are more likely to be explained by bias, including publication bias, or confounding).
- Dose–response relationship (a criterion that applies to both genotyping, when two mutations are more effective than one, and to phenotyping, when the association with cancer risk is proportional to enzyme activity).
- Reproducibility of the association (in studies conducted in different populations with a different design).
- Internal coherence (for example, the association is observed in both genders, unless there is some biological explanation to justify gender differences).
- Biological plausibility (see above about the need for sound biological hypotheses; a particularly important issue is knowledge of the relevant carcinogens metabolized by the enzyme involved in the polymorphism).
- Specificity of the association (for example, the *N*-acetyltransferase polymorphism is associated with bladder cancer, a target site for arylamines).
- Animal models (which are expected to become available for metabolic polymorphisms with the development of the transgenic mice technology).
- Time sequence: although the genotype does not change with time, and its measurement within a cross-sectional design is meaningful, there are subtle problems of interpretation. For example, there is evidence that some polymorphisms influence the survival of cancer patients (Kelsey *et al*., 1997); therefore, measurement of the genotype within a cross-sectional design may simply imply the observation of a survival effect.

To mention some examples, on the one hand we have the very well-established association between bladder cancer, arylamine exposure and the *N*-acetyltransferase (NAT) polymorphism. In this case, all Hill's criteria are met: the carcinogens metabolized are known, they induce the type of cancer (bladder) that has been associated with the NAT polymorphism, the association has been found in different populations with different study designs, the association is strong in at least a few of the studies and animal models are available. At the other extreme we have the CYP2D6 polymorphism, whose role in carcinogenesis is still doubtful because of unresolved issues (the carcinogens metabolized are not known, and potential confounding due to genetic linkage has not been ruled out).

Public Health Applications

Ethical issues, related to the potential application of metabolic polymorphisms in a public health context, go far beyond the obtainment of informed consent. The use of metabolic polymorphisms to identify highly susceptible individuals has several implications that should be discussed thouroughly before any field application is approved: (1) the distribution of polymorphisms is uneven according to ethnic groups, which means that any job-related selection – on the basis of genetic susceptibility – would imply ethnic discrimination; (2) the role of insurances must be clarified if personal information on genetic susceptibility to cancer is released; and (3) metabolic susceptibility to cancer is such a complex issues that it can be hardly used to select the susceptible individuals for any meaningful purpose.

DNA Repair

In addition to polymorphisms in carcinogen metabolism, a potentially important source of variability is also DNA repair capability, including the genetic instability syndromes. These are rare, recessive traits that include ataxia-telangiectasia, Fanconi anaemia and Bloom syndrome, all characterized by both chromosome instability and high risk of cancer, and xeroderma pigmentosum, characterized by extreme susceptibility to UV radiation-induced skin cancer (Friedberg *et al*., 1995). In addition to these rare syndromes, individuals differ for their capability in repairing DNA damage. At least part of such inter-individual differences are likely to have a genetic origin. A number of studies have been conducted on the subject, based on comparisons between cancer cases and healthy controls for their supposed DNA repair function. A variety of tests to measure DNA repair have been developed. In fact, these studies can be broadly grouped into three categories, depending on the tests used. The first category includes tests based on DNA damage (usually chromosome breaks) induced with chemical (bleomycin) or physical (radiation) mutagens; mutagen sensitivity assay, Unscheduled DNA synthesis, [^{3}H]thymidine incorporation or count of pyrimidine dimers are examples. In this category of tests DNA repair is simply inferred by the different frequency of DNA damage induced in cancer cases and controls, without direct evidence of repair. The second category encompasses tests based on some direct evidence of repair, e.g. the plasmid *cat* gene test, the ADPRT modulation test or immunoassays based on antigenicity of thymidine dimers. The third category is represented by genotype-based tests, in which the distribution of polymorphic alleles is the object of the test.

A systematic review on the subject has been published (Berwick and Vineis, 2000). According to this, all studies based on tests belonging to the first category showed highly statistically significant results; when ORs were available they were between 2.8 (Spitz *et al*., 1994) and 10.3 (Spitz *et al*., 1989), suggesting a fairly intense association. Concerning the second category, two of the eight studies belonging to it, and included in the review (Berwick and Vineis, 2000), did not attain statistical significance. The only investigation based on genotyping at the time of the review (Hu *et al*., 1997) did not find a statistically significant association with breast cancer, although the phenotypic expression (oligonucleotide-induced PARP activity) showed a weak association with cancer. (See also chapter *Genomic Instability and DNA Repair*.)

Methodological Issues

Only one study had a population-based design and, ironically, this was the only clearly negative study (Hall *et al*., 1994). One possible explanation is that all other studies are affected by selection bias. However, plausible reasons for the negative result have been given by the authors and may differ from the lack of selection bias. It is hard to conceive that selection bias affected all the positive studies (i.e. the vast majority), since they were based on different series of controls sampled from different populations. In addition, to justify ORs of in the order of four the bias should have been very important.

Confounding is related to the possibility that some exposure or characteristic of the patients is associated with DNA repair and is a risk factor for cancer, thus creating a spurious relationship between DNA repair and the disease. Repair enzymes can be induced in several ways by stresses that damage DNA, e.g. pro-oxidative stress. According to recent investigations based on microchip technology, in yeast treated with an alkylating agent, the expression of over 200 genes was upregulated, whereas that of nearly 100 genes was downregulated. However, no information is available on the persistence of gene induction (P. Hanawalt, personal communication).

In studies in humans, several tests of DNA repair were affected by characteristics such as age, sunlight, dietary habits (with a relationship between carotenoids and repair proficiency), exposure to pro-oxidants and cancer therapies. While age and therapies were usually controlled for in most studies, dietary habits might have acted as confounders, since both the intake and the plasma level of carotenoids and other antioxidants have been shown to be lowered in cancer patients compared with healthy controls. The extent of such potential confounding is hard to estimate. In one study (Schantz *et al*., 1997), dietary habits were not associated with mutagen sensitivity in controls; rather, vitamins seemed to act as effect modifiers, not as confounders. How persistent the effect of potential confounders could be is unknown, although it has been suggested (Paleologo *et al*., 1992) that DNA damage induced by coal tar treatment of psoriasis could persist for more than 3 months.

Biological Plausibility

A major limitation of many tests (those belonging to the first category as defined above) is that DNA repair is only indirectly inferred from DNA damage. The only investigation based on genotyping considered in the review (Hu *et al*., 1997) did not find a statistically significant association with breast cancer, although the phenotypic expression (oligonucleotide-induced PARP activity) showed a weak association with cancer. This is an example of lack of phenotype-genotype correlation.

To draw firm conclusions about a cause–effect relationship, therefore, more evidence about the biological meaning of tests is needed. In particular, evidence has not been provided that many of these tests really express DNA repair. One possible interpretation is that some tests refer to a general and aspecific impairment of the DNA repair machinery, whereas others would explore more specific segments of it. However, this working hypothesis requires further evidence.

CONCLUSIONS

Gene–Environment Interactions and Cumulative DNA Damage

As a rule of thumb, there is an inverse proportionality between the frequency of at risk alleles and the risk of cancer (penetrance) (**Figure 4**). In particular, single highly penetrant mutations in cancer genes explain a small proportion of cancers (Vogelstein and Kinzler, 1998). This consideration arises both from empirical observation and from general scientific knowledge. Highly penetrant gene mutations, which confer an exceptionally high risk of

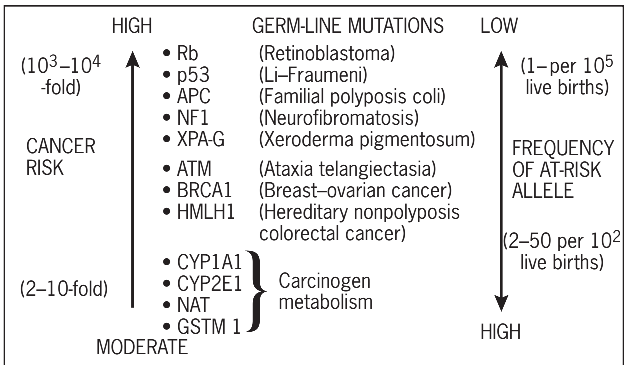


Figure 4 Examples of the inverse correlation between allelic frequency and cancer risk associated with selected cancer susceptibility genes. (From Hussain and Harris, 1998, *Cancer Research*, **58**, 4023-4037.)

cancer in the carriers, represent the tail of a distribution that includes (1) more common mutations in the same cancer genes (polymorphisms), that have a less disruptive effect on the protein function, or (2) rare or common mutations in genes that are less directly involved in the cancer process. There is increasing evidence in favour of both categories. Shen *et al*. (1998) have shown that even genes involved in rare and disruptive conditions such as xeroderma pigmentosum show common polymorphisms whose effects on the protein function (a DNA repair enzyme) are probably mild.

Concerning category (2), many metabolic polymorphisms are a clear example. Subjects with the *GSTM1* null genotype are frequent (about 50% of the caucasian population), have a serious genetic change (a deletion of the entire gene), but a slightly increased risk of some forms of cancer. In addition, many investigations (Berwick and Vineis, 2000) have been conducted on the greater susceptibility of cancer patients to DNA damage in comparison with controls. Most of these studies are based on the treatment of WBCs from cases and controls with a mutagen or a clastogen and on the observation of the frequencies of the induced DNA damage. The differences between cases and controls are interpreted as suggesting a greater susceptibility of cancer cases. As we have seen, the interpretation of such studies is not completely straightforward. In particular, the meaning of mutagen sensitivity tests is uncertain. Although virtually all studies show a greater sensitivity of cases, one cannot rule out that the apparent sensitivity is not due to a susceptibility factor, but is a consequence of cancer itself. Cancer is characterized by such a genetic instability that a mutagen-sensitive phenotype does not necessarily indicate a pre-existing susceptibility factor. In spite of such uncertainties, and of the aspecificity of mutagen sensitivity tests, the burden of investigations suggesting a greater sensitivity of some subjects to DNA damaging agents is impressive. Furthermore, the credibility of the observation is enhanced by the fact that in some investigations also healthy relatives of cancer patients, belonging to high-risk families, showed increased mutagen sensitivity or impaired DNA repair.

In addition to such studies, several investigations on DNA adducts (see the first section) suggest that cancer patients have a higher level of adducts after adjustment for relevant exposures. For example, in a study on bladder cancer we found that, after adjustment for smoking and dietary and occupational exposures, the level of WBC–DNA adducts was the variable that had the strongest association with the case vs control status (Peluso *et al*., 1998b). Similar observations have been made by others (Perera, 1996). In fact, adducts can be seen both as markers of cumulative internal dose and as markers of susceptibility.

Epidemiological studies have shown that in many instances duration of exposure is more important than the daily dose in increasing the risk of cancer. The paradigm for this general relationship is represented by smoking and lung cancer, but also experimental evidence has been produced. In the case of smoking, the incidence of lung cancer increased with the fourth power of duration and the second power of daily dose in one study (Doll and Peto, 1978). Other investigations did not find such a strong discrepancy between dose and duration, but the latter was nevertheless more relevant. Duration is mainly due to age at start; classical epidemiological studies have shown a very strong association between earlier age at start of smoking and the risk of lung cancer. In animals, fractionated and repeated doses induced tumours more frequently than the same total amount administered as a single dose (Lee and O'Neill, 1971; Lee *et al*., 1977). The latter observation is at odds with the general mechanisms of toxicity, according to which heavy exposure in a single administration has more devastating effects than repeated small doses (there are, however, some notable exceptions: not all carcinogens follow the rule suggested above).

In the light of such observations, a possible interpretation of the higher levels of adducts among cancer cases compared with controls is the concept of cumulative unrepaired DNA damage. What causes cancer, roughly speaking, would be the total burden of genotoxic chemicals that bind to DNA overcoming the repair processes. Such a burden may be higher because DNA repair is impaired (for genetic or acquired reasons) or because repeated exposures to the same agent occur.

Practical Consequences

If the premises are correct, important practical consequences follow. First, the contribution of genetic screening of populations is doomed to be rather limited. Genetic screening is sensible if at least two conditions are met: that the identification of a mutation is followed by effective preventive/therapeutic measures, that prolong survival and improve the quality of life; and that the population examined shows a high concentration of mutants so as to achieve a high predictive value of the screening test. If the prevalence of the mutation in the population is low, even in the case when we have effective preventive/therapeutic means, a screening strategy is unrealistic, since we have to screen hundreds of thousands to find one true positive plus (usually) a large number of false positives. Hence rare mutations can be reasonably sought in families, not in the general population.

Conversely, if the mutation is frequent (a polymorphism), its penetrance is likely to be very low and its effects to depend on interaction with external exposures. In such a case, even if the predictive value of the test is high, the success of screening is low: how can we deal effectively with 50% of the population (those with the *GSTM1* null genotype) who have a 30–40% excess risk of lung cancer? The best action is to advise them not to expose themselves to carcinogens, e.g. not to smoke, but such advice should

be obviously extended to the remaining 50% of the population. Another even better reason why genetic screening for common polymorphisms is not feasible is that multiple polymorphisms (tens or even hundreds) are involved in modulating the risk of cancer. Therefore, it makes little sense to identify a *GSTM1* null individual if the same person is at low risk for other metabolizing enzymes. It is obvious that for polymorphic conditions that interact with external exposures the only reasonable approach is avoidance of exposure.

The preceding considerations will be clearer with a quantitative example, based on the concept of 'number needed to treat' (NNT) (**Table 1**). Let us imagine that we have two different genetic traits, (A) one with low penetrance (1.4% cumulative lifetime risk in the carriers) and (B) the other with high penetrance (37% cumulative risk). Let us suppose that screening allows us to reduce the risk of cancer by 58% in both cases. This means that the absolute risk goes down to six per thousand in category A and to 15.5% in category B, with an absolute reduction of eight per thousand and 21.5%, respectively. The NNT is the inverse of such figures, i.e. 125 in category A and 4.5 in category B. This means that we have to screen 125 subjects to prevent one cancer in category A (even with a relative effectiveness of the intervention as high as 58%), whereas it is sufficient to screen 4.5 individuals in category B to achieve the same result.

In conclusion, what practical strategies can be proposed? If we accept that genetic screening should be limited to high-risk families, and that, apart from these, the risk of cancer depends on the total cumulative unrepaired DNA damage, then avoidance of exposure is the only realistic approach, even for low exposures. Low exposures have two properties that make them potentially highly relevant to the population cancer risk: they are frequent or even ubiquitous (e.g. dietary components, air or water pollution) and they are prolonged rather than limited in time. The average person in Western societies is exposed to low-level genotoxic pollutants for all their life. Such low-level exposure starts at a very young age and induces cumulative DNA damage which, if unrepaired, will be able to increase the cancer risk. Although many controversies have arisen about thresholds for carcinogens, little attention has been paid to long-lasting exposures.

Table 1 Calculation of the number needed to treat for a genetic test involving a highly penetrant gene or a low-penetrance gene

	Genotype	
	Low penetrance	**High penetrance**
Risk of cancer without screening (U)	0.014	0.37
Risk reduction due to screening ((U − T)/U)	58%	58%
Risk of cancer in the screenees (T)	0.006	0.155
Absolute reduction of risk (U − T)	0.008	0.215
NNT to prevent one cancer 1/ (U − T)	1/0.008 = 125	1/0.215 = 4.5

Advantages and Disadvantages of Molecular Epidemiology

Conventional epidemiology, based on simple tools such as interviews and questionnaires, has achieved extremely important goals, including the discovery of the causal relationships between smoking and lung cancer or asbestos and mesothelioma. Even a difficult issue such as the relationship between air pollution and chronic disease has been successfully dealt with by time-series analysis and other methods not based on the laboratory. Therefore, the use of molecular epidemiology should be carefully considered to answer specific scientific questions. The following are examples: (1) a better characterization of exposures, particularly when levels of exposure are very low or different sources of exposure should be integrated in a single measure; (2) the study of gene–environment interactions; and (3) the use of markers of early response, in order to overcome the main limitations of cancer epidemiology, i.e. the relatively low frequency of specific forms of cancer and the long latency period between exposure and the onset of disease. Limitations of molecular epidemiology should also be acknowledged: the complexity of many laboratory methods, with partially unknown levels of measurement error or interlaboratory variability; the scanty knowledge of the sources of bias and confounding; in some circumstances, the lower degree of accuracy (for example, urinary nicotine compared with questionnaires on smoking habits); and the uncertain biological meaning of markers, as in the case of some types of adducts or some early response markers (typically mutation spectra).

ACKNOWLEDGEMENTS

This work was made possible by a grant from the Associazione Italiana per le Ricerche sul Cancro and a grant from the European Union to P.V. for the project Gen-Air (QRLT-1999-00927)

REFERENCES

Begg, C. B. and Berlin, J. A. (1998). Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society*, **151**, (Part 3), 419–463.

Berwick, M. and Vineis, P. (2000). DNA repair in human cancer: an epidemiologic review. *Journal of the National Cancer Institute*, **92**, 874–897.

Brash, D. E., *et al*. (1991). A role for sunlight in skin cancer: UV-induced *p53* mutations in squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the USA*, **88**, 10124–10128.

Bressac, B., *et al*. (1991). Selective G to T mutations of *p53* gene in hepatocellular carcinoma from southern Africa. *Nature*, **350**, 429–431.

Case, R. A. M., *et al*. (1954). Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. I. The role of aniline, benzidine, alpha-naphthylamine and beta-naphthylamine. *British Journal of Industrial Medicine*, **11**, 75–212.

Coulondre, C. and Miller, J. A. (1977). Genetic studies of the lac repressor. IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli*. *Journal of Molecular Biology*, **117**, 577–606.

Denissenko, M. F., *et al*. (1996). Preferential formation of benzo[*a*]pyrene adducts at lung cancer mutational hotspots in *P53*. *Science*, **274**, 430–432.

Denissenko, M. F., *et al*. (1997). Cytosine methylation determines hot spots of DNA damage in the human *P53* gene. *Proceedings of the National Academy of Sciences of the USA*, **94**, 3893–3898.

Denissenko, M. F., *et al*. (1998). Slow repair of bulky DNA adducts along the nontranscribed strand of the human *p53* gene may explain the strand bias of transversion mutations in cancers. *Oncogene*, **16**, 1241–1247.

Doll, R. and Peto, R. (1978). Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. *Journal of Epidemiology and Community Health*, **32**, 303–313.

Farmer, P. B., *et al*. (1996). Biomonitoring human exposure to environmental carcinogenic chemicals. *Mutagenesis*, **11**, 363–338.

Friedberg, E. C., *et al*. (1995). *DNA Repair and Mutagenesis.* (ASM Press, Washington, DC).

Garte, S., *et al*. (1997). Gene–environment interactions in the application of biomarkers of cancer susceptibility in epidemiology. In: Toniolo, P., *et al*. (eds), *Application of Biomarkers in Cancer Epidemiology.* IARC Scientific Publication No. 142. (IARC, Lyon).

Hainaut, P., *et al*. (1998). IARC database of *p53* gene mutations in human tumors and cell lines: updated compilation, revised formats and new visualization tools. *Nucleic Acids Research*, **26**, 205–213.

Hall, J., *et al*. (1994). DNA repair capacity as a risk factor for non-melanocytic skin cancer – a molecular epidemiological study. *International Journal of Cancer*, **58**, 179–184.

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, **58**, 295–300.

Hsu, I. C., *et al*. (1991). Mutational hotspots in the *p53* gene in human hepatocellular carcinomas. *Nature*, **350**, 427–428.

Hu, J. J., *et al*. (1997). Poly(ADP-ribose) polymerase in human breast cancer: a case-control study. *Pharmacogenetics*, **7**, 309–316.

Hulla, J. E., *et al*. (1993). Aflatoxin B1-induced rat hepatic hyperplastic nodules do not exhibit a site-specific mutation within the *p53* gene. *Cancer Research*, **53**, 9–11.

Kelsey, T. K., *et al*. (1997). Glutathione-*S*-transferase class mu deletion polymorphism and breast cancer: results from prevalent versus incident cases. *Cancer Epidemiology Biomarkers and Prevention*, **6**, 511–516.

Lamb, P. and Crawford, L. (1986). Characterization of the human *p53* gene. *Molecular Cell Biology*, **6**, 1379–1385.

Lee, P. N. and O'Neill, J. A. (1971). The effect of both time and dose on tumour incidence rate in benzopyrene skin painting experiments. *British Journal of Cancer*, **25**, 759–770.

Lee, P. N., *et al*. (1977). Fractionation of mouse skin carcinogens in cigarette smoke condensate. *British Journal of Cancer*, **35**, 730–742.

Lewtas, J., *et al*. (1997). Air pollution exposure–DNA dosimetry in humans and rodents: evidence for non-linearity at high doses. *Mutation Research*, **378**, 51–63.

Lutz, W. K. (1990). Dose–response relationship and low-dose extrapolation in chemical carcinogenesis. *Carcinogenesis*, **11**, 1243–1247.

Marcus, P., *et al*. (2000). NAT2 slow acetylation and bladder cancer risk: a meta-analysis of 22 case-control studies conducted in the general population. *Pharmacogenetics*, **10(2)**, 115–122.

McLure, M., *et al*. (1989) Elevated blood levels of carcinogens in passive smokers. *American Journal of Public Health*, **79**, 1381–1384.

Miller, J. A., *et al*. (1979). Genetic studies of the lac repressor. IX. Generation of altered proteins by the suppression of nonsense mutations. *Journal of Molecular Biology*, **131**, 191–222.

Nielsen, P. S., *et al*. (1996). Environmental air pollution and DNA adducts in Copenhagen bus drivers: effect of GSTM1 and NAT2 genotypes on adduct levels. *Carcinogenesis*, **17**, 1021–1027.

Paleologo, M., *et al*. (1992). Detection of benzo(*a*)pyrene-diol-epoxide–DNA adducts in white blood cells of psoriatic patients treated with coal tar. *Mutation Research*, **281**, 11–16.

Pastorelli, R., *et al*. (1996). Hemoglobin adducts of benzo(*a*)-pyrene diolepoxide in newspaper vendors: association with traffic exhaust. *Carcinogenesis*, **17**, 2389–2394.

Peluso, M., *et al*. (1998a). 32P-postlabelling detection of aromatic adducts in the white blood cell DNA of nonsmoking police officers. *Cancer Epidemiology Biomarkers and Prevention*, **7**, 3–11.

Peluso, M., *et al*. (1998b). White blood cell DNA adducts, smoking, and NAT2 and GSTM1 genotypes in bladder cancer: a case-control study. *Cancer Epidemiology Biomarkers and Prevention*, **7**, 341–346.

Peluso, M., *et al*. (2000). White blood cell DNA adducts and fruit and vegetable consumption in bladder cancer. *Carcinogenesis*, **21**, 183–187.

Perera, F. P. (1996). Molecular epidemiology: insights into cancer susceptiblity, risk assessment, and prevention. *Journal of the National Cancer Institute*, **88**, 496–509.

Schantz, S. P., *et al*. (1997). Genetic susceptibility to head and neck cancer: interaction between nutrition and mutagen sensitivity. *The Laryngoscope*, **107**, 765–781.

Shen, I. M., *et al*. (1998). Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. *Cancer Research*, **58**, 604–608.

Spitz, M. R., *et al*. (1989). Chromosome sensitivity to bleomycin-induced mutagenesis, an independent factor for upper aerodigestive tract cancers. *Cancer Research*, **49**, 4626–4628.

Spitz, M. R., *et al*. (1994). Mutagen sensitivity as a risk factor for second malignant tumors following malignancies of the upper aerodigestive tract. *Journal of the National Cancer Institute*, **86**, 1681–1684.

Tokusashi, Y., *et al*. (1994). Absence of *p53* mutations and various frequencies of Ki-ras exon 1 mutations in rat hepatic tumours induced by different carcinogens. *Molecular Carcinogenesis*, **10**, 45–51.

Vineis, P., *et al*. (1999). *Metabolic Polymorphisms and Susceptibility to Cancer.* IARC Scientific Publication No. 148. (IARC, Lyon).

Vogelstein, B. and Kinzler, K. W. (1998). *The Genetic Basis of Human Cancer.* (McGraw-Hill, New York).

Yaeger, T. R., *et al*. (1998). Overcoming cellular senescence in human cancer pathogenesis. *Genes and Development*, **12**, 163–174.

FURTHER READING

In General on Molecular Epidemiology

Bartsch, H. (2000). Studies on biomarkers in cancer etiology and prevention: a summary and challenge of 20 years of interdisciplinary research. *Mutation Research*, **462**, 255–279.

Hulka, B., *et al*. (1990). *Biological Markers in Epidemiology.* (Oxford University Press, New York).

Perera, F. P. (2000). Molecular epidemiology: on the path to prevention? *Journal of the National Cancer Institute*, **92**, 602–612.

Schulte, P. and Perera, F. (1993). *Molecular Epidemiology. Principles and Practice.* (Academic Press, San Diego).

On Cancer Mechanisms

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Vogelstein, B. and Kinzler, K. W. (1998). *The Genetic Basis of Human Cancer.* (McGraw-Hill, New York).

On DNA Adducts

Denissenko, M. F., *et al*. (1996). Preferential formation of benzo[*a*]pyrene adducts at lung cancer mutational hotspots in *P53*. *Science*, **274**, 430–432.

Phillips, D. H. (1996). DNA adducts in human tissues: biomarkers of exposure to carcinogens in tobacco smoke. *Environmental Health Perspectives*, **104, Suppl. 3**, 453–458.

Vineis, P. and Perera, F. (2000). DNA adducts as markers of exposure to carcinogens and risk of cancer. *International Journal of Cancer*, **88**, 325–328.

On *p53*

Hussain, S. P. and Harris, C. C. (1998). Molecular epidemiology of human cancer: contribution of mutation spectra studies of tumor suppressor genes. *Cancer Research*, **58**, 4023–4037.

On Individual Susceptibility

Berwick, M. and Vineis, P. (2000). DNA repair in human cancer: an epidemiologic review. *Journal of the National Cancer Institute*, **92**, 874–897.

Shields, P. G. and Harris, C. C. (2000). Cancer risk and low-penetrance susceptibility genes in gene–environment interactions. *Journal of Clinical Oncology*, **18**, 2309–2315.

Vineis, P., *et al*. (1999). *Metabolic Polymorphisms and Susceptibility to Cancer.* IARC Scientific Publication No. 148. (IARC, Lyon).

Chapter 28

Dietary Genotoxins and Cancer

Takashi Sugimura and Keiji Wakabayashi
National Cancer Center Research Institute, Tokyo, Japan

Minako Nagao
Tokyo University of Agriculture, Tokyo, Japan

CONTENTS

- Introduction
- Microcomponents Affecting Carcinogenesis
- Macrocomponents Affecting Carcinogenesis
- Integration and Recommendation

INTRODUCTION

With regard to cancer causation, tobacco smoking, diet and infections/inflammation are three major factors (Doll and Peto, 1981). Cancer cells have multiple gene alterations (Sugimura, 1992), most of which occur owing to exposures encountered under life-style related conditions, although some of them are inherited through the germ-line. There are numerous genotoxic agents produced exogenously and endogenously. For instance, cigarette smoke contains many different types of genotoxic agents, and infection and inflammation yield reactive oxygen species and nitric oxide-related agents.

Dietary genotoxins exist in several situations: (1) as contaminants due to mould growth; (2) as edible plant components; (3) as substances formed during storage and fermentation of food, some of which are nitrosatable substances; (4) as products of cooking; and (5) as food additives, mainly of the preservative type.

Since the diet is one of three major cancer causative factors, it might be expected that large amounts of genotoxic substances are included. However, contrary to expectation, many kinds of genotoxic substance that do exist are present at very low levels, the only exception being aflatoxin B_1, where a positive association between exposure through the diet intake and liver cancer development is evident in epidemiological studies.

Nevertheless, diet is clearly of importance to human cancer occurrence. In addition to genotoxic agents present as microcomponents, macrocomponents such as fatty acids, salts and fibres have been indicated to be associated with human cancer. At the same time, many anticarcinogenic compounds are included in the diet. This chapter mainly deals with genotoxic substances and macrocomponents in the diet related to carcinogenesis. In the concluding comments, the significance of these dietary substances is described from the viewpoint of their contribution to human carcinogenic risk. (See also chapter *Mechanisms of Chemical Carcinogenesis*.)

MICROCOMPONENTS AFFECTING CARCINOGENESIS

Mycotoxins

Mycotoxins are toxic compounds produced by fungi. Typical examples are aflatoxins, metabolites formed by *Aspergillus flavus*. In 1960, numerous turkeys died in the UK, and the aetiological factor was found to be a contaminant in their diet, subsequently identified to be aflatoxins. Similarly, rainbow trout were reported to have died due to aflatoxin contamination of their diet. Among aflatoxins, aflatoxin B_1 (AFB_1), G_1 and M_1 are frequently detected in foods. The structure of AFB_1, the most potent carcinogen and mutagen of this family, is shown in **Figure 1**. AFB_1 induces hepatocellular carcinomas in

Figure 1 Structure of AFB_1.

many species of experimental animals, such as rats, monkeys and trout.

AFB_1 is metabolically activated to form its 8,9-epoxide derivative by cytochrome P-450s including CYP3A4 and 1A2, and the epoxide binds to guanine residues in DNA to produce 8,9-dihydro-8-(N^7-guanyl)-9-hydroxy-aflatoxin B_1. AFB_1 preferentially induces GC to TA transversions. In addition, G to T transversions in the third position of codon 249 of the *p53* tumour-suppressor gene in hepatocellular carcinoma have frequently been noted in inhabitants in Qidong, China and southern Africa, exposed to high levels of AFB_1. The data suggest that AFB_1 is involved in hepatocarcinogenesis in humans by inducing mutations in the *p53* gene. AFB_1 has been classified as a group 1 human carcinogen by the International Agency for Research on Cancer (IARC) (IARC, 1987). In the USA, the allowed levels of contamination of AFB_1, AFB_2 AFG_1 and AFG_2 in the diet are $< 20\,\mu g/kg$. In Japan, AFB_1 contamination is limited to $< 10\,\mu g/kg$ in the diet.

Sterigmatocystin is a toxic metabolite produced by moulds in the genera *Aspergillus* and *Penicillium* and is detected in various foods. This mycotoxin is mutagenic to *Salmonella* strains and induces hepatomas in rats when administered orally.

Pyrrolizidine Alkaloids

Mutagenic and carcinogenic pyrrolizidine alkaloids are distributed in a variety of plant species. Humans are exposed to some of them in foods and herbal remedies. Petasitenine is present in coltsfoot, *Petasites japonicus* Maxim, the young flower stalks of which have been used as a food and a herbal remedy in Japan. Senkirkine, also having an otonecine moiety, is present in another kind of coltsfoot, *Tsussilago farfara* L., the dried buds of which are taken as a herbal remedy for coughs in China and Japan. Petasitenine and senkirkine are mutagenic to *S. typhimurium* TA100 with S9 mix, and carcinogenic in rats, inducing liver tumours (Hirono *et al*., 1977). Petasitenine and senkirkine are responsible for tumour-induction by the two kinds of coltsfoot, *P. japonicus* Maxim and *T. farfara* L., respectively. The structures of petasitenine and senkirkine are shown in **Figure 2**.

Symphytine was isolated as a major alkaloid in comfrey, *Symphytum officinale* L., the leaves and roots of which are used as a green vegetable and a tonic. This alkaloid is mutagenic to *S. typhimurium* TA100 and carcinogenic in rats. Heliotrine similarly shows mutagenicity in *S. typhimurium* TA100 and causes tumours in rats. An outbreak of veno-occlusive disease which occurred in Afghanistan was due to intake of wheat flour heavily contaminated with seeds of a plant of the *Heliotropium* species containing heliotrine.

Petasitenine   Senkirkine

Figure 2 Structures of petasitenine and senkirkine.

Aquilide A/Ptaquiloside

The bracken fern, *Pteridium aquilinum*, is grown in many areas of the world and is eaten by residents of some countries, including Japan. However, haematuria with tumours of the urinary bladder are frequently observed in cows ingesting bracken fern in the fields in Turkey. Moreover, rats fed bracken fern develop tumours in the ileum, urinary bladder and mammary glands (Evans and Mason, 1965). It is also reported that a combination of papilloma virus and bracken fern could be involved in the development of the alimentary tract cancer in Scottish cattle.

The mutagenic principle in bracken fern was isolated and identified as a novel norsesquiterpene glucoside of the illudane type, named aquilide A (van der Hoeven *et al*., 1983). Since aquilide A itself is not mutagenic to *Salmonella*, but only after alkaline treatment, an aglycone of aquilide A was suggested to be the mutagenic component. By monitoring the carcinogenicity of various fractions from bracken fern in rats, the same compound was identified as the active agent and termed ptaquiloside. This substance induces multiple ileal adenocarcinomas and mammary carcinomas at high incidence in rats. The carcinogenicity of bracken fern was found to be reduced after boiling, and exposure levels of humans to aquilide A/ptaquiloside in cooked bracken fern could be small. The structure of aquilide A/ptaquiloside is shown in **Figure 3**.

Cycasin

Cycad is a plant which grows in the tropics. The inhabitants in some areas, including the Amami Oshima and Miyako Islands of Japan and Guam, have employed its nuts as a source of starch for food, but this practice is now very limited. Cycad nuts induce cancers in the liver, kidney and colon of rats when given in the diet. Cycasin, the β-D-glucoside of methylazoxymethanol, is also carcinogenic, inducing cancers in the same organs as cycad nuts (Laqueur *et al*., 1963). Cycasin is in fact not carcinogenic

Figure 3 Structure of aquilide A/ptaquiloside.

in germ-free rats and β-glucoside hydrolysis in the intestinal microflora is necessary to produce the aglycone, methylazoxymethanol, as the proximate carcinogenic form.

Alkenylbenzenes

Safrole (1-allyl-3,4-methylenedioxybenzene), a major component of sassafras oil, has been used as a fragrance in soft drinks and soaps. Estragole (1-allyl-4-methoxybenzene) is present in tarragon oil and sweet basil and is used as an essence. The structures of safrole and estragole are shown in **Figure 4**. These two alkenylbenzenes are carcinogenic in mice, and their 1′-hydroxy derivatives are mutagenic in *Salmonella* (Miller *et al*., 1983). In addition, three naturally occurring alkenylbenzenes, isosafrole (3,4-methylenedioxy-1-propenylbenzene), methyleugenol (1-allyl-3,4-dimethoxybenzene) and β-asarone (*cis*-1-propenyl-2,4,5-trimethoxybenzene), have been shown to be carcinogenic in rodents.

Mushroom Hydrazines

The cultivated edible mushroom *Agaricus bisporus* contains agaritine, β-*N*-[γ-L(+)-glutamyl]-4-hydroxymethyl-phenylhydrazine, and its decomposition products. Three hydrazine derivatives (the *N*′-acetyl derivative of 4-hydroxymethylphenylhydrazine, 4-methylphenylhydrazine hydrochloride and the tetrafluoroborate form of 4-(hydroxymethyl) benzenediazonium ion), derived from agaritine, are carcinogenic in mice. Moreover, uncooked cultivated *Agaricus bisporus* itself is carcinogenic in mice (Toth and Erickson, 1986). Agaritine is also present in the shiitake (*Cortinellus shiitake*), which is a popular edible mushroom in Japan. Another edible mushroom, false morel (*Gyromitra esculenta*), contains gyromitrin (acetaldehyde methylformylhydrazone). This compound is converted into the mutagenic and carcinogenic *N*-methyl-*N*-formylhydrazine and methylhydrazine under acidic conditions such as those prevailing in the stomach.

Safrole Estragole

Figure 4 Structures of safrole and estragole.

Flavonoids

Numerous kinds of flavonoids are present in plants. Among those, quercetin and its glycosides are the most common flavonoids, distributed in vegetables and fruits. Their daily intake by humans is estimated to be more than 100 mg per person. Many flavonoids including quercetin, kaempferol and isorhamnetin show mutagenicity in *Salmonella* with and without metabolic activation systems, but this is not the case with their glycoside compounds, which require digestion by glycosidase for mutagenic activity (Nagao *et al*., 1978).

Quercetin is mutagenic not only in bacteria *Salmonella* strains but also in some *in vitro* mutation test systems with Chinese hamster V79 cells and Chinese hamster lung cells. However, many studies have shown no carcinogenicity of quercetin in rodents such as rats, mice and hamsters, even with 10% in the diet. Moreover, the conclusion of the NTP Technical Report on toxicology and carcinogenesis studies of quercetin in F344 rats was that there is no clear evidence of carcinogenic activity of quercetin, the reported increased incidence of renal tubular cell adenomas being possible due to involvement of α_{2u}-globulin in male rats (NTP, 1991; http://ntp-server.niehs.nih.gov/htdocs/LT-studies/tr409.html). This is not considered relevant to human carcinogenesis, and therefore the risk potential of quercetin for human cancer must be negligible. These studies with quercetin offer a very important warning: it is not necessarily the case that mutagens are carcinogens.

Nitrosamines and Nitrosatable Mutagens and Carcinogen Precursors

Sodium nitrite has been used as a food preservative and colouring substance in meat. Severe liver disease was encountered in sheep fed a diet containing fishmeal preserved with nitrite and subsequently the toxic principle was identified as *N*-nitrosodimethylamine (Sakshaug *et al*., 1965). Vegetables are a major source of nitrite. Nitrite is also produced from nitrate by bacteria in the oral cavity. Under gastric acidic conditions, nitrite reacts with secondary amines to produce mutagenic and carcinogenic *N*-nitrosodialkylamines. Moreover, reactive nitric oxide

produced from L-arginine by nitric oxide synthase in inflammatory processes is involved in the generation of *N*-nitroso compounds. It has been reported that beer may contain *N*-nitrosodimethylamine. Oxides of nitrogen are produced during direct-fire drying of barley malt, and react with the barley malt alkaloids gramine and hordenine, both of which have a dimethylamine moiety as a common structure. However, improvements in drying malt have significantly decreased the levels of *N*-nitrosodimethylamine in beer.

In addition to secondary amines, nitrite reacts with indole and phenol derivatives to produce mutagenic compounds under acidic conditions (Wakabayashi *et al*., 1989). 1-Methyl-1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid in soy sauce and 4-chloro-6-methoxyindole in fava beans are examples of nitrosatable indole compounds, but carcinogenicity has not been proved. Tyramine, present in various fermented foods such as soy sauce, cheese and soybean paste, gives rise to 4-(2-aminoethyl)-6-diazo-2,4-cyclohexa-dienone (3-diazotyramine) after nitrite treatment. 3-Diazotyramine is mutagenic in *Salmonella* and carcinogenic in rats. Relatively large amounts of nitrite are required for the formation of diazo compounds *in vitro* and presumably also *in vivo* in the stomach.

Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs), which have been identified as carcinogenic compounds in coal tar, are in fact widely distributed in our environment such as in cigarette smoke, exhaust gas and cooked foods. The presence of PAHs including benzo[*a*]pyrene in charred parts of biscuits, roasted coffee beans and broiled steak has been reported. The average levels of benzo[*a*]pyrene is 8 ng per gram of broiled steak (Lijinsky and Shubik, 1964). Benzo[*a*]pyrene is converted into the ultimate diol epoxide derivative by metabolic activation, producing adducts at the N^2 position of the guanine residue in DNA. A high frequency of mutations in the *p53* tumour-suppressor gene due to G to T transversion is characteristic of mouse skin tumours induced by benzo[*a*]pyrene. Moreover, the aryl hydrocarbon receptor is required for its tumour induction. (See chapter on *Mechanisms of Chemical Carcinogenesis*.)

Heterocyclic Amines

Humans have used heat for cooking foods for over 500 000 years. Widmark (1939) reported carcinogenic activity of a solvent extract of broiled horse meat, with induction of mammary tumours in mice by painting on the skin. Production of mutagens by heating meat and fish was established in the 1970s by the use of bacterial mutagenesis assays (Sugimura *et al*., 1977). Methanol extracts from charred parts of grilled sun-dried sardine, beefsteak and hamburger were found to contain mutagens. Subsequent studies clarified that the mutagenic activity is mainly derived from heterocyclic amine (HCA) compounds which are produced during cooking from the meat constituents, creatin(in)e, amino acids and sugars.

Structures of mutagenic HCAs in cooked foods and preparations of heated food protein (soybean globulin) and heated amino acids have been determined. All have nitrogen(s) within aromatic rings and exocyclic amino groups, and they can be divided into five groups based on their structures: imidazoquinolines, imidazoquinoxalines, phenylimidazopyridines, pyridoindoles and dipyridoimidazoles (**Figure 5**).

Mutagenic activities of HCAs together with other typical carcinogens to *S. typhimurium* TA98 and TA100 are summarized in **Table 1**. 2-Amino-1-methyl-6-phenylimidazo-[4,5-*b*]pyridine (PhIP), 2-amino-3,8-dimethylimidazo[4,5-*f*]quinoxaline (MeIQx), 2-amino-3,4-dimethylimidazo[4,5-*f*]quinoline (MeIQ), 2-amino-3-methylimidazo[4,5-*f*]quinoline (IQ) and 2-amino-9*H*-pyrido[2,3-*b*]indole (AαC) also proved to be mutagenic *in vivo* in *lacI* or *lacZ* transgenic animals that were established by introducing the gene on a lambda shuttle vector. Some of these HCAs are mutagenic in mammalian cells, including human fibroblasts *in vitro*.

HCAs are *N*-hydroxylated by cytochrome P-450s, mainly CYP1A2, and esterified by acetyltransferase, sulfotransferase and others. All of the eight HCAs so far examined form DNA adducts by covalent binding to the C8 position of guanine. There is a linear correlation between DNA adduct levels and exposure in experimental animals, and guanine adducts of PhIP and MeIQx have been detected in human samples. These adducts result in mutations on cell division. (See chapter on *Mechanisms of Chemical Carcinogenesis*)

Carcinogenicities of these compounds in rodents are summarized in **Table 2**, along with TD_{50}s, the daily amounts required for induction of tumours with 50% prevalence per kilogram. It is worth noting that HCAs induce tumours in various organs, including the colon, mammary glands, lung, liver, bladder, prostate, blood vessels, haematopoietic system, forestomach and skin (Sugimura, 1997). Different species of animals and different strains of the same species show different susceptibilities of the various organs. PhIP and MeIQx are most abundant HCAs in foods, followed by AαC, average human daily intakes of these HCAs being around 1 μg per person per day. It is evident that HCAs may not be sufficient to induce cancers in humans by themselves. However, some epidemiological studies have indicated higher relative risks for colon, breast and bladder cancer in consumers of well-done meat, although contradictory data are also available. Production of mutations involved with any genomic instability may play an important role in carcinogenesis. Collaborative effects with other genotoxic agents may result in effective accumulation of mutations in cells.

HCA carcinogenesis is modified by various food factors, such as soybean isoflavonoids, chlorophyllin, diallyl

Trp-P-1
3-Amino-1,4-dimethyl-5*H*-pyrido[4,3-*b*]indole

Trp-P-2
3-Amino-1-methyl-5*H*-pyrido[4,3-*b*]indole

Glu-P-1
2-Amino-6-methyldipyrido[1,2-*a*:3′,2′-*d*]imidazole

Glu-P-2
2-Aminodipyrido[1,2-*a*:3′,2′-*d*]imidazole

AαC
2-Amino-9*H*-pyrido[2,3-*b*]indole

MeAαC
2-Amino-3-methyl-9*H*-pyrido[2,3-*b*]indole

IQ
2-Amino-3-methylimidazo[4,5-*f*]quinoline

MeIQ
2-Amino-3,4-dimethylimidazo[4,5-*f*]quinoline

MeIQx
2-Amino-3,8-dimethylimidazo[4,5-*f*]quinoxaline

PhIP
2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine

Figure 5 Structures of HCAs.

disulfide, docosahexaenoic acid, indole-3-carbinol tea cathechin and conjugated linoleic acids (CLAs). Since smaller amounts of HCAs are produced at lower cooking temperatures, a decrease in HCA exposure can be attained by modification of the way of cooking.

Dioxins

Contamination of food by dioxins, polychlorinated dibenzo-*p*-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs), has received much attention from the general public. 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (2,3,7,8-TCDD) is recognized as the most toxic compound among 210 congeners of PCDDs and PCDFs. To aid in the interpretation of the complex database and evaluation of the risk of exposure to mixtures of structurally related PCDDs and PCDFs, the concept of toxic equivalency factors (TEFs) is widely used, and exposure levels are frequently expressed as toxic equivalents (TEQ). TCDD itself is not genotoxic and its lethal toxicity and teratogenicity emerge through binding to the Ah receptor which is expressed in the liver and other organs in adults of various animals and during embryonic development in mice. Lethal doses in 50% of the animals tested (LD_{50} values) for 2,3,7,8-TCDD in the guinea pig and C57BL/6 and DBA2/2J mice are 0.6, 181 and 2570 $\mu g\,kg^{-1}$, respectively. Guinea pig is among the most sensitive species to 2,3,7,8-TCDD.

2,3,7,8-TCDD induces cancers in the skin, liver, haematopoietic system and lung of mice, the hard palate, nasal turbinates, tongue, liver, thyroid and lungs of rats and the skin of hamsters. Dose-related trend was observed with

oral administration of 2,3,7,8-TCDD twice a week for 104 weeks at 0.01, 0.05 and 0.5 μg kg^{-1} body weight (bw) in mice and rats (IARC, 1997).

2,3,7,8-TCDD accumulates mainly in the liver and the fat tissue. Its half-life in humans is about 10 years. Daily intakes of Europeans and Japanese are estimated to be around 120–160 pg TEQ, the major sources being fish and dairy products. Based on the data from animal experiments and considerations of mechanism of action, and epidemiological data from accidental exposure of humans to TCDD, the IARC concluded that 2,3,7,8-TCDD is 'carcinogenic to humans with only limited evidence.'

The ligand (TCDD)-activated Ah receptor forms complexes with the Arnt protein and functions as a transcription factor by binding to DNA. The enzymes CYP1A1, 1B1 and 1A2 that are involved in metabolic activation of chemical carcinogens and metabolism of hormones which have promotive effects on cancer development are among those genes that are transcriptionally activated by the Ah receptor. The activated Ah receptor also functions without DNA binding, activating the signal pathway of Src, with increased expression of *ras*, *erbA*, c-*fos*, c-*jun* and AP-1 which may result in cell growth. It also inhibits apoptosis and suppresses immune surveillance, resulting in expansion of preneoplastic cell populations. Mechanisms of TCDD carcinogenicity thus appear complex (IARC, 1997).

Table 1 Mutagenicities of HCAs and typical carcinogens in *Salmonella typhimurium* TA98 and TA100

	Revertants (per μg)	
Compound	**TA98**	**TA100**
MeIQ	661000	30000
IQ	433000	7000
MeIQx	145000	14000
Trp-P-2	104200	1800
Glu-P-1	49000	3200
Trp-P-1	39000	1700
Glu-P-2	1900	1200
PhIP	1800	120
AαC	300	20
MeAαC	200	120
Aflatoxin B_1	6000	28000
Benzo[a]pyrene	320	660
MNNG	0.00	870
N-Nitrosodimethylamine	0.00	0.23

MNNG: *N*-Methyl-*N*′-nitro-*N*-nitrosoguanidine.

MACROCOMPONENTS AFFECTING CARCINOGENESIS

Total Calories

Nutrition plays important roles in carcinogenesis through a variety of mechanisms. It is well established that excess calorie intake results in high incidences of spontaneously and chemically induced tumours in various organs in rats and mice, including the mammary glands, colon, lung, haematopoietic system and skin. Similar effects have also been suggested in humans by epidemiological studies, e.g. regarding incidence and mortality rates for cancers of the breast, colon, rectum, uterus and kidney in women and

Table 2 TD_{50} and target organs of HCA carcinogenesis in mice and rats

	Mice (CDF_1)		Rats (F344)	
	TD_{50} (mg kg^{-1} per day)	**Target organ**	**TD_{50} (mg kg^{-1} per day)**	**Target organ**
IQ	14.7	Liver, forestomach, lung	0.7	Liver, small and large intestines, Zymbal glands, clitoral gland, skin
MeIQ	8.4	Liver, forestomach	0.1	Zymbal glands, large intestine, skin, oral cavity, mammary glands
MeIQx	11.0	Liver, lung, haematopoietic system	0.7	Liver, Zymbal glands, clitoral gland, skin
PhIP	64.6	Lymphoid tissue	2.2	Large intestine, mammary glands, prostate
Trp-P-1	8.8	Liver	0.1	Liver
Trp-P-2	2.7	Liver	–	Urinary bladder, lymphoid tissue
Glu-P-1	2.7	Liver, blood vessels	0.8	Liver, small and large intestines, Zymbal glands, clitoral gland
Glu-P-2	4.9	Liver, blood vessels	5.7	Liver, small and large intestines, Zymbal glands, clitoral gland
AαC	15.8	Liver, blood vessels	–	
MeAαC	5.8	Liver, blood vessels	6.4	Liver

cancers of the colon, rectum, kidney and central nervous system in men.

Furthermore, delay in tumorigenesis by energy restriction has been observed in many models including nullizygous *p53* knockout mice, which show high rates of spontaneous neoplastic changes. Oestrogen-induced pituitary tumours were abolished by 40% energy restriction in F344 rats. However, no such effects were observed with ACI rats, suggesting that the antitumorigenic actions of energy restriction are strongly affected by the genetic background.

Through induction of a nitric oxide synthase isoform, energy restriction results in reduced generation of nitric oxide (NO), which is converted to reactive nitrogen oxide species (RNOS) such as N_2O_3 and peroxynitrite, by reaction with oxygen or superoxide. RNOS or NO induce(s) DNA strand breaks, 8-nitroguanine formation and deamination of guanine, cytosine and adenine in DNA. Peroxynitrite and NO are mutagenic, and RNOS inhibit various enzymes including DNA repair enzymes by reacting with their cysteine residues. Calorie restriction also reduces chemical induction of somatic mutations *in vivo*, although effects on oxidative DNA damage are limited.

Another mechanism of action involves suppression of cell proliferation with decrease in serum level of insulin-like growth factor I (IGF-I), and expression levels of EGF and cyclin D1, as well as upregulation of the cell-cycle growth arrest factor p27. The fidelity of DNA polymerases is also up-regulated and the decrease in cellular DNA repair capacity which occurs with ageing is to some extent prevented. All of these effects are highly implicated in suppression of tumour development.

Fat Intake

Dietary fat is thought to be one of the main risk factors for cancer development on the basis of positive correlations observed between intake and incidences of tumours of the breast, colon and prostate (Schottenfeld and Fraumeni, 1996). Enhancing effects of fat could be partly explained by the included calories, but many experiments under isocaloric conditions have demonstrated an importance of fats themselves as genotoxic and cell-proliferation stimulating agents.

Experimental animal studies have demonstrated that high intake of dietary fat increases the incidence of cancer of the breast, colon and prostate, with influences on many steps in the tumorigenic process, including initiation, promotion and metastasis. In general, the ω-6 polyunsaturated fatty acids (PUFAs), such as linoleic acid (C18:2) and arachidonic acid (C20:4), strongly and the saturated fats weakly enhance tumour development. However, ω-3 PUFAs such as docosahexaenoic acid (DHA, C22:6), eicosapentaenoic acid (EPA, C20:5) and α-linolenic acid (C18:3) generally have a tumour-protective effects while monounsaturated fats, such as oleic acid (C18:1 ω-9), have no significant influence. It is noteworthy, however, that several chemopreventive agents such as α-tocopherol and various classes of phenolics showing potent antioxidant activities are included as major components of the main source of ω-9 fatty acids, olive oil. Conjugated linoleic acid (CLA), present in dairy products, prevents mammary and colon cancer development in experimental animals. Linoleic acid has two double bonds at the 9 and 12 and CLA at the 9 and 11, 10 and 12 or 11 and 13 positions, being mainly present in the 9 and 11 form in food.

In fat-related carcinogenesis, induction of DNA damage by peroxidation products of PUFAs is one mechanism of action. This involves the generation of reactive α,β-unsaturated aldehydes. Thus, linoleic acid and arachidonic acid are sources of the major reactive aldehydes, *trans*-4-hydroxy-2-non-enal and malondialdehyde, which can form promutagenic exocyclic DNA adducts, such as M_1G (3-(2′-deoxy-β-D-*erythro*-pentofuranosyl)pyrimido[1,2-*a*]purin-10(3*H*)-one). *trans*-4-Hydroxy-2-non-enal is readily oxidized by fatty acid peroxides to form 2,3-epoxy-4-hydoxynonanal and the latter compound modifies DNA by forming etheno adducts such as 1,N^6-ethenoadenine, N^2,3-ethenoguanine and 3,N^4-ethnocytosine, which are highly miscoding.

The essential fatty acid linoleic acid is converted into arachidonic acid and then, with the aid of cycloxygenase and lipoxygenase, into prostaglandins, thromboxanes and leukotrienes that show hormone-like activity. One of them, prostaglandin E_2 (PGE_2), was recently shown to be involved in colon carcinogenesis using knockout mice for EP_1, one of the receptors for PGE_2. ω-6 PUFAs induce ornithine decarboxylase activity in colonic mucosa, activation of protein kinase C in mammary glands, and increases in number of oestrogen receptor binding sites, resulting in cell proliferation.

Oleic acid and ω-3 PUFAs, and specifically EPA, block the Δ^6-desaturase reaction, the first step from linoleic acid to eicosanoids. EPA, DHA and CLA suppress the arachidonic acid pathway by inhibiting cyclooxygenase activity. Stimulation of genotoxic NO production in a murine macrophage cell line with lipopolysaccharide is also suppressed by DHA, EPA and α-linolenic acid. Inhibition of inducible nitric oxide synthase gene expression may thus contribute to the cancer preventive activity of ω-3 PUFAs.

Sodium Chloride

A positive correlation between daily salt intake and gastric cancer incidence was noted by epidemiologists some time ago and urinary sodium chloride concentrations are well correlated with stomach cancer mortality. In experimental animal model systems, *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG) induces gastric cancers and this is enhanced by salt administration. High doses of salt disrupt the mucin layer covering and protecting the gastric

epithelium and further damage epithelial cells by generation of a high osmotic pressure. This in turn stimulates proliferation of stem cells of the gastric epithelium, providing favourable conditions for the occurrence of mutations. Prolonged damage results in chronic atrophic gastritis and intestinal metaplasia, both of which are understood to be precursor lesions for intestinal-type gastric cancer.

INTEGRATION AND RECOMMENDATION

It is very clear that genotoxic substances in the diet are related to human cancer development. Laboratory data offer particularly strong support.

In this chapter, information on genotoxic carcinogens in the diet and the significance of other nutritional factors have been concisely summarized. Carcinogenesis is due to the accumulation of multiple genetic changes in a cell, implying multiple steps. Each step could be caused by exposure to a minute amount of a genotoxic substance in the diet. Amounts of individual substances are generally not sufficiently high to induce human cancer alone, one exception being human exposure to contaminant mycotoxins in limited regions of the world. However, since many kinds of genotoxic compounds exist in the diet, efforts to lessen their combined impact should be made wherever possible. Since mutations may result in genomic instability, even a single such genetic alteration, for example, in a gene involved in DNA repair could produce spontaneous accumulation of further mutations with time.

Endogeneous genotoxic agents (autobiotics) such as reactive oxygen species are produced in line with the calorific intake. Consumption of appropriate amounts of substances working as radical scavengers, such as vitamin E, is essential for protection from carcinogenesis. Overintake of calories, fat and salt may act to enhance carcinogenesis through various pathways, while more fibre in the diet can suppress colon carcinogenesis. By integrating information described in this chapter, recommendations can be given with reference to eight of the 12 proposals from the National Cancer Center, Tokyo (Sugimura, 1986). These are: (1) have a nutritionally balanced diet, (2) eat a variety of types of food, (3) avoid excess calories, especially as fat, (4) avoid the excessive drinking of alcohol, (5) take vitamins in appropriate amounts and eat fibre and green and yellow vegetables rich in carotene, (6) avoid drinking fluids that are too hot and eating foods that are too salty, (7) avoid the charred parts of cooked food and (8) avoid food with possible contamination by fungal toxins. Similar recommendations were also proposed by the World Cancer Research Fund and the American Institute for Cancer Research (1997). It is noteworthy that, even for cancer-preventive compounds, the amount of intake should be appropriate, because it is reported that over-intake of β-carotene enhances carcinogenesis (The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group, 1994).

REFERENCES

Doll, R. and Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute*, **66**, 1191–1308.

Evans, I. A. and Mason, J. (1965). Carcinogenic activity of bracken. *Nature*, **208**, 913–914.

Hirono, I., *et al*. (1977). Carcinogenic activity of petasitenine, a new pyrrolizidine alkaloid isolated from *Petasites japonicus* Maxim. *Journal of the National Cancer Institute*, **58**, 1155–1157.

IARC (1987). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans; Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42*. IARC Scientific Publications Supplement 7, 83–87 (IARC, Lyon).

IARC (1997). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans; Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans*, Vol. 69. (IARC, Lyon).

Laqueur, G. L., *et al*. (1963). Carcinogenic properties of nuts from *Cycas circinalis* L. indigenous to Guam. *Journal of the National Cancer Institute*, **31**, 919–951.

Lijinsky, W. and Shubik, P. (1964). Benzo[*a*]pyrene and other polynuclear hydrocarbons in charcoal-broiled meat. *Science*, **145**, 53–55.

Miller, E. C., *et al*. (1983). Structure–activity studies of the carcinogenicities in the mouse and rat of some naturally occurring and synthetic alkenylbenzene derivatives related to safrole and estragole. *Cancer Research*, **43**, 1124–1134.

Nagao, M., *et al*. (1978). Environmental mutagens and carcinogens. *Annual Review of Genetics*, **12**, 117–159.

NTP (1991). *NTP Technical Report (No. 409) on the Toxicology and Carcinogenesis Studies of Quercetin in F344/N Rats*. NIH Publication No. 92–3140. (US Department of Health and Human Services, Public Health Service, National Toxicology Program, Research Triangle Park, NC).

Sakshaug, J., *et al*. (1965). Dimethylnitrosamine: its hepatotoxic effect in sheep and its occurrence in toxic batches of herring meal. *Nature*, **206**, 1261–1262.

Schottenfeld, D. and Fraumeni, J. F., Jr. (1996). *Cancer Epidemiology and Prevention*. (Oxford University Press, New York).

Sugimura, T. (1986). Studies on environmental chemical carcinogenesis in Japan. *Science*, **18**, 312–318.

Sugimura, T. (1992). Multistep carcinogenesis: a 1992 perspective. *Science*, **23**, 603–607.

Sugimura, T. (1997). Overview of carcinogenic heterocyclic amines. *Mutation Research*, **376**, 211–219.

Sugimura, T., *et al*. (1977). Mutagen–carcinogens in foods, with special reference to highly mutagenic pyrolytic products in

broiled foods. In: Hiatt, H. H., *et al*. (eds). *Origins of Human Cancer*. 1561–1577 (Cold Spring Harbor Laboratory, New York).

The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group (1994). The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *New England Journal of Medicine*, **330**, 1029–1035.

Toth, B. and Erickson, J. (1986). Cancer induction in mice by feeding of the uncooked cultivated mushroom of commerce *Agaricus bisporus*. *Cancer Research*, **46**, 4007–4011.

van der Hoeven, J. C. M., *et al*. (1983). Aquilide A, a new mutagenic compound isolated from bracken fern (*Pteridium aquilinum* (L.) Kuhn). *Carcinogenesis*, **4**, 1587–1590.

Wakabayashi, K., *et al*. (1989) Mutagens and carcinogens produced by the reaction of environmental aromatic compounds with nitrite. *Cancer Surveys*, **8**, 385–399.

Widmark, E. M. (1939). Presence of cancer-producing substances in roasted food. *Nature*, **143**, 984.

World Cancer Research Fund and American Institute for Cancer Research (1997). *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. (American Institute for Cancer Research, Washington, DC).

FURTHER READING

Nagao, M. and Sugimura, T. (eds) (2000). *Food Borne Carcinogens, Heterocyclic Amines*. (Wiley, Chichester).

Sugimura, T. (2000). Nutrition and dietary carcinogens. *Carcinogenesis*, **21**, 387–395.

Wakabayashi, K. and Sugimura, T. (1998). Heterocyclic amines formed in the diet: carcinogenicity and its modulation by dietary factors. *Journal of Nutritional Biochemistry*, **9**, 604–612.

Chapter 29

Tobacco Use and Cancer

Stephen S. Hecht
University of Minnesota Cancer Center, Minneapolis, MN, USA

CONTENTS

- Introduction
- Epidemiology of Tobacco and Cancer
- Tumour Induction in Laboratory Animals
- Chemistry of Tobacco Smoke
- Mechanisms of Tumour Induction
- Chemoprevention of Tobacco-Related Cancer
- Conclusion

INTRODUCTION

Worldwide tobacco use is staggering. According to estimates by the World Health Organisation (WHO), there are about 1100 million smokers in the world, representing approximately one-third of the global population aged 15 years or older (WHO, 1997). China alone has approximately 300 million smokers, about the same number as in all developed countries combined. **Table 1** summarizes the estimated number of smokers in the world, according to data available in the early 1990s. Globally, about 47% of men and 12% of women smoke. Smoking prevalence varies widely by country. For example, in Korea, 68% of men smoke daily, whereas the corresponding figure for Sweden is 22%. Male smoking prevalence varies from $< 30\%$ in the African Region (as defined by WHO) to 60% in the Western Pacific Region. Among women, the highest smoking prevalence is in Denmark, where 37% of women smoke, whereas in many Asian and developing countries, prevalence is reported to be $< 10\%$. Smoking among women is common in the former socialist countries of Central and Eastern Europe (28%), countries with established market economies (23%) and Latin American and Caribbean countries (21%) (WHO, 1997). **Figure 1** summarizes data on the number of cigarettes smoked per day per daily smoker in different regions of the world. Although smoking prevalence is lower in the less developed countries in general, it is expected that this will increase markedly as smoking takes hold and larger numbers of young smokers grow older.

Table 1 Estimated number of smokers (in millions) in the world (early 1990s) (from WHO, 1997)

Countries	Males	Females	Total
Developed countries	200	100	300
Developing countries	700	100	800
World	900	200	1100

Cigarettes are the main type of tobacco product worldwide (WHO, 1997). Manufactured cigarettes are available in all countries, but in some areas of the world roll-your-own cigarettes are still popular. Other smoked products include ‘kreteks,’ which are clove-flavoured cigarettes popular in Indonesia, and ‘sticks’ which are smoked in Papua New Guinea. ‘*Bidis*,’ which consist of a small amount of tobacco wrapped in temburni leaf and tied with a string, are very popular in India and neighbouring areas, and have recently taken hold in the USA. Cigars are currently increasing in popularity, and pipes are still used. A substantial amount of tobacco is consumed worldwide in the form of smokeless tobacco products. These include chewing tobacco, dry snuff used for nasal inhalation, moist snuff which is placed between the cheek and gum, a popular practice in Scandinavia and North America, and ‘*pan*’ or betel quid, a product used extensively in India. **Table 2** summarizes the estimated annual global consumption of various types of tobacco products. About six $\times 10^{12}$ cigarettes were consumed annually in the period 1990–1992 worldwide (WHO, 1997).

Most global tobacco manufacturing is controlled by a small number of state monopolies and multinational

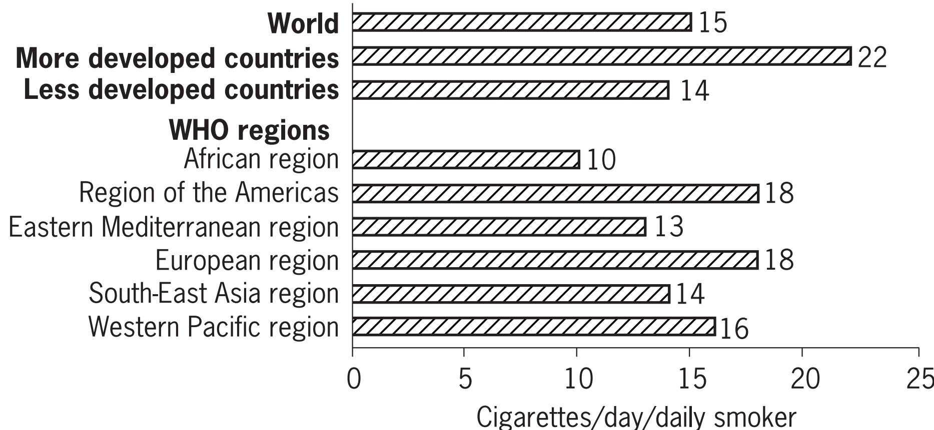


Figure 1 Number of cigarettes smoked per day per daily smoker, by region.

Table 2 Estimated annual global consumption of various types of tobacco products, 1990–1992 (from WHO, 1997)

Type of tobacco product	Estimated annual global consumption, 1990–1992 (kg × 10^9)	Estimated proportion of total global consumption (%)
All cigarettes, consisting of:	4.2-5.5	65-85
Manufactured cigarettes and kreteks	3.9-5.2	60-80
Bidis	0.2	3
Roll-your-owns	0.1	2
All other tobacco products	1.0-2.3	15-35
Al tobacco products	6.5	100

Table 3 Estimated percentage of deaths caused by smoking in 1995, all developed countries, by gender, age and major cause of death groupings (from WHO, 1997)

Gender	Age (years)	All causes	All cancer	Lung cancer	Upper aerodigestive cancer[a]	Other cancer	Chronic obstructive pulmonary diseases	Other respiratory diseases	Vascular diseases	Other causes
Men	35-69	36	50	94	70	18	82	29	35	35
	70+	21	36	91	59	13	73	11	12	12
	All ages	25	43	92	66	15	75	14	21	18
Women	35-69	13	13	71	34	2	55	16	12	15
	70+	8	13	74	38	2	54	7	5	6
	All ages	9	13	72	36	2	53	7	6	7
Both	35-69	28	35	89	65	10	73	25	28	27
	70+	13	25	86	52	7	65	9	8	8
	All ages	17	30	87	60	8	66	10	13	12

[a] Cancers of the mouth, oesophagus, pharynx and larynx.

corporations (WHO, 1997). The largest state monopoly is in China, where 1.7×10^{12} cigarettes were sold in 1993. This represents about 31% of the global market, about the same as the three largest multinational tobacco corporations combined. China produces about 36% of the world's tobacco, and the USA about 11%.

Worldwide, smoking is estimated to have caused about 1.05 million cancer deaths in 1990 (WHO, 1997). About 30% of all cancer deaths in developed countries are caused by smoking (**Table 3**). The corresponding figure for developing countries is 13%. Lung cancer is the dominant malignancy caused by smoking, with 514 000 lung cancer deaths attributed to smoking in developed countries in 1995 (WHO, 1997). Smoking also causes other types of cancer. This is discussed further below. Lung cancer was rare at the beginning of the twentieth century (Anonymous,

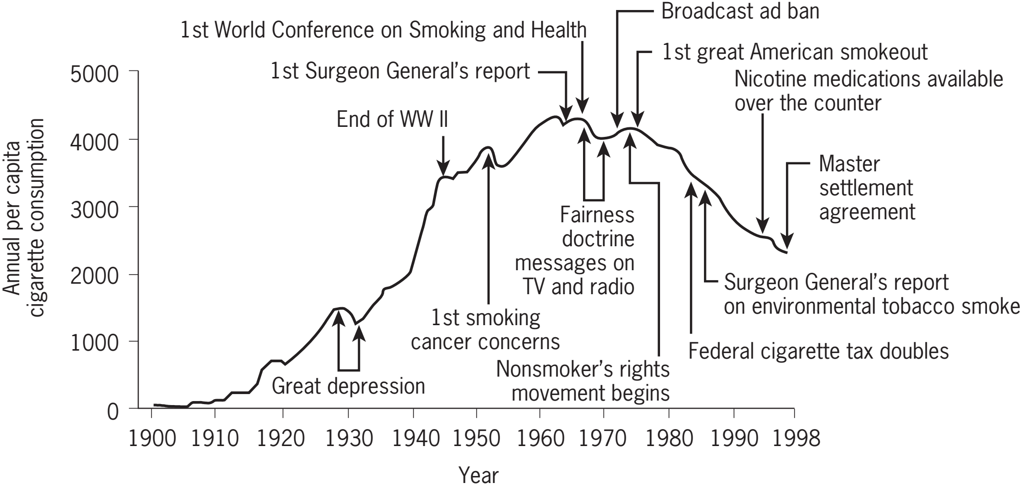


Figure 2 Annual adult per capita cigarette consumption and major smoking and health events, USA, 1900-1988.

1999). However, the incidence and death rates increased as smoking became more popular. In the USA, the lung cancer death rate in 1930 for men was 4.9 per 100 000. By 1990, this had increased to 75.6 per 100 000. The lung cancer death rate can be shown to parallel the curves for cigarette smoking prevalence, with an approximate 20-year lag time. In 1964, the first Surgeon General's report on the health consequences of cigarette smoking was published. Following this landmark report, smoking prevalence began to decrease in the USA. **Figure 2** summarizes annual adult per capita cigarette consumption and major smoking and health events in the USA in the twentieth century. Although smoking prevalence in the USA decreased from 1964 to 1990, there has been no overall change since then. There are still 48 million adult smokers in the USA (Anonymous, 1999).

EPIDEMIOLOGY OF TOBACCO AND CANCER

How do we know that smoking causes cancer? The strongest evidence comes from prospective epidemiological studies. The prospective study design is powerful. Individuals are asked questions about their lifestyle and other factors, then followed for long periods of time. Cancer deaths are recorded and relative risks are determined with respect to the answers given on the questionnaires. Therefore, the answers cannot be influenced by disease state, bias and many other confounding factors. Large studies of this type conclusively demonstrate that all three main types of lung cancer – squamous cell, small cell and adenocarcinoma – are caused mainly by tobacco smoking (IARC, 1986; Shopland, 1995; Blot and Fraumeni, 1996). These studies have consistently demonstrated a dose–response relationship between numbers of cigarettes consumed daily and relative risk of lung cancer. These data, which are based on over 20 million person years of observation, are summarized in **Table 4** (IARC, 1986; Shopland, 1995; Blot and Fraumeni, 1996).

Case-control epidemiological studies ask questions of people with lung cancer and corresponding matched controls without cancer. Although this study design is not as powerful as the prospective method, it is more economical. Hundreds of case-control studies consistently demonstrate a higher risk for lung cancer in smokers than non-smokers.

The incidence of lung cancer also depends on the duration of smoking (IARC, 1986; Shopland, 1995; Blot and Fraumeni, 1996). The greatest risk is among those who start smoking in adolescence and continue throughout their lives. Cessation of smoking gradually decreases the risk of lung cancer, but no change is seen during the first 4 years. Data on the relative risk of developing lung cancer by time since stopping and total duration of smoking are summarized in **Table 5** (Blot and Fraumeni, 1996).

While cigarette smoking has been definitively established as the major cause of lung cancer, a number of issues remain to be investigated (Wynder and Hoffmann, 1994). Since 1975, the incidence of adenocarcinoma has greatly increased in the USA. The ratio of adenocarcinoma to squamous cell carcinoma was 1 : 2.3 among white males in 1969–1971, whereas it was 1 : 1.4 in 1984–1986. Adenocarcinoma now exceed squamous cell carcinoma of the lung in the USA. This changing histology of cigarette smoke-induced lung cancer is due to changes in cigarette design, particularly the introduction of lower nicotine and higher *N*-nitrosamine-containing cigarettes. There are important geographical and ethnic differences in lung cancer which require investigation. Lung cancer incidence in Japan is considerably lower than would be expected by comparison with US rates in spite of a dramatic increase in smoking. Diet may be one factor which enhances risk in the USA; laboratory studies have shown that a high-fat diet similar to that consumed in the USA can enhance lung cancer induction by constituents of tobacco smoke. In the

Table 4 Lung cancer mortality ratios in men and women, by number of cigarettes smoked daily, major prospective mortality studies. (From Shopland, 1995, *Environmental Health Perspectives*, **103**, 131–142.)

	Men		Women	
Study, population	**Cigarettes/day**	**Ratio**	**Cigarettes/day**	**Ratio**
American Cancer Society 25-State study, 1 million	Nonsmokers	1.00	Nonsmokers	1.00
	All smokers	8.53	All smokers	3.58
	1-9	4.62	1-9	1.30
	10-19	8.62	10-19	2.40
	20-39	14.7	20-39	4.90
	40 +	18.7	40 +	7.50
British Physicians' Study, 40 000	Nonsmokers	1.00	Nonsmokers	1.00
	All smokers	14.9	All smokers	5.00
	1-14	7.50	1-14	1.28
	15-24	14.9	15-24	6.41
	25 +	25.4	25 +	29.7
US veterans study, 290 000	Nonsmokers	1.00		
	All smokers	11.3		
	1-9	3.89		
	10-19	9.63		
	21-39	16.7		
	40 +	23.7		
Japanese study, 270 000	Nonsmokers	1.00	Nonsmokers	1.00
	All smokers	3.76	All smokers	2.03
	1-9	2.06	1-9	2.25
	10-19	4.00	10-19	2.56
	20 +	6.24	20 +	4.47
American Cancer Society 50-State Study, 1.2 million	Nonsmokers	1.00	Nonsmokers	1.00
	All smokers	22.4	All smokers	11.9
	1-20	18.8	1-10	5.50
	20 +	26.9	11-19	11.2
			20	14.2
			21-30	20.4
			31 +	22.0

Table 5 Relative risk of developing lung cancer according to years since quitting smoking among males in three cohort studies of smokers[a]. (From Blot and Fraumeni, 1996, *Cancer Epidemiology and Prevention*, **637–665**, Oxford University Press, New York.)

	Years since quitting smoking					
Cohort	**0**	**1–4**	**5–9**	**10–14**	**15–19**	**20+**
British physicians	15.8	16.0	5.9	5.3	2.0	2.0
US veterans	11.3	18.8	7.5	5.0	5.0	2.1
American Cancer Society[b]	13.7	12.0	7.2	1.1	1.1	1.1

[a]All risks relative to lifelong nonsmokers.
[b]Excludes those who smoked less than one pack of cigarettes per day.

USA, lung cancer rates among African-Americans are substantially higher than among Caucasians in spite of the fact that African-Americans smoke less. Some evidence suggests that ethnic differences in the metabolism of tobacco smoke carcinogens may be involved in these differences. Recent studies have also indicated that there are gender differences in susceptibility to cigarette smoke, with women apparently being at greater risk than men, for a given level of cigarette consumption. It has been proposed that hormonal effects may mediate these differences.

In addition to its firmly established and widely recognized role as a major cause of lung cancer, cigarette smoking is also an important cause of bladder cancer, cancer of the renal pelvis, oral cancer, oropharyngeal cancer, hypopharyngeal cancer, laryngeal cancer, oesophageal cancer and pancreatic cancer (IARC, 1986). Other cancers that may be caused by smoking include renal adenocarcinoma, cancer of the cervix, myeloid leukaemia and stomach cancer (Doll, 1996). Relative risks for major smoking-related cancers are summarized in **Table 6** (Shopland, 1995). In the USA, about 30% of all cancer deaths are caused by smoking, similar to worldwide estimates for developed countries (**Table 7**; Shopland, 1995).

Establishing whether cause and effect exists between environmental tobacco smoke and lung cancer has been more difficult, given the limitations of conventional epidemiological studies (EPA, 1992; Boffetta *et al*., 1998).

Environmental tobacco smoke is mainly a composite of the smoke generated between puffs, called sidestream smoke, and a minor portion of mainstream smoke constituents that are exhaled by a smoker. Although the levels of certain carcinogens in environmental tobacco smoke are greater than in mainstream smoke per gram of tobacco burned, environmental tobacco smoke is diluted by air. Therefore, the carcinogen dose received by a nonsmoker exposed to environmental tobacco smoke may be only 1% of that received by an active smoker and the risk of cancer will be less. Epidemiological studies of environmental tobacco smoke and lung cancer have typically found relative risks between 1 and 2 for exposed versus non-exposed people (EPA, 1992; Boffetta *et al*., 1998). Several panels, including the National Research Council, the US Surgeon General and the Environmental Protection Agency, have concluded that environmental tobacco smoke is a cause of lung cancer.

Table 6 Relative risks for major smoking-related cancer sites among male and female smokers: American Cancer Society 50-State study, 4-year follow-up. (From Shopland, 1995, *Environmental Health Perspectives*, **103**, 131–142.)

Gender	Cancer site	Current smokers	Former smokers
Male	Lung	22.4	9.36
	Oral	27.5	8.80
	Oesophagus	7.60	5.83
	Larynx	10.5	5.24
	Bladder	2.86	1.90
	Pancreas	2.14	1.12
	Kidney	2.95	1.95
Female	Lung	11.9	4.69
	Oral	5.59	2.88
	Oesophagus	10.3	3.16
	Larynx	17.8	11.9
	Bladder	2.58	1.85
	Pancreas	2.33	1.78
	Kidney	1.41	1.16
	Cervix	2.14	1.94

Unburned tobacco is a cause of oral cavity cancer (IARC, 1985). The annual mortality from tobacco chewing in southern Asia, where it is used primarily in the form of betel quid, is estimated to be of the order of 50 000 deaths per year (WHO, 1997). Oral cavity cancer is the leading cancer killer in India. Snuff-dipping, as practised in North America, is also an accepted cause of oral cavity cancer. The prevalence of snuff-dipping has increased markedly in the USA, especially among young males (Hatsukami and Severson, 1999).

TUMOUR INDUCTION IN LABORATORY ANIMALS

Experimental studies evaluating the ability of cigarette smoke and its condensate to cause cancer in laboratory

Table 7 1995 US cancer deaths caused by cigarette smoking. (From Shopland, 1995.)

Gender	Site and ICD disease category	1995 cancer deaths expected	Smoking attributable risk (%)	Estimated deaths due to smoking
Male	Oral, 140-149	5480	90.6	4965
	Oesophagus, 150	8200	76.6	6282
	Pancreas, 157	13200	25.9	3419
	Larynx, 161	3200	79.6	2547
	Lung, 162	95400	89.4	85288
	Bladder, 188	7500	43.8	3285
	Kidney, 189	7100	45.1	3202
	Total cancer deaths expected	289000		108988
Female	Oral, 140-149	2890	58.5	1691
	Oesophagus, 150	2700	71.5	1931
	Pancreas, 157	13800	31.0	4278
	Larynx, 161	890	85.5	761
	Lung, 162	62000	76.1	47182
	Cervix, 180	4800	30.6	1469
	Bladder, 188	3700	34.2	1265
	Kidney, 189	4600	10.7	492
	Total cancer deaths expected	258000		159069
	Total male and female cancer deaths expected in 1995			547000
	Total excess deaths due to cigarette smoking			168 057
	Percentage of cancer deaths due to cigarette smoking in 1995			30.7

animals have collectively demonstrated that there is sufficient evidence that inhalation of tobacco smoke as well as topical application of tobacco smoke condensate cause cancer in experimental animals (IARC, 1986; Hecht, 1998b). The Syrian golden hamster has been the model of choice for inhalation studies of cigarette smoke because it has a low background incidence of spontaneous pulmonary tumours and little interfering respiratory infection. Inhalation of cigarette smoke has repeatedly caused carcinomas in the larynx of hamsters and this model system has been widely applied. It is the most reliable model for induction of tumours by inhalation of cigarette smoke. Studies in mice, rats and dogs have been less frequent.

There are a number of operational problems inherent in inhalation studies of cigarette smoke (IARC, 1986; Hecht, 1998b). The smoke must be delivered in a standardized fashion and this has been accomplished in different ways. Both whole-body exposure and nose-only exposure designs have been used. Generally, a 2-s puff from a burning cigarette is diluted with air and forced into the chamber. Animals will undergo avoidance reactions and will not inhale the smoke the way humans do. Thus, the dose to the lung is less than in humans, and this partially explains the occurrence of larynx tumours rather than lung tumours in hamsters. Unlike humans, rodents are obligatory nose breathers. Their nasal passages are more complex than those of humans, thereby affecting particle deposition in the respiratory tract. Tobacco smoke is irritating and toxic, creating further problems in inhalation studies with rodents.

Inhalation studies have reproducibly demonstrated that cigarette smoke, especially its particulate phase, causes laryngeal carcinomas in hamsters (IARC, 1986; Hecht, 1998b, 1999). Some experiments with mice resulted in low incidences of lung tumours, in tests of both mainstream smoke and environmental tobacco smoke. Respiratory tract tumours were produced in one long-term exposure of rats to cigarette smoke. Studies in rabbits and dogs were equivocal. Treatment-related tumours other than those of the respiratory tract have not been consistently observed.

Cigarette smoke condensate (CSC) has been tested extensively for tumour induction (IARC, 1986; Hecht, 1998b). CSC is produced by passing smoke through cold traps and recovering the material in the traps by washing with a volatile solvent which is then evaporated. Some volatile and semivolatile constituents may be lost during this process. CSC is roughly equivalent to cigarette total particulate matter (TPM), the material collected on a glass-fibre filter which has had smoke drawn through it. The term 'tar,' which is often used in official reports on cigarette brands, is equivalent to TPM but without nicotine and water.

CSC generation and collection techniques have been standardized (IARC, 1986; Hecht, 1998b). The most widely used test system for carcinogenicity of CSC is mouse skin. Consistently, CSC induces benign and malignant skin tumours in mice. This test system has been employed to evaluate the carcinogenic activities of cigarettes of different designs and to investigate mechanisms of carcinogenesis by cigarette smoke. For example, mouse skin studies led to the identification of carcinogenic polycyclic aromatic hydrocarbons (PAHs) in cigarette smoke as well as the demonstration that CSC has co-carcinogenic and tumour-promoting activity. The overall carcinogenic effect of CSC on mouse skin appears to depend on the composite interaction of the tumour initiators such as PAH, tumour promoters and co-carcinogens.

There are drawbacks to the mouse skin assay (Hecht, 1998b). Since CSC lacks volatile and semivolatile components, contributions of these compounds to total activity is lost. Furthermore, mouse skin is insensitive to certain carcinogens in tobacco smoke, such as nitrosamines, which show high selectivity for tissues such as lung. Mouse skin is on the other hand a relatively sensitive tumour induction site for PAHs. Mouse skin studies also ignore the complexity of the respiratory system, where different cell types are known to respond differently to various carcinogens in tobacco smoke. CSC has also been tested by direct injection into the rodent lung, generally in a lipid vehicle. This caused squamous cell carcinomas of the lung in rats. Tumours were not observed in rats treated with the vehicle (IARC, 1986; Hecht, 1998b).

Many studies have evaluated tumour induction in rodents by extracts of unburned tobacco (IARC, 1985). Although some positive results have been obtained, there is currently no widely accepted and reproducible model for the induction of oral cavity cancer in rodents by tobacco extracts, in spite of the strong human data. There are probably cofactors that contribute to human oral cancer upon tobacco use, which are not reproduced in animal studies.

CHEMISTRY OF TOBACCO SMOKE

When cigarette tobacco is burned, mainstream smoke and sidestream smoke are generated (IARC, 1986; Hecht, 1998b). Mainstream smoke is the material drawn from the mouth end of a cigarette during puffing. Sidestream smoke is the material released into the air from the burning tip of the cigarette plus the material which diffuses through the paper. The material emitted from the mouth end of the cigarette between puffs is sometimes also considered as sidestream smoke.

The mainstream smoke emerging from the cigarette is an aerosol containing about 1×10^{10} particles ml^{-1}, ranging in diameter from 0.1 to 1.0 μm (mean diameter 0.2 μm) (IARC, 1986; Hecht, 1998b, 1999). About 95% of the smoke is made up of gases, mainly nitrogen, oxygen and carbon dioxide. For chemical analysis, the smoke is arbitrarily separated into a vapour phase and a particulate phase, based on passage through a glass-fibre filter pad called a Cambridge filter. This retains 99.7% of all

particles with diameters of $\geq 0.1\,\mu m$. Individual smoke components, of which more than 50% appear in the vapour phase of fresh mainstream smoke, are considered volatile smoke components whereas all others are considered particulate phase components. The particulate phase contains more than 3500 compounds, and most of the carcinogens. Standardized machine smoking conditions have been used for measurement of cigarette smoke constituents. These conditions are also arbitrary and it is recognized that each smoker may puff in ways that are widely different from the standardized conditions, thereby changing the yield of individual smoke constituents. In addition to nitrogen, oxygen and carbon dioxide, the gas phase contains substantial amounts of carbon monoxide, water, argon, hydrogen, ammonia, nitrogen oxides, hydrogen cyanide, hydrogen sulfide, methane, isoprene, butadiene, formaldehyde, acrolein, pyridine and other compounds. Some major constituents of the particulate phase include nicotine and related alkaloids, hydrocarbons, phenol, catechol, solanesol, neophytadienes, fatty acids and others. Many of the components are present in higher concentration in sidestream smoke than in mainstream smoke; this is especially true of nitrogen-containing compounds. However, a person's exposure to sidestream smoke is generally far less than to mainstream smoke because of dilution with room air.

Among the many compounds in tobacco smoke are carcinogens, which are agents capable of inducing cancer in laboratory animals or humans. There are 55 carcinogens in cigarette smoke that have been evaluated by the International Agency for Research on Cancer (IARC) and for which there is 'sufficient evidence for carcinogenicity' in either laboratory animals or humans (Hecht, 1999). The types of carcinogens, based on their chemical classes, are listed in **Table 8**. Carcinogens specifically associated with lung cancer are listed in **Table 9**. The 20 compounds included in this list have been found convincingly to induce lung tumours in at least one animal species and

Table 8 Summary of carcinogens in cigarette smoke. (From Hecht, 1999, *Journal of the National Cancer Institute*, **91**, 1194–1210.)

Type	No. of compounds
Polycyclic aromatic hydrocarbons (PAHs)	10
Azaarenes	3
N-Nitrosamines	7
Aromatic amines	3
Heterocyclic aromatic amines	8
Aldehydes	2
Miscellaneous organic compounds	15
Inorganic compounds	7
Total	55

have been positively identified in cigarette smoke. The structures of the organic compounds are shown in **Figure 3**. These compounds are most likely involved in lung cancer induction in people who smoke.

PAHs are condensed ring aromatic compounds that are formed during all incomplete combustion reactions, such as occur in the burning cigarette. Among the PAHs, benzo[*a*]pyrene (BaP) is the most extensively studied compound. Its ability to induce lung tumours upon local administration or inhalation is well documented (Hecht, 1999). It causes lung tumours in mice, but not in rats, when administered systemically. In studies of lung tumour induction by implantation in rats, BaP is more carcinogenic than several other PAHs of tobacco smoke. In analytical studies, it has often been used as a surrogate for other PAHs and extensive data on its occurrence in cigarette smoke are available (IARC, 1986; Hecht, 1999). Thus BaP is a potent lung carcinogen, the occurrence of which is well documented. The vast literature on BaP tends to distract attention from other PAHs. However, PAHs such as dibenz[*a*, *h*]anthracene, 5-methylchrysene and dibenzo[*a*, *i*]pyrene are substantially stronger lung tumorigens than BaP in mice or hamsters, although the levels of these compounds in cigarette smoke are lower than those of BaP (Hecht, 1999).

Azaarenes are nitrogen-containing analogues of PAHs. Two azaarenes, dibenz[*a*,*h*]acridine and 7*H*-dibenzo-[*c*,*g*]carbazole, are pulmonary tumorigens when tested by implantation in the rat lung and instillation in the hamster trachea, respectively (Hecht, 1999). The activity of dibenz[*a*,*h*]acridine is significantly less than that of BaP, whereas that of 7*H*-dibenzo[*c*, *g*]carbazole is greater than BaP. The levels of both compounds in cigarette smoke are relatively low.

N-Nitrosamines are a large group of potent carcinogens formed by nitrosation of amines. Among the *N*-nitrosamines, *N*-nitrosodiethylamine (NDEA) is an effective pulmonary carcinogen in the hamster, but not the rat (Hecht, 1999). Its levels in cigarette smoke are low compared with those of other carcinogens. The tobacco-specific *N*-nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) is a potent lung carcinogen in rats, mice and hamsters (Hecht, 1998a, 1999). NNK is called a tobacco-specific *N*-nitrosamine because it is a chemical derivative of nicotine, and thus occurs only in tobacco products. It is the only compound in **Table 9** which induces lung tumours systemically in all three commonly used rodent models. The organospecificity of NNK for the lung is remarkable; it induces tumours of the lung, mainly adenoma and adenocarcinoma, independent of the route of administration and in both susceptible and resistant strains of mice (Hecht, 1998a, 1999). The systemic administration of NNK to rats is a reproducible and robust method for the induction of lung tumours. Cigarette smoke contains substantial amounts of NNK (IARC, 1986; Hecht, 1998b) and the total dose experienced by a smoker in a lifetime of

Table 9 Pulmonary carcinogens in cigarette smoke. (From Hecht, 1999, *Journal of the National Cancer Institute*, **91**, 1194–1210.)

Carcinogen class	Compound	Amount in mainstream cigarette smoke (ng/cigarette)	Sidestream/ mainstream ratio	Representative lung tumorigenicity in species
PAHs	Benzo[*a*]pyrene (BaP)	20-40	2.5-3.5	Mouse, rat, hamster
	Benzo[*b*]fluoranthane	4-22		Rat
	Benzo[*j*]fluoranthane	6-21		Rat
	Benzo[*k*]fluoranthane	6-12		Rat
	Dibenzo[*a, i*]pyrene	1.7-3.2		Hamster
	Indeno[1,2,3-*cd*]pyrene	4-20		Rat
	Dibenz[*a,h*]anthracene	4		Mouse
	5-Methylchrysene	0.6		Mouse
Azaarenes	Dibenz[*a,h*]acridine	0.1		Rat
	7H-Dibenzo[*c,g*]carbazole	0.7		Hamster
N-Nitrosamines	*N*-Nitrosodiethylamine	ND-2.8	< 40	Hamster
	4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)	80-770	1-4	Mouse, rat, hamster
Miscellaneous organic compounds	1, 3-Butadiene	$(20\text{-}70) \times 10^3$		Mouse
	Ethyl carbamate	20-38		Mouse
Inorganic compounds	Nickel	0-510	13-30	Rat
	Chromium	0.2-500		Rat
	Cadmium	0-6670	7.2	Rat
	Polonium-210	0.03-1.0 pCi	1.0-4.0	Hamster
	Arsenic	0-1400		None
	Hydrazine	24-43		Mouse

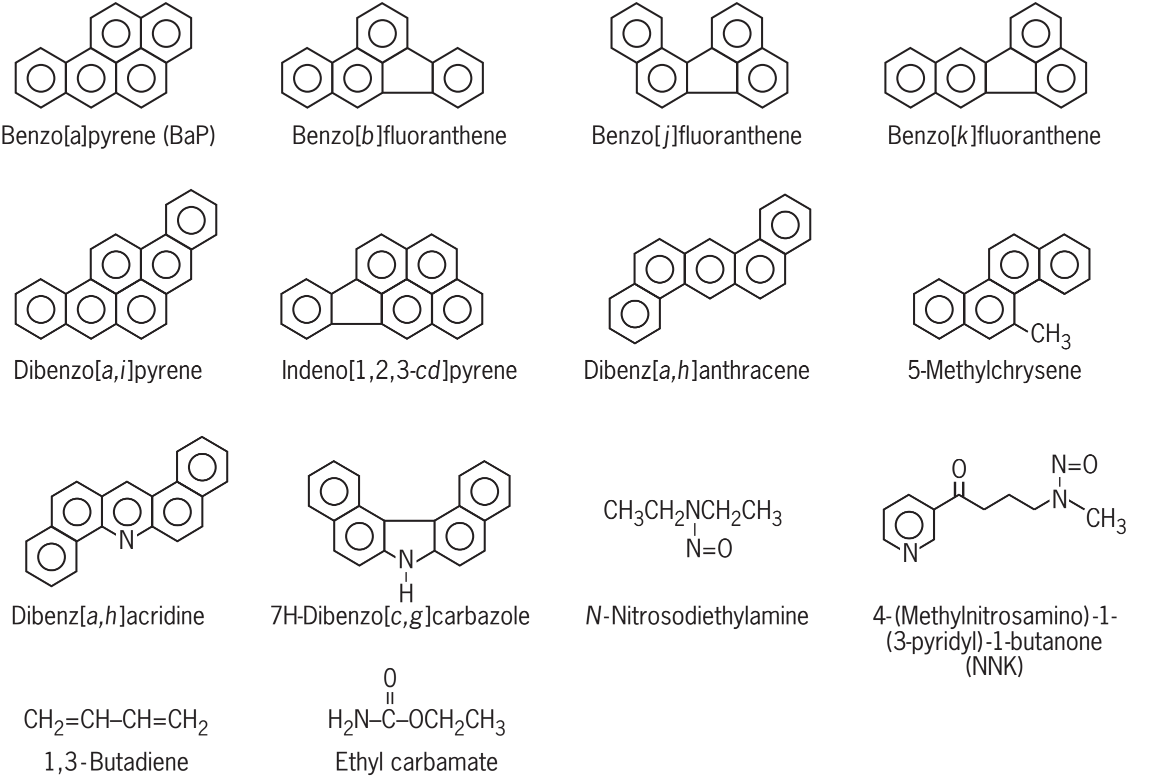


Figure 3 Structures of organic pulmonary carcinogens in tobacco smoke.

smoking is remarkably close to the lowest total dose shown to induce lung tumours in rats (Hecht, 1998a). Levels of NNK and total PAHs in cigarette smoke are similar (IARC, 1986; Hecht, 1998b).

Lung is one of the multiple sites of tumorigenesis by 1,3-butadiene in mice, but is not a target in the rat (Hecht, 1999). 1,3-Butadiene is a component of the vapour phase of cigarette smoke, but in most inhalation studies, the particulate phase shows more overall carcinogenic activity. Ethyl carbamate is a well established pulmonary carcinogen in mice but not in other species (Hecht, 1999). Nickel, chromium, cadmium and arsenic are all present in tobacco and a percentage of each is transferred to mainstream smoke; arsenic levels have been substantially lower since discontinuation of its use as a pesticide in 1952 (Hecht, 1999). Metal carcinogenicity depends on the valence state and anion; these are poorly defined in many analytical studies of tobacco smoke. Thus, although some metals are effective pulmonary carcinogens, the role of metals in tobacco-induced lung cancer is unclear. Levels of polonium-210 in tobacco smoke are not believed to be great enough to impact lung cancer significantly in smokers (Hecht, 1999). Hydrazine is an effective lung carcinogen in mice and has been detected in cigarette smoke in limited studies (Hecht, 1999).

Considerable data indicate that PAHs and NNK play very important roles as causes of lung cancer in people who smoke. The other compounds discussed above may also contribute, but probably to a lesser extent.

PAHs and *N*-nitrosamines such as NNK and *N*′-nitrosonornicotine (NNN) are probably involved as causes of oral cavity cancer in smokers (Hoffmann and Hecht, 1990). *N*-Nitrosamines such as NNN and NDEA are likely causes of oesophageal cancer in smokers (Hoffmann and Hecht, 1990). The risk of oral cavity cancer and oesophageal cancer in smokers is markedly enhanced by consumption of alcoholic beverages. NNK is also believed to play a prominent role in the induction of pancreatic cancer in smokers, whereas aromatic amines such as 4-aminobiphenyl and 2-naphthylamine are the most likely causes of bladder cancer (Hoffmann and Hecht, 1990).

Cigarette smoke is also a tumour promoter (Hecht, 1998b). The majority of the activity seems to be due to uncharacterized weakly acidic compounds. Substantial levels of cocarcinogens such as catechol are present in cigarette smoke (Hecht, 1998b). Co-carcinogens enhance the activity of carcinogens when administered simultaneously. Other co-carcinogens in tobacco smoke include methylcatechols, pyrogallol, decane, undecane, pyrene, benzo[*e*]pyrene and fluoranthene. In addition, cigarette smoke contains high levels of acrolein, which is toxic to the pulmonary cilia, and other agents such as nitrogen oxides, acetaldehyde and formaldehyde that could contribute indirectly to pulmonary carcinogenicity through their toxic effects (Hecht, 1998b, 1999).

Whereas cigarette smoke is extraordinarily complex, unburned tobacco is simpler. With respect to carcinogens, the tobacco-specific nitrosamines NNK and NNN are the most prevalent strong cancer-causing agents in products such as smokeless tobacco (IARC, 1985). A mixture of NNK and NNN induces oral tumours in rats, and consequently these compounds are considered to play a significant role as causes of oral cavity cancer in people who use smokeless tobacco products (IARC, 1985; Hecht, 1998a).

MECHANISMS OF TUMOUR INDUCTION

The mechanisms by which tobacco causes cancer can best be illustrated by considering the relationship between cigarette smoking and lung cancer, because it is here that the most information is available. The overall framework for discussing this information is illustrated in **Figure 4** (Hecht, 1999). Carcinogens form the link between nicotine addiction and cancer. Nicotine addiction is the reason why people continue to smoke. While nicotine itself is not considered to be carcinogenic, each cigarette contains a mixture of carcinogens, including a small dose of PAHs and NNK among other lung carcinogens, tumour promoters and co-carcinogens (Hecht, 1999). Carcinogens such as NNK and PAHs require metabolic activation, that is, they must be enzymatically transformed by the host into reactive intermediates, in order to exert their carcinogenic effects. There are competing detoxification pathways which result in harmless excretion of the carcinogen. The balance between metabolic activation and detoxification differs among individuals and will affect cancer risk.

We know a great deal about mechanisms of carcinogen metabolic activation and detoxification (Hecht, 1999). The metabolic activation process leads to the formation of

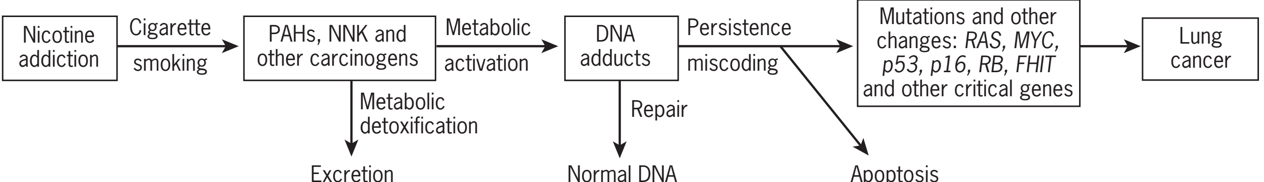


Figure 4 Scheme linking nicotine addiction and lung cancer via tobacco smoke carcinogens and their induction of multiple mutations in critical genes. PAHs = polycyclic aromatic hydrocarbons; NNK = 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone.

DNA adducts, which are carcinogen metabolites bound covalently to DNA, usually at guanine or adenine. There have been major advances in our understanding of DNA adduct structure and its consequences in the past two decades and we now have a large amount of mechanistic information (Hecht, 1999). If DNA adducts escape cellular repair mechanisms and persist, they may lead to miscoding, resulting in a permanent mutation. This occurs when DNA polymerase enzymes read an adducted DNA base incorrectly, resulting in the insertion of the wrong base, or other errors. As a result of clever strategies that combine DNA adduct chemistry with the tools of molecular biology, we know a great deal about the ways in which carcinogen DNA adducts cause mutations. Cells with damaged DNA may be removed by apoptosis, or programmed cell death (Sekido *et al*., 1998). If a permanent mutation occurs in a critical region of an oncogene or tumour-suppressor gene, it can lead to activation of the oncogene or deactivation of the tumour-suppressor gene. Oncogenes and tumour-suppressor genes play critical roles in the normal regulation of cellular growth. Changes in multiple tumour-suppressor genes or oncogenes lead to aberrant cells with loss of normal growth control and ultimately to lung cancer. Although the sequence of events has not been well defined, there can be little doubt that these molecular changes are important (Sekido *et al*., 1998). There is now a large amount of data on mutations in the human K-*ras* oncogene and *p53* tumour-suppressor gene in lung tumours from smokers (Hecht, 1999).

Blocking any of the horizontal steps in **Figure 4** may lead to decreased lung cancer, even in people who continue to smoke. In the following discussion, some of these steps will be considered in more detail.

Upon inhalation, cigarette smoke carcinogens are enyzmatically transformed to a series of metabolites as the exposed organism attempts to convert them to forms that are more readily excreted. The initial steps are usually carried out by cytochrome P450 (CYP) enzymes which oxygenate the substrate (Guengerich and Shimada, 1998). These enzymes typically are responsible for metabolism of drugs, other foreign compounds and some endogenous substrates. Other enzymes such as lipoxygenases, cyclooxygenases, myeloperoxidase and monoamine oxidases may also be involved, but less frequently. The oxygenated intermediates formed in these initial reactions may undergo further transformations by glutathione-*S*-transferases, uridine-5′-diphosphate glucuronosyltransferases, sulfatases and other enzymes which are typically involved in detoxification (Hecht, 1999). Some of the metabolites produced by the CYPs react with DNA or other macromolecules to form covalent binding products known as adducts. This is referred to as metabolic activation (see **Figure 4**). Metabolic pathways of BaP and NNK, representative pulmonary carcinogens in cigarette smoke, have been extensively defined through studies in rodent and human tissues. The major metabolic activation pathway of BaP is conversion to a reactive diol epoxide metabolite called BPDE; one of the four isomers produced is highly carcinogenic and reacts with DNA to form adducts with the N^2-atom of deoxyguanosine (Hecht, 1999). The major metabolic activation pathways of NNK and its main metabolite, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), occur by hydroxylation of the carbons adjacent to the *N*-nitroso group (α-hydroxylation), which leads to the formation of two types of DNA adducts: methyl adducts such as 7-methylguanine or O^6-methylguanine, and pyridyloxobutyl adducts (Hecht, 1998a, 1999).

Considerable information is available on pulmonary carcinogen metabolism *in vitro*, in both animal and human tissues, but fewer studies have been carried out on uptake, metabolism and adduct formation of cigarette smoke lung carcinogens in smokers (Hecht, 1999). Various measures of cigarette smoke uptake in humans have been used, including exhaled carbon monoxide, carboxyhaemoglobin, thiocyanate and urinary mutagenicity. However, the most specific and widely used biochemical marker is the nicotine metabolite cotinine (IARC, 1986; Hecht, 1999). While continine and other nicotine metabolites are excellent indicators of tobacco smoke constituent uptake by smokers, the NNK metabolites NNAL and its *O*-glucuronide (NNAL-Gluc) are excellent biomarkers of tobacco smoke lung carcinogen uptake (Hecht, 1999). NNAL is a potent pulmonary carcinogen like NNK, whereas NNAL-Gluc is a detoxified metabolite of NNK (Hecht, 1998a, 1999). Since NNK is a tobacco-specific carcinogen, its metabolites NNAL and NNAL-Gluc are found only in the urine of individuals exposed to tobacco products. Urinary NNAL and NNAL-Gluc have been quantified in several studies of smokers and in nonsmokers exposed to environmental tobacco smoke. The latter results demonstrate that uptake of NNAL-Gluc is 1–3% of that in smokers, consistent with the weaker epidemiological evidence for a role of environmental tobacco smoke, compared with mainstream cigarette smoke, as a cause of lung cancer (Hecht, 1999).

BaP has been detected in human lung; no differences between smokers and nonsmokers were noted (Hecht, 1999). 1-Hydroxypyrene and its glucuronide, urinary metabolites of the noncarcinogen pyrene, have been widely used as indicators of PAH uptake. 1-Hydroxypyrene levels in smokers are generally higher than in nonsmokers (Hecht, 1999). Overall, there is considerable evidence that pulmonary carcinogens in cigarette smoke are taken up and metabolized by smokers as well as by nonsmokers exposed to environmental tobacco smoke.

Less than 20% of smokers will get lung cancer (IARC, 1986). Susceptibility will depend in part on the balance between carcinogen metabolic activation and detoxification in smokers. This is an important area requiring further study. Most investigations have focused on the metabolic activation pathways by quantifying DNA or protein adducts. There are considerable data demonstrating

activation of BaP to DNA adducts in the lungs of smokers. Earlier investigations demonstrated that cigarette smoke induces aryl hydrocarbon hydroxylase (AHH) activity and proposed a relationship between AHH activity and lung cancer (IARC, 1986; Hecht, 1999). AHH metabolizes BaP to 3-hydroxyBaP, and is equivalent to CYP 1A1. Cigarette smoking induces expression of this enzyme. Cancer patients who stopped smoking within 30 days of surgery had elevated levels of AHH activity compared with nonsmoking cancer patients (Hecht, 1999). Lung tissue from recent smokers with elevated AHH activity also metabolically activated BaP to a greater extent than lung tissue from nonsmokers or ex-smokers. DNA adduct levels correlated with AHH activity in the same samples. Collectively, these results support the existence of a cigarette smoke-inducible pathway leading to BaP–DNA adducts in smokers' lungs, as illustrated in **Figure 4**.

A large number of studies have used immunoassays and ^{32}P postlabelling, which are sensitive but relatively nonspecific techniques, to estimate levels of 'PAH–DNA adducts' or 'hydrophobic DNA adducts' in white blood cells and other human tissues including lung (Hecht, 1999). Many of these have shown elevated adduct levels in smokers. However, none of the studies using immunoassays and ^{32}P postlabelling has identified the structures of the compounds leading to DNA adduct formation. Probably some are PAHs, but individual PAH–DNA adducts have not been characterized in these studies.

Several studies have detected 7-methylguanine in human lung (Hecht, 1999). Levels were higher in smokers than in nonsmokers in two studies, suggesting that NNK may be one source of these adducts. While 7-methylguanine is not generally considered as an adduct that would lead to miscoding in DNA and the introduction of a permanent mutation, other methyl adducts which do have miscoding properties such as O^6-methylguanine are formed at the same time, but at lower levels. Pyridyloxobutylated DNA also has been detected in lung tissue from smokers in one study, reflecting metabolic activation of NNK or NNN. The detection of methyl and pyridyloxobutyl adducts in DNA from smokers' lungs is consistent with the ability of human lung tissue metabolically to activate NNK, but the quantitative aspects of the relationship of metabolism to DNA adduct levels are unclear (Hecht, 1998a, 1999).

DNA repair processes are important in determining whether DNA adducts persist. Because smoking is a chronic habit, one would expect a steady-state DNA adduct level to be achieved by the opposing effects of damage and repair. There are three mechanisms of DNA repair: direct repair, base excision repair and nucleotide excision repair. With respect to smoking and lung cancer, direct repair of O^6-methylguanine by O^6-methylguanine–DNA alkyltransferase and nucleotide excision repair of PAH–DNA adducts would appear to be the most relevant processes (Hecht, 1999).

As indicated in **Figure 4**, the direct interaction of metabolically activated carcinogens with critical genes such as the *p53* tumour-suppressor gene and the K-*ras* oncogene is central to the hypothesis that specific carcinogens form the link between nicotine addiction and lung cancer (Hecht, 1999). The *p53* gene plays a central role in the delicate balance of cellular proliferation and death. It is mutated in about half of all cancer types, including over 50% of lung cancers, leading to loss of its activity for cellular regulation. Point mutations at guanine (G) are common. In a sample of 550 *p53* mutations in lung tumours, 33% were G → T transversions, and 26% were G → A transitions. (A purine → pyrimidine or pyrimidine → purine mutation is referred to as a transversion, and a purine → purine or pyrimidine → pyrimidine mutation is called a transition.) A positive relationship between lifetime cigarette consumption and the frequency of *p53* mutations and of G → T transversions on the nontranscribed DNA strand has also been noted. These observations are generally consistent with the fact that most activated carcinogens react predominantly at G, and that repair of the resulting adducts would be slower on the nontranscribed strand, and thus support the hypothesis outlined in **Figure 4**.

Mutations in codon 12 of the K-*ras* oncogene are found in 24–50% of human primary adenocarcinomas but are rarely seen in other lung tumor types (Hecht, 1999). When K-*ras* is mutated, a complex series of cellular growth signals are initiated. Mutations in K-*ras* are more common in smokers and ex-smokers than in nonsmokers, which suggests that they may be induced by direct reaction with the gene of an activated tobacco smoke carcinogen. The most commonly observed mutation is GGT → TGT, which typically accounts for about 60% of the codon 12 mutations, followed by GGT → GAT (20%), and GGT → GTT (15%).

The $p16^{INK4a}$ tumour-suppressor gene is inactivated in more than 70% of human non-small cell lung cancers, via homozygous deletion or in association with aberrant hypermethylation of the promoter region (Hecht, 1999). In the rat, 94% of adenocarcinomas induced by NNK were hypermethylated at the *p16* gene promoter. This change was frequently detected in hyperplastic lesions and adenomas which are precursors to the adenocarcinomas induced by NNK. Similar results were found in human squamous cell carcinomas of the lung. The *p16* gene was coordinately methylated in 75% of carcinoma *in situ* lesions adjacent to squamous cell carcinomas which had this change. Methylation of *p16* was associated with loss of expression in tumours and precursor lesions indicating functional inactivation of both alleles. Aberrant methylation of *p16* has been suggested as an early marker for lung cancer. The expression of cell cycle proteins is related to the *p16* and retinoblastoma tumour-suppressor genes; NNK-induced mouse lung tumours appear to resemble human non-small-cell lung cancer in the expression of cell

cycle proteins. The oestrogen receptor gene is also inactivated through promoter methylation. There was concordance between the incidence of promoter methylation in this gene in lung tumours from smokers and from NNK-treated rodents.

Loss of heterozygosity and exon deletions within the fragile histidine triad (*FHIT*) gene are associated with smoking habits in lung cancer patients and have been proposed as a target for tobacco smoke carcinogens (Hecht, 1999). However, point mutations within the coding region of the *FHIT* gene were not found in primary lung tumours.

Collectively, the evidence favouring the sequence of steps illustrated in **Figure 4** as an overall mechanism of tobacco-induced cancer is extremely strong, although there are important aspects of each step that require further study. These include carcinogen metabolism and DNA binding in human lung, the effects of cigarette smoke on DNA repair and adduct persistence, the relationship between specific carcinogens and mutations in critical genes and the sequence of gene changes leading to lung cancer.

Using a weight-of-the-evidence approach, specific PAHs and the tobacco-specific nitrosamine NNK can be identified as probable causes of lung cancer in smokers, but the contribution of other agents cannot be excluded (**Table 9**). The chronic exposure of smokers to the DNA-damaging intermediates formed from these carcinogens is consistent with our present understanding of cancer induction as a process which requires multiple genetic changes. Thus, it is completely plausible that the continual barrage of DNA damage produced by tobacco smoke carcinogens causes the multiple genetic changes that are associated with lung cancer. While each dose of carcinogen from a cigarette is extremely small, the cumulative damage produced in years of smoking will be substantial.

Aspects of the scheme illustrated in **Figure 4** are well understood for PAHs and NNK. A great deal is known about the metabolic activation and detoxification of these compounds. There is a good general understanding of the mechanisms by which they interact with DNA to form adducts and considerable information is available about the repair, persistence and miscoding properties of these adducts. There are many aspects of these processes that require further study, however. In particular, little is known about the levels, persistence, and repair of specific carcinogen DNA adducts in the lungs of smokers or the effects of chronic smoking on these factors. The location of carcinogen adducts at specific sites in human DNA has not been studied, mainly owing to limitations on sensitivity. Nevertheless, one can reasonably conclude that metabolically activated tobacco smoke carcinogens directly cause mutations observed in tumour-suppressor genes and oncogenes, although details remain elusive since numerous DNA-damaging agents in tobacco smoke cause similar mutations. (See also chapter *The Formation of DNA Adducts*.)

CHEMOPREVENTION OF TOBACCO-RELATED CANCER

Avoidance of tobacco products is clearly the best way to prevent tobacco-related cancers. Smoking cessation programmes have enjoyed some success in this regard, particularly with the advent of the nicotine patch. Political pressure, especially related to the potentially harmful effects of environmental tobacco smoke, has also had an impact on decreasing smoking. However, approximately 25% of the adult population in the USA continues to smoke, and many of these people are addicted to nicotine. For the addicted smoker who has failed smoking cessation, and for the ex-smoker, chemoprevention may be a way to reduce the risk for cancer (Hecht, 1998b). Chemoprevention involves administration of a non-toxic agent capable of blocking or reversing any of the steps illustrated in **Figure 4**.

Chemoprevention of lung cancer and other tobacco-related cancers in humans is attractive because epidemiological studies have consistently demonstrated a protective effect of vegetables and fruits against these cancers (Hecht, 1998b). This indicates that there are compounds in vegetables and fruits which can inhibit carcinogenesis by NNK, PAHs and other carcinogens. There are already a substantial number of compounds that have been shown to inhibit lung carcinogenesis induced by NNK and BaP in rats and mice, and many of these are naturally occurring (Hecht, 1998b). Phenethyl isothiocyanate (PEITC) is a relatively nontoxic compound which occurs in cruciferous vegetables as a thioglucoside conjugate. It is released upon chewing of the vegetable. PEITC is an effective inhibitor of lung cancer induced by NNK in both rats and mice (Hecht, 1998b). Its major mode of action is inhibition of metabolic activation of NNK by selectively inhibiting cytochrome CYP enzymes of the lung. PEITC does not inhibit carcinogenesis by BaP, but a related naturally occurring isothiocyanate, benzyl isothiocyanate (BITC), is a good inhibitor of BaP-induced lung tumorigenesis in mice (Hecht, 1998b). Other isothiocyanates are known inhibitors of tumour development at other sites. Notable among these is sulforaphane, a constituent of broccoli.

Some of the other compounds which have been shown to inhibit lung carcinogenesis by NNK include butylated hydroxyanisole, an antioxidant used in food preservation, (+)-limonene, a constituent of orange juice and other citrus products, and diallyl sulfide, a constituent of garlic. Inhibition of NNK carcinogenesis has also been observed in animals treated with green and black tea, as well as its major polyphenolic constituents. Inhibitors of lung tumorigenesis induced by BaP include β-naphthoflavone, butylated hydroxyanisole, ethoxyquin, diallyl sulfide and *myo*-inositol. It seems likely that properly designed combinations of some of these inhibitors will be effective chemopreventive agents against lung cancer in humans.

A number of human trials are already in progress, and several of these have centred on β-carotene as a potential chemopreventive agent for lung cancer. The results have not been encouraging, as β-carotene had an enhancing effect on lung cancer, and vitamin E and a combination of β-carotene and vitamin E had no inhibitory effect (Hecht, 1998b). However, it should be noted that animal studies have not demonstrated efficacy against lung cancer for these two agents. In future trials, it will be important to select chemopreventive agents carefully based on their effectiveness in animal studies and on known mechanisms of action that would be applicable to smokers and ex-smokers. In addition, subjects may need to be selected based on intermediate biomarker profiles and smoking history.

CONCLUSION

The sheer magnitude of the tobacco and cancer problem is difficult to grasp because the numbers are so large. While significant progress has been made in tobacco control, especially in the USA and some other developed countries, the worldwide problem is still immense. Tobacco products cause about 30% of all cancer deaths in developed countries. The epidemic of lung cancer, while possibly starting to abate in the USA and UK, is only beginning to develop fully in other parts of the world such as China. Oral cavity cancer due to tobacco products is a major cause of cancer deaths in parts of Asia. Tobacco products play a significant role as causes of cancers of the upper respiratory and digestive tract, pancreas, bladder and kidney. Three general strategies can be envisioned to decrease tobacco-related cancer death: (1) prevent people from starting to use tobacco; (2) if they do start, find ways to treat their dependence resulting in cessation; and (3) develop ways to decrease the risk of cancer in people who continue to use tobacco products. The first two approaches are likely to be the most effective, but they may not be universally successful. Taking the USA as an example, there has been no change in smoking prevalence since 1990. The third approach depends on an understanding of mechanisms of tobacco-induced cancer, as discussed here. Blocking any of the horizontal steps in **Figure 4** should result in decreased cancer incidence. This can potentially be achieved by identifying particularly susceptible individuals and by developing effective chemoprevention strategies. Even if these approaches are only partly successful, they would significantly impact overall cancer death rates.

REFERENCES

Blot, W. J. and Fraumeni, J. F. Jr (1996). Cancers of the lung and pleura. In: Schottenfeld, D. and Fraumeni, J. F., Jr (eds), *Cancer Epidemiology and Prevention*. 637–665 (Oxford University Press, New York).

Boffetta, P., *et al*. (1998). Multicenter case-control study of exposure to environmental tobacco smoke and lung cancer in Europe. *Journal of the National Cancer Institute*, **90**, 1440–1450.

Doll, R. (1996). Cancers weakly related to smoking. *British Medical Journal*, **52**, 35–49.

EPA (1992). *Respiratory Health Effects of Passive Smoking: Lung Cancer and Other Disorders*. Report No. EPA/600/6-90/006F. (Environmental Protection Agency. Office of Health and Environmental Assessment, Office of Research and Development, Washington, DC).

Guengerich, F. P. and Shimada, T. (1998). Activation of procarcinogens by human cytochrome P450 enzymes. *Mutation Research*, **400**, 201–213.

Hatsukami, D. K. and Severson, H. H. (1999). Oral spit tobacco: addictions prevention and treatment. *Nicotine and Tobacco Research,* **1**, 21–44.

Hecht, S. S. (1998a). Biochemistry, biology, and carcinogenicity of tobacco-specific *N*-nitrosamines. *Chemical Research in Toxicology*, **11**, 559–603.

Hecht, S. S. (1998b). Cigarette smoking and cancer. In: Rom, W. N. (ed.), *Environmental and Occupational Medicine*. 1479–1499. (Lippincott-Raven, New York).

Hecht, S. S. (1999). Tobacco smoke carcinogens and lung cancer. *Journal of the National Cancer Institute*, **91**, 1194–1210.

Hoffmann, D. and Hecht, S. S. (1990). Advances in tobacco carcinogenesis. In: Cooper, C. S. and Grover, P. L. (eds), *Handbook of Experimental Pharmacology*, 94/I. 63–102 (Springer, Heidelberg).

IARC (1985). Tobacco habits other than smoking: betel quid and areca nut chewing and some related nitrosamines. In: *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Vol. 37 (International Agency for Research on Cancer, Lyon).

IARC (1986) Tobacco smoking. In: *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Vol. 38 (International Agency for Research on Cancer, Lyon).

Sekido, Y. *et al*. (1998). Progress in understanding the molecular pathogenesis of human lung cancer. *Biochimica Biophysica Acta*, **1378**, F21–F59.

Shopland, D. R. (1995). Tobacco use and its contribution to early cancer mortality with a special emphasis on cigarette smoking. *Environmental Health Perspectives*, **103** (**supplement 8**), 131–142.

WHO (1997). *Tobacco or Health: A Global Status Report*. 1–48 (World Health Organization, Geneva).

Wynder, E. L. and Hoffmann, D. (1994) Smoking and lung cancer: scientific challenges and opportunities. *Cancer Research*, **54**, 5284–5295.

FURTHER READING

Blot, W. J. and Fraumeni, J. F., Jr (1996) Cancers of the lung and pleura. In: Schottenfeld, D., and Fraumeni, J. F., Jr (eds),

Cancer Epidemiology and Prevention. 637–665 (Oxford University Press, New York).

Hecht, S. S. (1998a). Biochemistry, biology, and carcinogenicity of tobacco-specific *N*-nitrosamines. *Chemical Research in Toxicology*, **11**, 559–603.

Hecht, S. S. (1998b). Cigarette smoking and cancer. In: Rom, W. N. (ed.), *Environmental and Occupational Medicine*. 1479–1499 (Lippincott-Raven, New York).

Hecht, S. S. (1999). Tobacco smoke carcinogens and lung cancer. *Journal of the National Cancer Institute*, **91**, 1194–1210.

IARC (1986). Tobacco smoking. In: *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Vol. 38 (International Agency for Research on Cancer, Lyon).

Sekido, Y., *et al*. (1998). Progress in understanding the molecular pathogenesis of human lung cancer. *Biochimica Biophysica Acta*, **1378**, F21–F59.

WHO (1997). *Tobacco or Health: A Global Status Report*. 1–48 (World Health Organization, Geneva).

Chapter 30

Occupational Causes of Cancer

Harri Vainio
International Agency for Research on Cancer, Lyon, France

Tony Fletcher
London School of Hygiene and Tropical Medicine, London, UK

Paolo Boffetta
International Agency for Research on Cancer, Lyon, France

CONTENTS

- Historical and Current Patterns on Occupational Cancer Risks
- Known and Suspected Occupational Carcinogens
- Estimates of the Burden of Cancer Attributable to Occupation
- Prevention of Occupational Cancer
- Prospects for Cancer Prevention

HISTORICAL AND CURRENT PATTERNS ON OCCUPATIONAL CANCER RISKS

Exposures encountered at the workplace are a substantial source of cancer, as has been known for over 200 years (Pott, 1775). Occupational cancers were initially detected by clinicians. From early findings of Pott of scrotal cancer among chimney sweeps in 1775 to Creech and Johnson's identification of angiosarcoma of the liver among vinyl chloride workers in 1974, unusual cancer patterns among persons with unusual occupations amounted to sufficient evidence to judge that the occupational exposure had caused the cancer (Creech and Johnson, 1974). Pott was a physician treating chimney sweeps and Creech was a physician who treated vinyl chloride monomer workers. The era of initial identification of occupational cancer by clinicians extended into the last quarter of the twentieth century. The period of formal epidemiological assessment of the occurrence of cancer in relation to workplace exposures started after the Second World War, and knowledge of the occupational and other environmental causes of cancer then grew rapidly in the next few decades. It has been clear for about 20 years that it should be possible to reduce the incidence of cancer in middle and old age by 80–90% (Doll and Peto, 1981); however, the precise means of how to bring about such a large reduction were not known. It might be helpful to recall here that most known occupational carcinogens were identified in the 1950–1970s, with few carcinogens identified later. Cancer hazards in the workplaces in the earlier decades of the twentieth century were substantial, causing, in the extreme case, all the most heavily exposed workers to develop cancer, as occurred in some groups of manufacturers of 2-naphthylamine and benzidine, while coal tar fumes and asbestos were so widespread that many thousands of skin and lung cancers were produced. Although the remainder of these cases are now starting to disappear through elimination of these substances and/or exposure to them, some of the consequences of the earlier exposure are still evident. Most estimates of the burden occupational cancer in industrialized countries are around 5% overall, with higher proportions among subgroups (Doll and Peto, 1981; Tomatis, 1990; Boffetta and Kogevinas, 1999). Mesotheliomas due to exposure to asbestos are typical examples of persisting occupational risks.

The workplace is an environment that provides unusual opportunities for the causation and thus the prevention and control of cancer. Occupational exposures tend to be greater, sometimes by orders of magnitude, than exposures in the general environment. This chapter provides a brief review of the occurrence and causes of occupational cancer based on epidemiological studies in humans and characteristics of occupational exposures. (See also chapter on *Identifying Cancer Causes through Epidemiology*.)

KNOWN AND SUSPECTED OCCUPATIONAL CARCINOGENS

The International Agency for Research on Cancer (IARC) has established, within the framework of its Monographs programme, a set of criteria to evaluate the evidence for carcinogenicity of specific agents. The IARC Monographs programme represents one of the most comprehensive efforts to review cancer data systematically and

consistently, and is highly reputed in the scientific community (IARC, 1972–2000). It also has an important impact on national and international occupational cancer control activities. Agents, mixtures and exposure circumstances are evaluated within the IARC Monographs if there is evidence of human exposure, and data on carcinogenicity (either in humans or in experimental animals) are available.

The evaluation of carcinogenicity within the IARC Monographs programme includes several steps. First, the evidence for the induction of cancer in humans, which obviously plays an important role in the identification of human carcinogens, is reviewed and is classified into one of the following categories:

Sufficient evidence of carcinogenicity: a causal relationship has been established between exposure to the agent, mixture or exposure circumstance and human cancer. That is, a positive relationship has been observed between the exposure and cancer in studies in which chance, bias and confounding could be ruled out with reasonable confidence.

Limited evidence of carcinogenicity: a positive association has been observed between exposure to the agent, mixture or exposure circumstance and cancer for which a causal interpretation is considered to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.

Inadequate evidence of carcinogenicity: the available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association, or no data on cancer in humans are available.

Evidence suggesting lack of carcinogenicity: there are several adequate studies covering the full range of levels of exposure that human beings are known to encounter, which are mutually consistent in not showing a positive association between exposure to the agent and the studied cancer at any observed level of exposure.

Second, studies in which experimental animals (mainly rodents) are exposed chronically to potential carcinogens and examined for evidence of cancer are reviewed and the degree of evidence of carcinogenicity is then classified into categories similar to those used for human data.

Third, data on biological effects in humans and experimental animals that are of particular relevance are reviewed. These may include toxicological, kinetic and metabolic considerations and evidence of DNA binding, persistence of DNA lesions or genetic damage in exposed humans. Toxicological information, such as that on cytotoxicity and regeneration, receptor binding and hormonal and immunological effects, and data on structure–activity relationship are used when considered relevant to the possible mechanism of the carcinogenic action of the agent.

Finally, the body of evidence is considered as a whole, in order to reach an overall evaluation of the carcinogenicity to humans of an agent, mixture or circumstance of exposure. The agent, mixture or exposure circumstance is described according to the wording of one of the following categories:

Group 1: The agent (mixture) is carcinogenic to humans. The exposure circumstance entails exposures that are carcinogenic to humans. This evaluation is mainly reached when the evidence in humans is considered sufficient.

Group 2A: The agent (mixture) is probably carcinogenic to humans. The exposure circumstance entails exposures that are probably carcinogenic to humans. Agents classified in this category are typically experimental carcinogens for which there is either limited epidemiological evidence in humans or mechanistic data suggesting that they also operate in humans.

Group 2B: The agent (mixture) is possibly carcinogenic to humans. The exposure circumstance entails exposures that are possibly carcinogenic to humans. Agents in this group are mainly experimental carcinogens with no or inadequate data in humans.

Group 3: The agent (mixture, exposure circumstance) is not classifiable as to its carcinogenicity to humans. Agents is this category are typically suspected experimental carcinogens with no or inadequate human data.

Group 4: The agent (mixture, exposure circumstance) is probably not carcinogenic to humans. This category includes agents with evidence suggesting lack of carcinogenicity.

At present, 26 chemicals, groups of chemicals and mixtures for which exposures are mostly occupational are established human carcinogens (**Table 1**). Whereas some agents such as asbestos, benzene and heavy metals are currently widely used in many countries, other agents are of mainly historical interest (e.g. mustard gas and 2-naphthylamine).

Twenty-six additional agents are classified as probably carcinogenic to humans (Group 2A): they are listed in **Table 2**, and include exposures that are currently prevalent in many countries, such as formaldehyde and buta-1,3-diene. A large number of important occupational agents are classified as possible human carcinogens (Group 2B), e.g. acetaldehyde, carbon black, chloroform, chlorophenoxy herbicides, DDT, dichloromethane, mineral wools, inorganic lead compounds, polychlorophenols and styrene. The complete list can be found at the IARC website (www.iarc.fr).

The distinction between occupational and environmental carcinogens is not always straightforward. Several of the agents listed in **Tables 1** and **2** are also present in the general environment, although exposure levels tend to be higher at the workplace. This is the case, for example, with 2,3,7,8-TCDD, diesel engine exhaust, radon and asbestos. On the other hand, there are agents that have been evaluated in IARC Group 1 or 2A for which exposure is not primarily occupational, but which are often

Table 1 Agents, groups or agents and mixtures classified as established human carcinogens (IARC group 1) for which exposure is mainly occupational

Exposure	Target organ/cancer	Main industry or use
4-Aminobiphenyl	Bladder	Rubber
Arsenic and arsenic compounds	Lung, skin	Glass, metals, pesticides
Asbestos	Lung, pleura	Insulation, construction
Benzene	Leukaemia	Solvent, fuel
Benzidine	Bladder	Pigment
Beryllium and beryllium compounds	Lung	Aerospace, metals
Bis(chloromethyl) ether[a]	Lung	Chemical
Chloromethyl methyl ether[a]	Lung	Chemical
Cadmium and cadmium compounds	Lung	Pigment, battery
Chromium(VI) compounds	Nasal cavity, lung	Metal plating, pigment
Coal-tar pitches	Skin, lung, bladder	Construction, electrodes
Coal-tars	Skin, lung	Fuel
Ethylene oxide	Leukaemia	Chemical, sterilant
Mineral oils, untreated and mildly treated	Skin	Lubricant
Mustard gas (sulfur mustard)[a]	Pharynx, lung	War gas
2-Naphthylamine[a]	Bladder	Pigment
Nickel compounds	Nasal cavity, lung	Metal, alloy
Radon and its decay products	Lung	Mining
Shale-oils	Skin	Lubricant, fuel
Silica, crystalline (inhaled in the form of quartz or cristobalite)	Lung	Construction, mining
Soots	Skin, lung	Pigment
Strong inorganic acid mists containing sulfuric acid	Larynx, lung	Chemical
Talc containing asbestiform fibres	Lung	Paper, paint
2,3,7,8-Tetrachlorodibenzo-*p*-dioxin	Several	Chemical
Vinyl chloride	Liver	Plastic
Wood dust	Nasal cavity	Wood

[a]Agent mainly of historical interest.

encountered in the occupational environment. They include the following:

- drugs, such as cyclophosphamide, combined chemotherapy including alkylating agents such as MOPP, and cyclosporin: occupational exposure can occur in pharmacies and during their administration by nursing staff;
- food contaminants, such as aflatoxins, to which food processors can be exposed;
- biological agents, such as hepatitis B virus, hepatitis C virus and human immunodeficiency virus, to which medical personnel can be exposed;
- environmental agents, in particular solar radiation (exposure in agriculture, fishing and other outdoor occupations);
- environmental tobacco smoke, deriving from smoking by fellow employees or by the public in bars, restaurants and other public settings.

Polycyclic aromatic hydrocarbons (PAHs) represent a specific problem in the identification of occupational carcinogens. This group of chemicals include several potent experimental carcinogens, such as benzo[*a*]pyrene, benz[*a*]anthracene and dibenz[*a,h*]anthracene. However, humans are always exposed to mixtures of PAHs (several of which are listed in **Tables 1** and **2**, e.g. coal-tars, soots, creosotes) and an assessment of the carcinogenicity of individual PAHs in humans cannot be done at present.

Current understanding of the relationship between occupational exposures and cancer is far from complete; in fact, only 26 individual agents are established occupational carcinogens (**Table 1**) and for many more experimental carcinogens no definitive evidence is available from exposed workers. In some cases, there is considerable evidence of increased risks associated with particular industries and occupations, although no specific agents can be identified as aetiological factors. **Table 3** reports occupations and industries which entail (or are suspected to entail) a carcinogenic risk on the basis of the IARC Monograph programme.

Constructing and interpreting lists of chemical or physical carcinogenic agents and associating them with specific occupations and industries is complicated by a number of factors: (1) information on industrial processes and exposures is frequently poor, not allowing a complete evaluation of the importance of specific carcinogenic exposures in different occupations or industries; (2) exposures to well-known carcinogenic agents, such as vinyl chloride and benzene, occur at different intensities in different occupational situations; (3) changes in exposure

Table 2 Agents, groups or agents and mixtures classified as probable human carcinogens (IARC group 2A) for which exposure is primarily occupational

Exposure	Suspected target organ/cancer	Main industry or use
Acrylamide	–	Plastic
Benzidine-based dyes	Bladder	Pigment, leather
Buta-1,3-diene	Leukaemia	Plastic, rubber
Captafol	–	Pesticide
α-Chlorinated toluenes (benzal chloride, benzotrichloride, benzyl chloride, benzoyl chloride)	–	Pigment, chemical
p-Chloro-*o*-toluidine	Bladder	Pigment, textile
Creosotes	Skin	Wood
Diesel engine exhaust	Lung	Transport, mining
Diethyl sulfate	–	Chemical
Dimethylcarbamoyl chloride	–	Chemical
1,2-Dimethylhydrazine	–	Research
Dimethyl sulfate	–	Chemical
Epichlorohydrin	–	Plastic
Ethylene dibromide	–	Fumigant
Formaldehyde	Nasopharynx	Plastic, textile
Methyl methane-sulfonate	–	Chemical
4,4′-Methylenebis2-chloroaniline (MOCA)	Bladder	Rubber
Non-arsenical insecticides	Leukaemia	Agriculture
Polychlorinated biphenyls	Liver, lymphoma	Electrical components
Styrene-7,8-oxide	–	Plastic
Tetrachloroethylene	Oesophagus, lymphoma	Solvent
Trichloroethylene	Liver, kidney lymphoma	Solvent, dry cleaning
Trichloropropane	–	Solvent
Tris(2,3-dibromopropyl) phosphate	–	Plastic, textile
Vinyl bromide	–	Plastic, textile
Vinyl fluoride	–	Chemical

occur over time in a given occupational situation, either because identified carcinogenic agents are replaced by other agents or (more frequently) because new industrial processes or materials are introduced; and (4) any list of occupational exposures can only refer to the relatively small number of chemical exposures which have been investigated with respect to the presence of a carcinogenic risk. (See also chapter on *Mechanisms of Chemical Carcinogenesis*.)

Table 3 Industrial processes and occupations evaluated in IARC Monographs Volumes 1–78 (IARC 1972–2001)

Industry/occupation	Target organs/cancer[a]
Group 1	
Aluminium production	Lung, bladder
Auramine manufacture	Bladder
Boot and shoe manufacture and repair	Nasal cavity, leukaemia
Coal gasification	Skin, lung, bladder
Coke production	Skin, lung, kidney
Furniture and cabinet making	Nasal cavity
Haematite mining (underground) with exposure to radon	Lung
Iron and steel founding	Lung
Magenta manufacture	Bladder
Painter	Lung
Propan-2-ol manufacture (strong-acid process)	Nasal cavity
Rubber industry	Bladder, leukaemia
Group 2A	
Art glass, glass containers and pressed ware, manufacture of	(Lung, stomach)
Hairdressers or barbers	(Bladder, lung)
Petroleum refining	(Leukaemia, skin)
Group 2B	
Carpentry and joinery	(Nasal cavity)
Dry cleaning	(Oesophagus, bladder)
Printing processes	(Lung, bladder)
Textile manufacturing industry	(Nasal cavity, bladder)

[a]Suspected target organs are given in parentheses.

ESTIMATES OF THE BURDEN OF CANCER ATTRIBUTABLE TO OCCUPATION

It is instructive to estimate the number of cancers that might be prevented by avoiding workplace exposure to carcinogens. In practice this is most readily estimated in terms of the number of cancers which can be attributed to past exposures to workplace carcinogens. Estimating the total proportion of cancers attributable to occupation involves some extrapolation, and there are essentially two approaches which may be taken. One method draws on studies of specific occupational groups (usually ‘cohort studies’) in which the numbers of attributable cases can be estimated. Along with some estimate of the total number of exposed workers, the total burden can thus be estimated. This approach is somewhat uncertain as there is usually very limited quantitative information on the extent and level of exposure across occupational groups collected in

a comparable way to the specific epidemiological studies from which risk estimates derive.

A more satisfactory approach is to estimate the attributable fraction directly from case control studies in communities. This fraction is, of course, specific to the community where the study was conducted, but if there are a number of such studies which can be synthesized, then a global estimate may be made. There are now a wealth of occupational case-control studies and the proportions summarized here derive from these types of studies.

The sites of cancer which contribute most numbers to the estimated burden of occupational cancer include mesothelioma, lung, bladder, sinonasal and laryngeal cancers. Estimated burdens in terms of the proportions of cases attributable to occupation are summarized here.

Mesothelioma

Pleural mesothelioma death rates are rising in several European countries where surveillance is most effective, in particular Finland, UK and The Netherlands. Extrapolation of current trends suggests that death rates will continue to rise, reaching a peak around 2018, before falling again. Asbestos is the overwhelming cause of mesothelioma and the proportion attributable to asbestos may be considered to be over 80%. Some of these are not directly occupational (e.g. children exposed while living adjacent to asbestos factories or family members exposed to dust brought home by asbestos workers), but most are due to occupational exposures. As a proportion of cancers, the approximately 1200 per year in the UK amount to about 0.5% of all cancers per annum and this is expected to rise to 2% of cancers at the mortality peak in 20–30 years time. In terms of absolute numbers it has been projected that asbestos-related mesothelioma deaths will amount to 250 000 in total in Western Europe alone over the next 35 years (Albin *et al*., 1999; Peto *et al*., 1999).

Lung Cancer

Case control studies in different communities show a range of estimates of proportions of lung cancers attributable to known occupational carcinogens. Two recent studies have looked at very large populations to overcome the small numbers inherent in individual studies. One is a four-country analysis of the entire Nordic population followed prospectively for cancer incidence from the 1970 censuses, and using exposures inferred from the occupation reported at the census. From this, 18% of male lung cancers and <1% of female lung cancers are attributed to occupational exposures. This compares with an estimate derived from a reanalysis of eight case-control studies in five European countries of 13% for male cancers and 3% for females. Both values are close to the earlier estimate for the USA of 15% for males, although lower than that estimate of 5% for females. Overall, an estimate of 15% for males and 2–3% for females would seem reasonable (Doll and Peto, 1981; Vineis and Simonato, 1991; Boffetta and Kogevinas, 1999).

Bladder

For cancer of the bladder, a pooled reanalysis of 11 case control studies of occupational risks for males found an overall attributable fraction for known occupational risks of 4%. The Nordic study found 2%. These are both somewhat lower than the 10% estimated by Doll and Peto (1981) for the USA. For females, both studies estimated the attributable fractions overall as very low.

Sinonasal Cancers

While most adenocarcinomas of the nose are caused by occupational exposures, the attributable fraction for all sinonasal cancers together is estimated as 30–41% for males by the Nordic and European pooled reanalysis of eight case control studies, respectively. For females the results are less consistent, with the equivalent estimates being 2% and 7% of all sinonasal cancers.

Laryngeal Cancers

The attributable fraction for laryngeal cancers together is estimated as 6–8% for males by the Nordic and European pooled reanalysis of six case control studies. For females, the attributable fraction was close to zero in both cases.

In summary, estimates for Europe of the proportion of cancers attributable to occupation varies between cancer site, from close to 100% for mesothelioma, close to 50% for sinonasal cancers among males, around 15% for male lung cancer and a little under 10% for larynx and bladder cancer, again among males. For females the risks are generally lower (apart from mesothelioma, which as for males is attributable entirely to asbestos), with attributable fractions under 1% except for lung cancer, which is around 1–2%, and sinonasal cancer, between 2 and 7%. For most other cancer sites, the attributable fractions calculated in this way are lower. Summing across all sites to estimate the total proportion of cancers which may be attributable to occupational exposures has led to estimates around 5%.

Cigarette smoking has been dealt with to an extent in these studies, in that in virtually all of them active smoking was assessed in cases and controls and controlled for in the analysis. However, in most cases these studies were conducted before awareness had grown about environmental tobacco smoke (ETS), and they focused on other occupational carcinogens. Thus, if the contribution of ETS exposure at the workplace were included, the attributable fraction of workplace exposures would need to be revised upwards, especially for lung cancer. As tobacco cessation programmes at workplaces are a relatively recent

phenomenon, the health impact of ETS exposure at work will take some years before its impact on occupational cancers will have waned.

Recent studies of occupational cancer estimate the burden on current disease of past exposures. The burden on future disease due to current exposure would be expected to fall in countries where controls on occupational carcinogens have been successful. However, estimates of the total numbers of workers exposed to carcinogens at work remain high. Recent work carried out in the European Union to estimate these numbers has led to a European database on occupational exposures in the workplace, CAREX (available at http://www.occuphealth.fi/list/data/CAREX/). This is based on a systematic estimate in each country of the numbers of employees exposed to a defined list of established carcinogens at work. The grand total is 32 million, representing 23% of the working population, estimated for the early 1990s. The most important contributions come from solar radiation (from working outdoors) and environmental tobacco smoke (8.8 and 7.1 million exposed, respectively), followed by more classical occupational pollutants such as silica, diesel exhaust, radon decay products and wood dust (about 3 million each). These are very substantial numbers, emphasizing the remaining scope for primary prevention.

Outside Europe, the burden of occupational cancer may be different. Recent epidemiological studies of occupational cancer risk in developing countries revealed situations where exposures and risks might be much higher than currently found in Northern countries (Kogevinas *et al*., 1994). Some carcinogens are being substantially phased out, in some but not all countries. For example, asbestos use in Europe has fallen dramatically over the last 20 years and is now negligible in some countries, e.g. 0.004 kg per person per year in the Nordic countries in 1996, although in other countries consumption remains high, e.g. 2.4 kg per person per year in the former Soviet Union in 1996 (Albin *et al*., 1999). Therefore, the proportions of cancers attributable to occupation estimated above for Europe would underestimate the proportions in more poorly regulated developing country contexts.

PREVENTION OF OCCUPATIONAL CANCER

Primary prevention of occupational cancer is defined as aetiological prevention – prevention is directed against the source of the disease. Primary prevention contrasts with later (secondary and tertiary) forms of prevention that involve, respectively, the early detection of disease and treatment of disease to prevent death and reduce disability. Primary prevention is inherently more effective than secondary and tertiary prevention. Several avenues exist for the prevention of occupational cancer.

Legal Initiatives

Restriction of the use of carcinogens at the workplace, after they have been identified, is the simplest but bluntest tool available for risk reduction. Banning the use of a substance has been successfully used with occupational carcinogens.

Regulation is different from restriction. Regulation requires that anyone who deals with the substance should keep to certain minimum standards to minimize the exposure and consequently the toxic effects. Known carcinogens are strictly regulated at workplaces, through standards, at least in industrialized countries.

Industrial Hygiene

Industrial hygiene, workplace technology and general knowledge on the safety issues have been continuously improving in industrialized countries over the last 30 years. As a consequence, exposures in the workplace are now less intense than in the past. However, the possibility of low-level exposures exists, sometimes to a multitude of chemicals or mixtures of chemicals with possibilities for various types of interactions.

The most efficient of the industrial hygiene tools is the substitution of a new, less hazardous material for a material of known carcinogenicity. It is important, however, to ensure the new material is indeed less hazardous than that which is replaces. Other approaches include process enclosure or isolation, or use of ventilation.

Use of Personal Protective Equipment

Respirators, gloves and other forms of protective clothing are all common forms of protective equipment in use throughout industry. They can be important in reducing carcinogenic exposures provided that carefully designed equipment is in use and that equipment is properly used and maintained.

Cancer Surveillance and Monitoring at the Workplace

Biological monitoring and medical screening of workers uses information from health history and results from periodical physical and laboratory examinations to estimate the levels of exposure and to assess the early health effects. Surveillance of workers is useful for identifying unforeseen hazards and to protect workers who are at increased risk – with the idea of detecting the cancer in its presymptomatic stages when it still can be controlled or cured. Screening for occupational cancer in exposed populations for purposes of early diagnosis is rarely applied, but has been tested in some situations. Medical surveillance of populations at risk of getting cancer is only effective in the following situations: (1) if the screening

test is easy to perform and sensitive, (2) if it detects premalignant abnormalities or tumours at an early stage and (3) if there is an effective intervention that reduces morbidity and mortality when applied to early tumours.

Cancer surveillance at the workplace has been explored, e.g., for bladder cancer among people exposed to 2-napthylamine and to benzidine, and for lung cancer among workers exposed to asbestos. Chest X-rays, cytology and urinalysis have been proposed (Schulte *et al*., 1990). Urine cytology has good sensitivity and specificity for invasive bladder cancer, but no survival advantage has been demonstrated. Also, chest X-rays and cytological examination of sputum has not reduced mortality from lung cancer significantly. Judgments on the value of screening depend also on the intensity of exposure: it may be more justified in small groups exposed to high levels of carcinogens than among large groups exposed to low levels.

PROSPECTS FOR CANCER PREVENTION

In the future, new technology and molecular biomarkers (of exposure, effect and susceptibility) promise to revolutionize the practice of cancer prevention and provide new tools for screening and prevention of occupational cancer. Lung cancer remains the leading cause of death from cancer, and recent developments have generated substantial interest in the use of spiral computed tomography (CT) to screen for lung cancer. The use of lung cancer biomarkers to identify the early clonal phase of progression of lung cancer in high-risk populations has also been proposed, enabling cancers to be detected earlier than is possible with spiral CT, since the latter is not as sensitive for small central cancers as it is for small peripheral cancers. A hybrid CT–biomarker approach may improve the robustness with which lung cancer of any type can be detected early (Mulshine and Henschke, 2000). The high-risk populations screened for lung cancer will generally include large numbers of people who share risks of other, possibly coexisting, smoking- and/or asbestos-related disorders. Spiral CT is being used routinely in the investigation of emphysema and asbestosis. Simultaneous screening for coexisting disorders would be of great public health value. A rational integrated use of spiral CT and biomarkers in lung cancer screening is a research priority. However, for the purposes of practical prevention at workplaces, the avoidance of exposure remains the main and the primary means for occupational cancer prevention.

REFERENCES

Albin, M., *et al*. (1999). Asbestos and cancer: an overview of current trends in Europe. *Environmental Health Perspectives*, **107**, Suppl. 2, 289–298.

Boffetta, P. and Kogevinas, M. (1999). Introduction: epidemiologic research and prevention of occupational cancer in Europe. *Environmental Health Perspectives*, **107**, Suppl. 2, 229–231.

Creech, J. L. Jr and Johnson, M. N. (1974). Angiosarcoma of liver in the manufacture of polyvinyl chloride. *Journal of Occupational Medicine*, **16**, 150–151.

Doll, R. and Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute*, **66**, 1191–1308.

IARC (1972–2000). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Vols 1–78 (International Agency for Research on Cancer, Lyon).

Kogevinas, M., *et al*. (1994). *Occupational Exposure to Carcinogens in Developing Countries*. IARC Scientific Publication No. 129. 63–95 (International Agency for Research on Cancer, Lyon).

Mulshine, J. L. and Henschke, C. L. (2000). Prospects for lung-cancer screening. *Lancet*, **355**, 592–593.

Peto, J., *et al*. (1999). The European mesothelioma epidemic. *British Journal of Cancer*, **79**, 666–672.

Pott, P. (1775). *Chronological Observations Relative to the Cataract, the Polypus of the Nose, the Cancer of the Scrotum, the Different Kinds of Ruptures, and the Mortification of the Toes and Feet* (Clarke and Collins, London).

Schulte, P., *et al*. (1990). Final discussion: where do we go from here? *Journal of Occupational Medicine*, **32**, 936–945.

Tomatis, L. (1990). *Cancer: Causes, Occurrence and Control*. IARC Scientific Publication No. 100 (International Agency for Research on Cancer, Lyon).

Vineis, P. and Simonato, L. (1991). Proportion of lung and bladder cancers in males resulting from occupation: a systematic approach. *Archives of Environmental Health*, **46**, 6–15.

FURTHER READING

Andersen, A., *et al*. (1999). Work-related cancer in the Nordic countries. *Scandinavian Journal of Work and Environmental Health*, **25**, Suppl. 2, 1–116.

Boffetta, P. and Merler, E. (1999). Occupational cancer in Europe. *Environmental Health Perspectives,* **107**, Suppl. 2, 227–298.

Boffetta, P., *et al*. (1998). Occupational carcinogens. In: Stellman, J. M. (ed.), *Encyclopedia of Occupational Health and Safety*, 4th edn, Vol 1, 2.4–2.8 (International Labour Office, Geneva).

Stellman, J. M. and Stellman, S. D. (1996). Cancer and the workplace. *CA Cancer Journal for Clinicians*, **46**, 70–92.

Websites

http://www.occuphealth.fi/list/data/CAREX/. Exposure information on occupational carcinogens.

http://www.monographs.iarc.fr/. Information about carcinogens.

Chapter 31

Antigenotoxins and Cancer

Wilbert H. M. Peters and Esther M. M. van Lieshout
St Radboud University Hospital, Nijmegen, The Netherlands

CONTENTS

- Introduction
- Antigenotoxins
- Test systems for Antigenotoxins
- Mechanisms by which Antigenotoxins may Act
- Different Classes of Antigenotoxins
- Conclusions and Future Perspectives

INTRODUCTION

It is now widely accepted that the bulk of all cancer is related to environmental factors such as lifestyle practices; smoking, dietary habits, physical exercise, etc. In addition, cancer is one of the degenerative diseases of old age and increases dramatically with age. In a minority of cases, varying between 5 and 30%, depending on the type of cancer, genetic factors may contribute to cancer development, and in certain rare forms of human cancer hereditary factors play a decisive role. Now that such knowledge is becoming increasingly available, it is encouraging that theoretically most cancers could be prevented. Recent (epidemiological) studies have identified many factors that are likely to reduce cancer rates, and many others will follow. Therefore, prevention of cancer is the main goal for many researchers and clinicians in this new era (Anonymous, 1999).

The first method of choice in cancer prevention is to reduce exposure to potentially DNA-damaging sources by changing the lifestyle with respect to smoking, sun exposure, physical activity, diet, etc. Dietary factors alone have been estimated to account for approximately 30% of cancer risk, especially for cancers of the gastrointestinal tract. Two different types of dietary factors should be clearly discerned: those which initiate or promote carcinogenesis and those which are preventive against cancer. Of course, our diet should be formulated such that the latter substances dominate. Therefore, it is of utmost importance to gain more knowledge about both types of factors in order to be able eventually to modify our diets in favour of preventive substances and thus reduce cancer incidence in the future. In this chapter we will focus on factors (antigenotoxins) that are preventive in carcinogenesis.

According to the stage in the carcinogenic process at which they are effective, Wattenberg (1985) classified preventive compounds into three different categories: (1) compounds that prevent the formation of carcinogens, (2) compounds that inhibit carcinogenesis by preventing carcinogens from reaching and reacting with critical targets such as DNA (these inhibitors were called ‘blocking agents’) and (3) compounds acting subsequent to exposure to carcinogens. These compounds appear to inhibit the carcinogenic process after initiation, and were called ‘suppressing agents’ since they act by suppressing neoplasia in the exposed tissues. In **Table 1** some examples of these three classes of inhibitors are shown.

Before a normal cell is transformed into a cancer cell, it has to be irreversibly damaged, e.g. by a genotoxin. A genotoxic substance can be defined as a compound that is damaging to DNA and thereby may cause a mutation or cancer. Consequently, an antigenotoxin prevents or restores damage to DNA, and therefore substances stimulating DNA repair mechanisms may also be considered as antigenotoxins. In a strict sense, ‘suppressing agents’ should be called anticarcinogens rather than antigenotoxins. However, since the mechanism of many anticarcinogens, compounds preventing cancer, is still unclear, we will use the words antigenotoxin and anticarcinogen interchangeably. Adding to this complexity is the occurrence of species and test system dependences with respect to antigenotoxic properties of a certain compound, i.e. a particular substance has antigenotoxic characteristics in test system A but not in test system B; it will still be denoted an antigenotoxin. Several test systems derived from a variety of species will be discussed later.

ANTIGENOTOXINS

Antigenotoxins can be divided into endogenous and exogenous antigenotoxins according to whether they can be formed somewhere in the human body or are taken up from outside, such as food constituents or food additives.

Table 1 Inhibitors of carcinogen-induced neoplasias. (Adapted from Wattenberg, 1985.)

Category of inhibitor	Chemical class	Inhibitory compound
Compounds preventing formation of carcinogen	Reductive agents	Vitamin C[a]
	Tocopherols	α-Tocopherol[a], γ-tocopherol[a]
	Phenols	Caffeic acid[a], ferulic acid[a], gallic acid[a]
Blocking agents	Phenols	t-Butylhydroxyanisole[b], butylated hydroxytoluene[b],ellagic acid[a], caffeic acid[a], ferulic acid[a]
	Indoles	Indole-3-acetonitrile[a], indole-3-carbinol[a]
	Coumarins	Coumarin[a], limettin[a]
	Flavones	Quercetin[a], rutin[a], catechin[a]
	Aromatic isothiocyanates	Benzyl isothiocyanate[a], phenyl isothiocyanate[a]
	Dithiolthiones	Oltipraz[b]
Suppressing agents	Retinoids and carotenoids	Retinyl palmitate[a], retinyl acetate[a], β-carotene[a]
	Protease inhibitors	Soybean protease inhibitors[a]
	Inhibitors of arachidonic acid metabolism	Indomethacin[b], aspirin[b]
	Phenols	t-Butylhydroxyanisole[b],
	Methylated xanthines	Caffeine[a]
	Plant sterols	β-Sitosterol[a]
	Selenium salts	Sodium selenite[a], selenium dioxide[a], selenious acid[a]

[a]Naturally occurring compound present in food or formed during digestion.
[b]Synthetic compound.

Endogenous Antigenotoxins

Evidence is increasing that DNA damage can occur as a result of normal cellular functions. There are three main sources of DNA damage. (1) many cellular processes require or consume oxygen. Sometimes oxygen and other small molecules such as hydroxyl (OH) groups become electron deficient and may escape normal cellular pathways such as those of the respiratory chain, and these oxygen radicals can react with DNA and may cause mutations or cancer. (2) A second very important source of oxygen radicals in the body is the normal burst of free radicals generated by neutrophils, in order to kill bacteria, viruses, etc., entering the body, as part of the normal defence against such intruders. (3) In the normal metabolism of endogenous or exogenous molecules, often very reactive metabolites are formed intracellular by cytochrome P450-mediated reactions. Sometimes such metabolites may be mutagenic or carcinogenic and need to be further metabolized or detoxified by appropriate endogenous systems.

Most organisms have endogenous systems that prevent DNA damage caused by mechanisms as mentioned above. Such systems and their enhancers may belong to the class of antigenotoxins.

Some endogenous antigenotoxins, as outlined above, often have antioxidant or radical-scavenging properties. These can be relatively simple molecules or very complex proteins or enzyme systems. (See also chapter on *Mechanisms of Chemical Carcinogenesis*.)

Simple Molecules

Relative simple endogenous molecules with antioxidant and potential antigenotoxic properties include the metabolites uric acid, porphyrins and bilirubin. The most important endogenous antigenotoxin, however, is probably the tripeptide glutathione and to a lesser extent other thiols such as cysteine and cysteinylglycine. These molecules are able to scavenge free radicals or other oxidative compounds directly. In addition, glutathione is an important cofactor in many enzymatic detoxification reactions.

Complex Molecules

More complex endogenous antigenotoxins include the proteins or enzymes that are involved in the rapid detoxification of reactive metabolites resulting from oxidative stress or from phase I enzyme-mediated reactions (cytochrome P450). Important enzymes are glutathione peroxidase, glutathione *S*-transferase (GST), catalase and superoxide dismutase, but many others can also be involved.

Exogenous Antigenotoxins

Exogenous antigenotoxins can be divided into synthetic and naturally occurring substances present in food or food constituents or additives.

Synthetic Antigenotoxins

Most antigenotoxins are taken up via consumption of food, and more especially from fruit and vegetables as part of the diet. Recently however, synthetic antigenotoxins such as nonsteroidal anti-inflammatory drugs (NSAIDs) such as aspirin, indomethacin, sulindac and oltipraz have been recognized. Also food supplements in the form of tablets containing vitamins, minerals, trace elements and other additives can be included in this group. NSAIDs have been shown to prevent the growth and formation of colon adenomas or carcinomas in several epidemiological and intervention studies. The mechanism of this protective effect may be very complex and multifactorial but enhancement of detoxification enzymes by NSAIDs was evident. Oltipraz, which is both an inhibitor of activating systems and an enhancer of detoxification enzymes, was shown to have strong antigenotoxic potential by preventing aflatoxin B_1-induced DNA damage (**Figure 1**). Aflatoxin B_1 is a fungal toxin, thought to be the cause of many liver cancers in China and other parts of the world.

Naturally Occurring Antigenotoxins from Food

From epidemiological studies it has now been firmly established that individuals who consume diets rich in fruits and vegetables have a reduced cancer risk (Steinmetz and Potter, 1991). Many chemical substances from fruits and vegetables with antigenotoxic or antimutagenic properties have been identified, but the bulk of the potential anticarcinogens from these sources remain to be identified. Since the mechanism of most of these food-derived anticarcinogens is not known in detail, it cannot be clearly established which compound has purely antigenotoxic properties or is an otherwise cancer-preventive compound, such as a stimulator of apoptosis, inhibitor of cell proliferation, inhibitor of metastasis, etc.

Antigenotoxins from fruit and vegetables can be divided into vitamins, minerals and trace elements, and other compounds based on their putative working mechanism.

Vitamins (see Odin, 1997)

Several vitamins have been shown to have antigenotoxic potential, including vitamins A, C, E, B_2, B_6, B_{11} and B_{12}.

Vitamin A (retinol) and its related structures are constituents of many orange- and red-coloured fruits and vegetables. They have shown antigenotoxic properties in many different test systems and experiments, mainly owing to their free radical scavenging properties. However,

Figure 1 Effect of oltipraz on the metabolism of aflatoxin B_1.

it should be mentioned that vitamin A and other retinoids appear to have genotoxic properties, as was shown in some test systems. In addition, high doses of vitamin A may be toxic owing to hepatic storage, and these latter properties may possibly explain the disappointing results so far in cancer chemoprevention trials using high doses of this vitamin (see the chapter *Intervention and Chemoprevention of Cancer*).

Vitamin C or ascorbic acid is also a constituent of many fruits and vegetables. In analogy with vitamin A, the antigenotoxic properties of vitamin C are mainly due to the scavenging of radicals, and high concentrations have also been to have mutagenic properties in some test systems. It is now generally accepted that this water-soluble vitamin is beneficial unless very high doses (up to 10 g daily) are consumed.

Vitamin E. The fat-soluble vitamin E is a naturally occurring mixture of several tocol derivatives among which α-tocopherol is the most biologically active. The basic function of vitamin E in all living organisms is as an antioxidant in protection against oxidative damage of hydrophobic molecules or structures, generated either by free radicals or by radical-inducing radiation. Thus, by preventing radical-induced damage, vitamin E was shown to be an antimutagen or antigenotoxin in many experimental studies.

Folic acid or vitamin B_{11}. Dietary factors which directly influence methyl group availability, in particular folic acid and the essential amino acid methionine, may also be associated with cancer incidence. Folic acid, which is a component of many green vegetables such as beans, spinach and Brussels sprouts, appears to be an important antigenotoxin. Several epidemiological studies have shown that people consuming large amounts of folic acid are better protected against cancer of the colon and possibly also of the smoking-related cancers of the lungs, head and neck area than people with a low intake (Glynn and Albanes, 1994). Folic acid is an essential factor in a number of critical pathways in the cell that involves the transfer of one-carbon groups, such as methyl groups, e.g. in the biosynthesis of DNA building blocks. In this way folic acid plays a key role in DNA replication and cell division and sufficient levels of folate do prevent mutations in DNA and thus may prevent cancer.

Vitamin B_2 or riboflavin has been shown to prevent binding of metabolically activated and highly carcinogenic aflatoxins to DNA. It also inhibits the damaging or mutagenic effects of irradiation or cigarette smoke condensate in *in vitro* test systems.

Vitamins B_6 and B_{12} have an important role in the conversion of homocysteine into methionine and in other steps of the folate metabolism pathway, ultimately leading to an optimal supply of methyl groups, which is essential for prevention of DNA damage.

Minerals and Trace Elements

Minerals such as calcium and phosphorus and trace elements such as zinc, selenium and molybdenum have important functions in the synthesis of biomolecules and tissue structures, and are essential in numerous cellular and enzymatic processes. Antigenotoxic properties of these minerals and trace elements have attracted little attention. In many studies, however, deficiencies of the trace element selenium have been associated with an increased cancer risk. Since selenium is an essential component of the selenium-dependent glutathione peroxidases, which detoxify highly reactive peroxides and thus may prevent oxidation and damage to DNA, it may be considered an antigenotoxin.

Other Compounds

Most antigenotoxins cannot be categorized into the above-mentioned groups. Several hundred agents have been identified as anticarcinogens with potential antigenotoxic properties, mainly based on data derived from one of the following study categories: (1) basic mechanism; (2) observational studies in laboratory animals; (3) studies in humans; and (4) selective screening systems. With the exception of antipromoter, antiprogression or tumour cell-killing activity, most anticarcinogens are antigenotoxins.

Apart from vitamins and trace elements, plant foods such as fruits, vegetables and cereals contain many bioactive compounds with antigenotoxic properties. Since these constituents do not belong to the classes of proteins, fats or carbohydrates, they are often called non-nutrients, microconstituents, phytochemicals or phytoprotectants. Exogenous antigenotoxins of plant origin can be roughly categorized as follows:

- Sulfides present in the allium vegetables which include onions, garlic and chives.
- Dithiolthiones, which are also sulfur-containing compounds of a more complex structure, are found in cruciferous vegetables such as cauliflower, cabbage and Brussels sprouts.
- Glucosinolates are also found in cruciferous vegetables. During cooking and chewing these glucosinolates are converted into the bioactive metabolites isothiocyanates and indoles.
- Terpenoids such as (+)-limonene, which are components of citrus fruits.
- Phyto-oestrogens including isoflavones and lignans. Isoflavones are found in cereals and pulses such as soy bean. The main dietary sources of lignans are wholegrain products, seeds, fruits and berries.
- Flavonoids such as quercetin, kaempferol, rutin, tangeritin and myricetin are widely distributed in fruits and vegetables. Rich sources are berries, potatoes, tomatoes, onions, broccoli, beans and citrus fruits.
- Other phenols and polyphenols, such as ellagic acid, caffeic acid, ferulic acid, resveratrol are found in nuts, fruits, wine and tea.

TEST SYSTEMS FOR ANTIGENOTOXINS

In the search for and characterization of antigenotoxic, substances, both *in vitro* and *in vivo* test systems can be used. *In vitro* test systems can be divided into (1) tests with bacterial cultures such as the Ames test (Ames *et al*., 1973), (2) tests with stable cell lines of animal or human origin, (3) short-term organ or cell cultures of both animal and human origin and (4) *in vivo* experiments, which include laboratory animal and human clinical trials or intervention studies. **Figure 2** shows a schematic overview of the strategy for detecting antigenotoxins.

There are two main problems with these studies: (1) data obtained in bacterial or animal test systems (both *in vitro* and *in vivo*) cannot automatically be extrapolated to humans; and (2) human studies are difficult to perform since the potential anticarcinogenic/genotoxic test compounds may themselves be toxic or mutagenic.

In Vitro Test Systems

In general, a substance selected on promising epidemiological data is tested at first using *in vitro* systems, since these tests are most easiest to interpret, the most readily available and the cheapest.

The most common test employed is the Ames test, where the effects of added substances on bacterial mutation rates are tested (see later).

Another category of *in vitro* tests use cell lines, which most often are derived from human or animal tumour cells, but sometimes originate from foetal cells or virus-transformed and thereby immortalized ‘normal’ cells. By adding the substance of interest to the culture medium the eventual antigenotoxic effect can be studied at several levels as outlined later.

In vitro test systems based on short-term cultures of slices of normal tissue or freshly isolated cells from patients or controls have been helped by the development of very sensitive detection methods for monitoring changes in DNA, RNA or other important biomolecules. Since such studies use normal human tissues, misinterpretation of data extrapolated from results in abnormal cells such as tumour-derived cell lines is avoided as much as possible.

In Vivo Test Systems

Next, when the *in vitro* data on compound efficacy combined with toxicity data are convincing, *in vivo* testing in laboratory animals is generally performed. Preferentially, different species should be used in parallel, in order to exclude incidental positive or negative results in one strain of animals.

Traditionally, laboratory animals are treated with various chemical carcinogens generating tumours at most common epithelial sites and used for the evaluation of antigenotoxic or anticarcinogenic test agents. Potential antigenotoxic test compounds can be added before, during or after the exposure to the carcinogen and in this way information on the working mechanism is also obtained.

More recently, genetically modified animal models that mimic specific characteristics of human risk groups or diseases have been used, specifically, transgenic animals, in which new properties have been introduced into the genome by genetic engineering, and gene knockout animals in which specific genes, and consequently specific functions, have been knocked out. Key aspects in evaluating the use of these experimental animals in chemoprevention research is the match between the (modified) animal model and the human disease entity: the closer the match, the more reliable the information obtained will be for transfer to the human situation.

There are two types of human *in vivo* study: (1) human clinical trials and (2) human intervention studies. Human clinical trials with antigenotoxic substances, owing to their high costs, intensive preparation, guidance and long duration, have rarely been performed (see the chapter *Intervention and Chemoprevention of Cancer*). In addition, there are often ethical problems due to the unknown properties of the test compounds with respect to short- or long-term toxicity. There are very few human intervention studies of short-term dietary supplementation with potential anticarcinogens which have examined biomarkers for genetic damage.

MECHANISMS BY WHICH ANTIGENOTOXINS MAY ACT

Many carcinogens are not directly active in their parental form but are metabolically activated to highly reactive

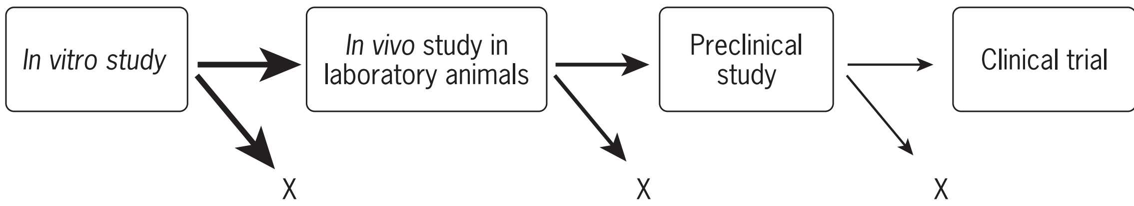


Figure 2 Common strategy for detecting and testing potential antigenotoxic compounds. X = compounds not suitable. (Adapted from Ito and Imaida, 1992.)

intermediates, which are the ultimate carcinogenic species (Miller and Miller, 1977). These compounds have gained the ability to react with DNA and to cause mutations that subsequently might lead to cancer. Approximately 90% of the chemical compounds that are mutagenic are also carcinogenic (Ames *et al*., 1973). The ultimate carcinogens are electrophilic molecules that bind to electron-rich DNA, forming covalently bound adducts. Since this modified DNA is error-prone, subsequent replication can result in mutations, which are structural changes in the genetic material that are inherited by the daughters of these damaged cells. Different kinds of mutations can be discriminated: substitution of an incorrect base, the so-called point mutation; addition or deletion of genetic material, the so-called frameshift or missense mutation; or even DNA strand breakage with resultant aberrations in the pattern and organisation of the chromosomes, such as karyotypic abnormalities, sister chromatid exchange or micronucleus formation. When repair of the damage does not occur, replication of DNA may lead to permanent DNA lesions. In the presence of a tumour promoter these cells may change into pre-neoplastic cells and eventually neoplastic cells and may subsequently gain the potential to metastasize.

There are many ways of assessing genotoxicity. For example, one can extract DNA from tissues or cells and quantify the amount of DNA adducts formed. Mutations, chromosomal aberrations and DNA strand breaks can be examined at the DNA level by polymerase chain reaction, DNA sequencing and several other methods. Another two examples of genetic alterations are sister chromatid exchange and micronucleus formation. Sister chromatid exchange takes place due to crossing-over events during both mitosis and meiosis and results in two chromatids that no longer contain the same genes. Several techniques allow the visualization of this exchange on a cellular base. Micronucleus formation reflects the amount of DNA damage. Using staining techniques it is possible to visualize gross changes in number and sizes of individual chromosomes.

In addition to genetic changes, it is also important to study proliferation rates of cell populations. The proliferating cell fraction can be determined by bromodeoxyuridine (BrdU) incorporation and by immunostaining for MIB-1, Ki-67 or proliferating cell nuclear antigen (PCNA), three antigens that are present during different phases of the cell cycle. Proliferation is also reflected in nucleolar organizer regions (NORs), which are loops of DNA that contain ribosomal RNA genes that are involved in protein synthesis. The proteins that are associated with these NORs can be visualized using a silver staining technique. The resulting silver reaction products (AgNORs) are visible as black dots. The number and size of AgNOR dots in a nucleus reflect cell proliferation. Several studies have demonstrated that AgNOR and PCNA both increase during the progression from normal epithelium to hyperplasia, dysplasia and carcinoma *in situ*. More specifically, they may be indicative of the stage of the carcinogenesis process. To gain more objective results, computer programs have been designed to analyse staining results.

The parameters mentioned above are generally not direct targets for antigenotoxins, but can be used to assess the beneficial effects of antigenotoxic compounds.

A large number of antigenotoxic and anticarcinogenic agents have been detected, some of which have proved effective in clinical trials. The multistage nature of carcinogenesis offers the possibility for intervention at each stage of the process, as shown in **Figure 3**. At the initiation stage chemoprevention can be accomplished by inhibition of metabolic activating enzymes or the enhancement of detoxification enzymes, radical scavengers and binding

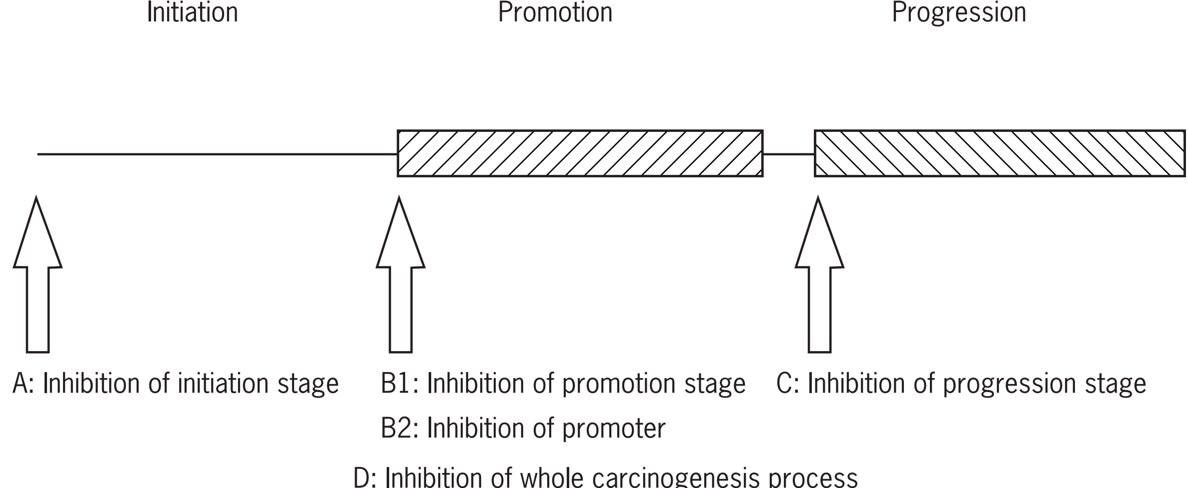


Figure 3 Classification of antigenotoxins and anticarcinogens based on the stage where they exert their protective effects. (Adapted from Ito and Imaida, 1992.)

capacity to active chemical carcinogens (A). At the promotion stage proposed mechanisms for chemoprevention include inhibition of DNA synthesis, cell proliferation and alteration of cell differentiation and communication as well as signal transduction pathways (B). At the progression stage, inhibition of cell proliferation has been proposed (C). Further, some compounds have chemopreventive capacity over the whole carcinogenic process (D). Potential antigenotoxic compounds need to be tested extensively before conclusions with respect to their suitability and applicability as antigenotoxin can be drawn, as was described in the previous section. Many of the target molecules by which antigenotoxins and anticarcinogens act are described below.

DIFFERENT CLASSES OF ANTIGENOTOXINS

Inhibitors of Carcinogen Formation

With the exception of preventing intake, preventing carcinogen formation is by far the most effective and desirable way of controlling cancer, e.g. preventing the formation of nitrosamines from secondary amines and nitrite in an acid environment, such as present in the stomach. Ascorbic acid, ferulic acid, gallic acid, caffeic acid, *N*-acetylcysteine, proline and thioproline can all exert antigenotoxic effects by inhibiting carcinogen formation.

Blocking Agents

Many procarcinogens need metabolic activation in order to react with DNA. Chemoprevention can be accomplished by interfering with the metabolism of carcinogens or mutagens. This process of biotransformation can be divided into phase I and phase II reactions. In phase I metabolism, oxidation, reduction or hydrolysis reactions transform endogenous or exogenous molecules into more hydrophilic compounds that are readily accessible for phase II reactions. The most important phase I biotransformation enzymes are the cytochromes P450 (Guengerich, 1988). In phase II, parent molecules or their derivatives are conjugated with endogenous compounds such as UDP-glucuronic acid and glutathione, catalysed by the phase II enzymes GSTs and UDP-glucuronyltransferases, respectively. The net result of biotransformation is a polar and water-soluble compound that is readily excreted in urine or faeces and unable to bind covalently with DNA. **Figure 4** shows a schematic overview of the biotransformation of procarcinogens. The antigenotoxins that prevent the carcinogens from reaching and reacting with critical targets such as DNA are called blocking agents. Most blocking agents can be assigned to one or more of the following major categories (**Table 2**).

Inhibitors of Cytochrome P450 Enzymes

Inhibitors of cytochrome P450 enzymes will reduce the activation of procarcinogens (Rendic and Di Carlo, 1997).

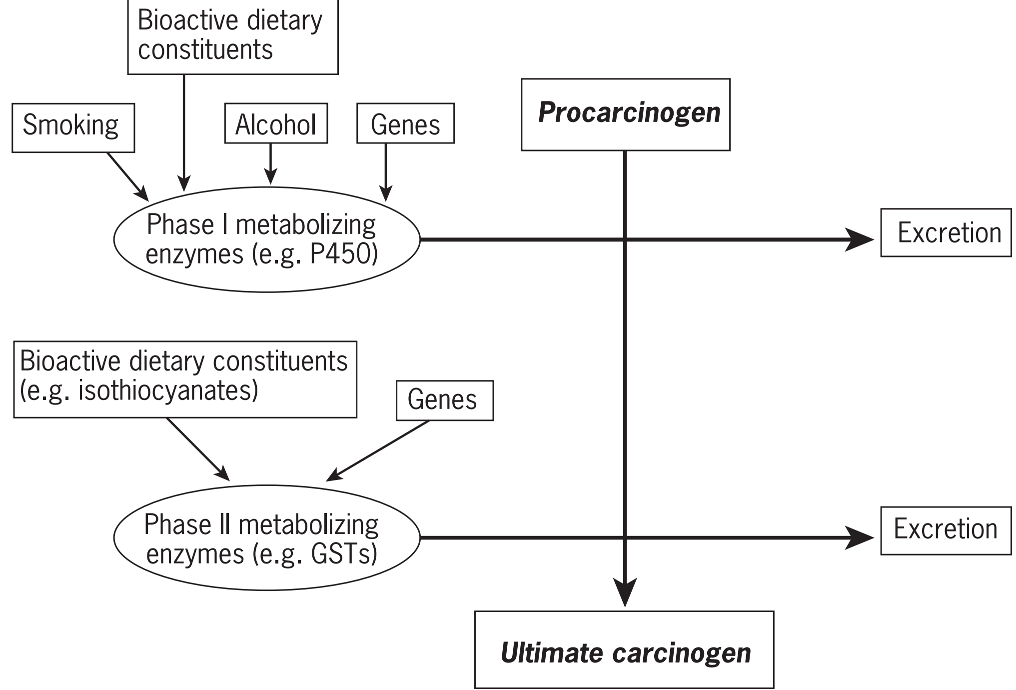


Figure 4 Biotransformation of procarcinogens. (Adapted from *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. American Institute for Cancer Research/World Cancer Research Fund, Washington, DC, 1997.)

Table 2 Working mechanisms and examples of blocking agents. (Adapted from Stoner *et al.*, 1997.)

Mechanism	Examples
Inhibition of cytochrome P450	Ellagic acid, diallyl sulfide, isothiocyanates
Induction of cytochrome P450	Indole-3-carbinol
Induction of phase II enzymes:	
Glutathione *S*-transferase	Allyl sulfides, dithiolthiones, isothiocyanates
UDP-glucuronyltransferase	Polyphenols
Glutathione peroxidase	Selenium
Scavenging of electrophiles	Ellagic acid, *N*-acetylcysteine, glutathione
Scavenging of free radicals	Glutathione, polyphenols, vitamin E
Increased overall levels of DNA repair	Vanillin
Increased poly(ADP-ribosyl) transferase	*N*-Acetylcysteine
Suppression of error-prone DNA repair	Protease inhibitors

As a consequence, less genetic damage will occur. This category of blocking agents consists mainly of dietary anticarcinogens such as ellagic acid, diallyl sulfide and isothiocyanates.

Inducers of Cytochrome P450 Enzymes

Inducers of cytochrome P450 such as indole-3-carbinol may increase the metabolic activation of procarcinogens (Rendic and Di Carlo, 1997). Ironically, this often is a necessary step in the detoxification process. In other words, more ultimate carcinogenic derivatives are formed, but they are constructed in such a way that phase II biotransformation enzymes can dispose of them.

Inducers of Phase II Enzymes

Inducers of phase II detoxification enzymes might be more beneficial over inducers of cytochrome P450 because inducers of phase II systems are less likely to result in highly bioactive compounds. Of utmost importance are enhancers of GST activity (Hayes and Pulford, 1995). GSTs catalyse the conjugation between electrophilic highly reactive compounds with glutathione, giving rise to a conjugate that can be excreted via urine or faeces. Many dietary anticarcinogens such as vitamins, indole-3-carbinol, limonene and phenyl isothiocyanate are inducers of GSTs in the gastrointestinal tract (Van Lieshout *et al.*, 1996, 1998), but also NSAIDs such as aspirin, sulindac and indomethacin have this capacity (Van Lieshout *et al.*, 1997).

Other important phase II enzymes include the glutathione peroxidases, which can be divided into selenium-dependent and non-selenium-dependent enzymes (Mannervik, 1985). Selenium is a biological trace element that is necessary for cell growth. Selenium has preventive potential in carcinogenesis: it may stimulate the immune system, effect carcinogen metabolism, inhibit protein synthesis, induce apoptosis and protect against oxidative damage. Selenium-dependent glutathione peroxidases catalyse the rapid detoxification of very reactive peroxides such as hydrogen peroxides and lipid peroxides. The non-selenium-dependent activity mainly involves the GSTs.

Under some circumstances, phase II enzymes can have a stimulating effect on carcinogenesis. For example, some products of *N*-acetyltransferase-, sulfotransferase- and GST-mediated reactions have been associated with a higher cancer risk, mainly from epidemiological studies.

Scavengers of Electrophiles and Free Radicals

Free radicals may damage lipids, proteins, DNA and other important biomolecules and thereby contribute to the initiation of cancer. Free radicals may arise exogenously, e.g. in cigarette smoke, or endogenously as a consequence of inflammation or as a by-product of normal respiration. Scavengers or trapping agents are compounds that physically react with electrophilic forms of carcinogens and oxygen free radicals, preventing them from damaging DNA or other molecules. Vegetables and fruit contain natural antioxidants such as vitamin C and E, carotenoids and flavonoids, which appear to be good radical scavengers. Oxidative damage can be detected in the form of oxidized bases in the DNA. A common example is the formation of 8-oxodeoxyguanosine (8-oxo-dG) from its parent compound deoxyguanine in human DNA.

Inducers of DNA Repair

If a procarcinogen has been metabolically activated into an electrophilic ultimate carcinogen, the latter compound can initiate carcinogenesis by forming DNA adducts and subsequent somatic alterations of oncogenes, tumour-suppressor genes and DNA repair genes (**Figure 5**). Damaged DNA can be repaired to prevent the formation of (pre)neoplastic cells. There are three possible chemopreventive mechanisms that involve DNA repair. The first is an increase in the overall level of DNA repair enzymes, e.g. by vanillin. Second, the enzyme poly(ADP-ribosyl) transferase is involved in modulation of DNA damage and the level of this enzyme is reduced by chemical carcinogens, but may be stabilized by antigenotoxins. *N*-Acetylcysteine is an example of this group of antigenotoxins that act by preventing the decrease in poly(ADP-ribosyl) transferase. A third mechanism is found in bacteria, where proteases specifically activate error-prone repair systems, and in this respect protease inhibitors such as soy bean trypsin inhibitors may be shown beneficial in the future.

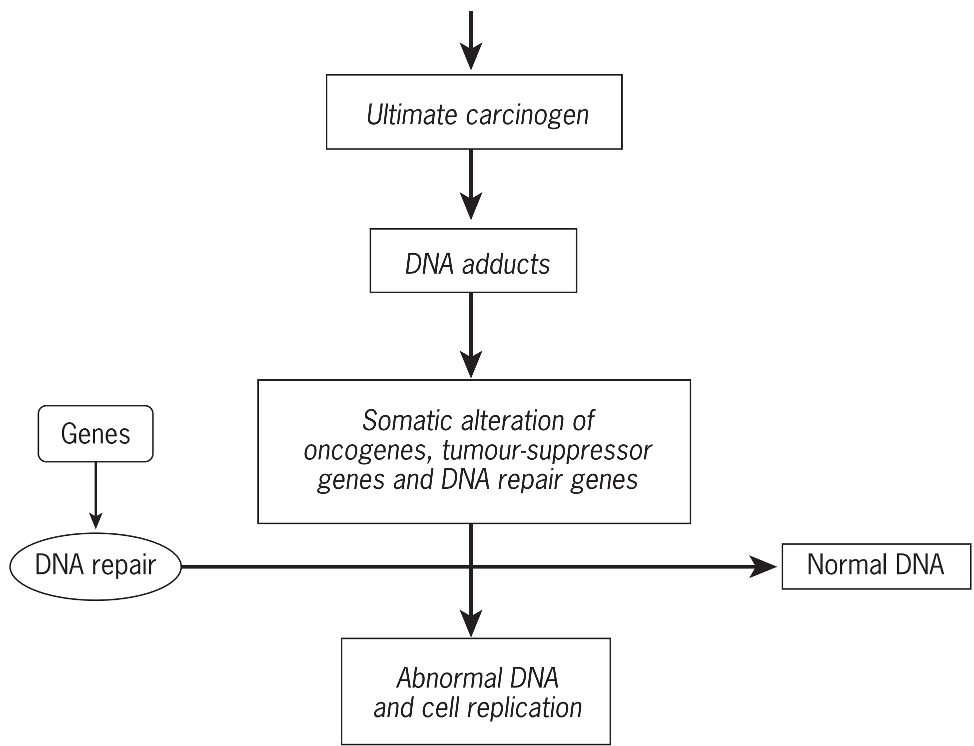


Figure 5 Initiation of carcinogenesis and the possibility of DNA repair. (Adapted from *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. American Institute for Cancer Research/World Cancer Research Fund, Washington, DC, 1997.)

Suppressing Agents

Classification of suppressing agents is more difficult because the critical events and their exact sequence in the process of tumour promotion and progression are not yet well understood. However, the classification proposed by Kelloff *et al*. (1994) (**Table 3**) is generally accepted. After completion of the initiation phase, tumour cells have gained the ability to replicate uncontrolled. **Figures 6** and **7** show schematic overviews of the effects of antigenotoxic and anticarcinogenic compounds on the promotion and progression phases of the carcinogenic process, respectively.

Inhibitors of Polyamine Metabolism

Polyamines play an important role in the regulation of cell proliferation and differentiation. A key enzyme in the polyamine biosynthetic pathway, ornithine decarboxylase (ODC), catalyses the conversion of ornithine to the polyamine putrescine. The levels of ODC and polyamines are frequently elevated in tumour tissues compared with their normal counterparts, indicating the role of ODC in stimulating or maintaining high proliferation rates. Owing to the rapid turnover of ODC of several minutes, high levels of an ODC inhibitor must remain present in the target organ to achieve the desired antiproliferative activity. Examples of this group of anticarcinogens are α-difluoromethylornithine (DFMO) and polyphenols.

Inducers of Terminal Cell Differentiation

Terminal differentiation is one of the steps in the normal, regulated cell proliferation of epithelia. Cancer cells often have lost the ability to differentiate. The carcinogenic process may be stopped by substances which restore the ability of abnormal proliferating cells to differentiate, such as calcium, retinoids and vitamin D_3.

Modulators of Signal Transduction

Signal transduction is a very complex cascade of events leading to a cellular response on extracellular signals. The components of signal transduction pathways are attractive sites for chemopreventive activity since they can result in restoration of normal cellular growth control. For example, one of the steps in signal transduction involves activation of protein kinase C by diacylglycerol. Several chemopreventive agents such as the flavonoids and glycyrrhetinic acid cause inhibition of protein kinase C activity, which leads to suppression of carcinogenesis.

Modulators of Hormonal/Growth Factor Activity

This group of chemopreventive agents inhibit neoplastic cell proliferation by directly regulating the levels and biological activity of specific hormones and growth factors. As a result of specific binding of these hormones and

Table 3 Working mechanisms and examples of suppressing agents. (Adapted from Stoner *et al.*, 1997 and Kelloff *et al.*, 1994.)

Mechanism	Examples[a]
Inhibit polyamine metabolism	DFMO, polyphenols, substituted putrescines, ellagic acid, curcumin
Induce terminal cell differentiation	Calcium, retinoids, vitamin A and D_3
Modulate signal transduction	Glycyrrhetinic acid, NSAIDs, polyphenols, retinoids
Modulate hormonal/ growth factor activity	NSAIDs, retinoids, tamoxifen
Inhibit oncogene activity	Genistein, NSAIDs, monoterpenes, (+)-limonene, retinoic acid
Promote intracellular communication	Carotenoids, polyphenols, retinoids
Restore immune response	NSAIDs, selenium, vitamin E
Induce apoptosis	Butyric acid, genistein, selenium, sulindac sulfone, retinoids
Correct DNA methylation imbalances	Folic acid, choline, methionine, vitamin B_{12}
Inhibit basement membrane degradation	Protease inhibitors
Inhibit arachidonic acid metabolism	Glycyrrhetinic acid, *N*-acetylcysteine, NSAIDs, polyphenols, flavonoids, vitamin E

[a]DFMO, α-difluoromethylornithine; NSAIDs, nonsteroidal anti-inflammatory drugs.

growth factors to their receptor, very distinct signal transduction pathways will be started. For example, transforming growth factor-β has antiproliferative activity in both normal and neoplastic cells *in vitro* and *in vivo*. In addition, overexpression of the epidermal growth factor receptor, a product of the *erb* oncogene, may be involved in the pathogenesis of certain epithelial neoplasms (Shin *et al*., 1994). For many growth factor receptors, antibodies are available that can be used for immunohistochemical staining and subsequent qualitative and quantitative analysis. NSAIDs and retinoids are among the compounds that have beneficial effects by modulating the activity of hormones or growth factors, e.g. by competing for binding sites to their receptors.

Inhibitors of Oncogene Activity

Protein inhibitors and retinoids exert their chemopreventive action by inhibiting oncogene expression. Oncogenes are genes that may cause cancer. Present in the genome as proto-oncogenes, they are not harmful as such. However, they can be converted to oncogenes when they acquire specific mutations. The *ras* oncogene protein, which plays a role in cell growth and differentiation, has been extensively studied. To be activated, the Ras protein must be farnesylated. The enzyme farnesyl protein transferase (FPTase) catalyses farnesylation of Ras precursors in a critical step during post-translational modification of Ras oncoproteins, resulting in their anchorage to the plasma membrane. FPTase inhibitors therefore inhibit Ras-induced colonic tumorigenesis. Singh *et al*. (1998) demonstrated that ingestion of a high amount of corn oil enhances FPTase expression. This in turn leads to increased levels of functional Ras and promotes chemically induced colon carcinogenesis. Dietary fish oil, on the other hand, suppresses FPTase expression and causes decreased production of biologically active Ras, thereby inhibiting colon tumorigenesis. FPTase and type I and type II geranylgeranyl-protein transferases (GGTases) catalyse protein prenylation reactions. As much as 2% of the cellular proteins are isoprenylated in certain cell types. Some of these proteins perform basic cellular functions, whereas others may be involved in the signal transduction pathways that regulate cell proliferation, cell differentiation and cell death. Several isoprenylated proteins, including Ras, which is a substrate of FPTase, RhoA and Rac1 (two substrates of type I GGTase), were shown to have transforming properties. Chemoprevention by interfering with prenylation reactions effects proliferation, differentiation and cell death. Examples of compounds capable of such activity are genistein, NSAIDs, (+)-limonene and retinoic acid.

Promoters of Intercellular Communication

Communication between cells is mediated through gap junctions. Gap junctions are pores or channels in the cell membrane, which join channels of adjacent cells and allow the passage of molecules up to approximately 1000 Da in size. Growth regulatory signal molecules are among the molecules that move between cells. It has been postulated that inhibition of gap junctional communication between cells occurs during carcinogenesis. Several carotenoids, such as β-carotene and canthaxanthine, and several retinoids, such as vitamin A, enhance gap junctional communication *in vitro* and inhibit cell transformation.

Restorers of Immune Response

The immune system may also be a target for chemopreventive agents. For example, retinoic acid increases natural killer cell cytotoxicity which can be involved in eliminating (pre)cancerous cells. Retinoids can also cause leukaemic promyelocytes to differentiate into mature granulocytes, which behave normally. Compounds, as mentioned above, that stimulate the elimination of transformed or otherwise abnormal cells may prove to be chemopreventive.

Inducers of Apoptosis

Apoptosis or programmed cell death has an important function in preventing cells from changing into cancer cells. A desirable way of chemoprevention is by stimulating specific

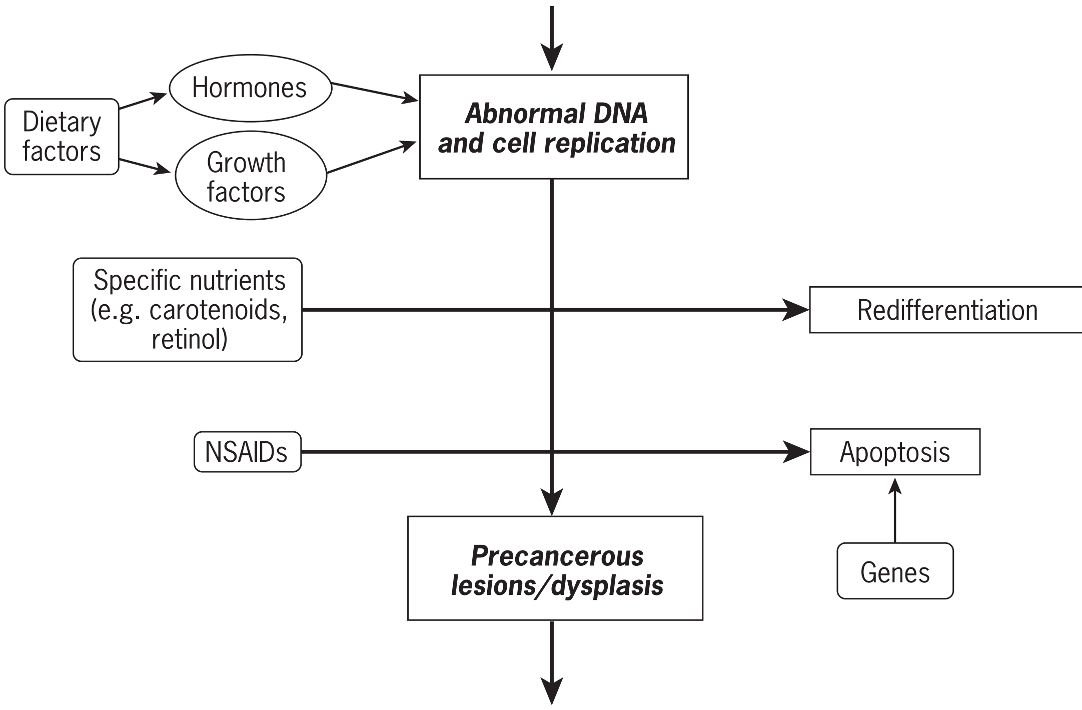


Figure 6 Promotion. (Adapted from *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. American Institute for Cancer Research/World Cancer Research Fund, Washington, DC, 1997.)

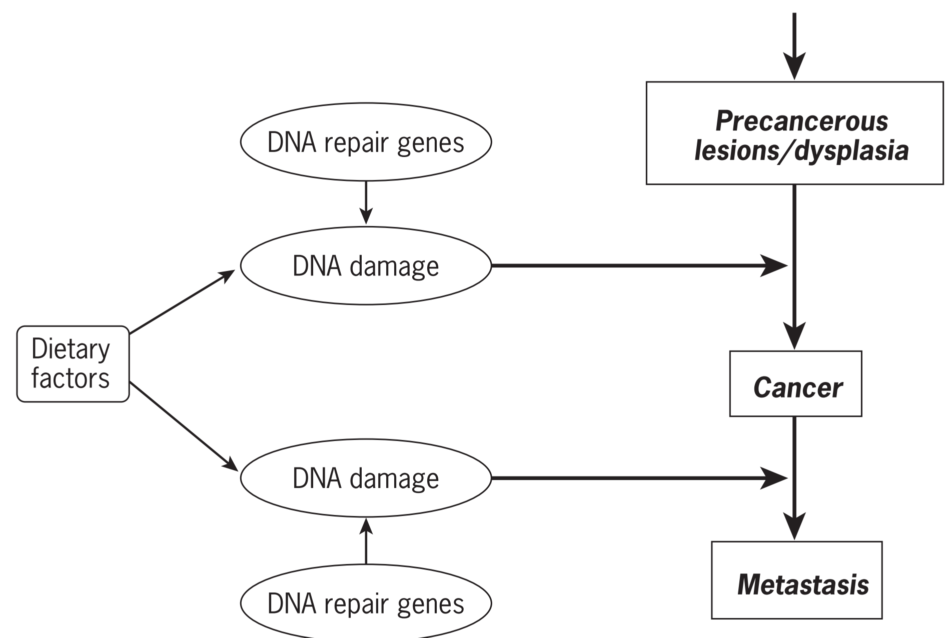


Figure 7 Progression. (Adapted from *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. American Institute for Cancer Research/World Cancer Research Fund, Washington, DC, 1997.)

elimination of potential cancerous cells by apoptosis. Tumour-suppressor proteins such as wild-type p53, and growth factors such as TGF-β have been implicated as inducers of apoptosis. The regulation of apoptosis is very complex. This complexity, on the other hand, also presents many options for interfering with apoptotic pathways. NSAIDs such as sulindac have been demonstrated to be powerful stimulators of apoptosis. Other compounds capable

of such activity include butyric acid, genistein, selenium and several retinoids.

Correctors of DNA Methylation Imbalances

A significant loss of methyl groups in DNA appears to occur very early in colorectal tumorigenesis; the smallest adenomas examined (a few millimetres in diameter) have already lost approximately 10–15 million methyl groups, compared with a similar amount of normal colon mucosa cells (Goelz *et al.*, 1985). Hypomethylated DNA may cause undercondensation of chromosomes during mitosis, which may lead to aneuploidy. Aneuploid cells contain an abnormal number of chromosomes, which enhances the cancer risk. Hypomethylation generally coincides with increased proto-oncogene expression and decreased expression of growth factors and their receptors. These observations imply that compounds such as folic acid are likely to be chemopreventative and indeed it has been shown that certain compounds that serve as methyl group donors inhibit tumorigenesis (see earlier). In addition, a number of studies have shown that methyl-deficient diets increase cell turnover and promote the development of carcinogen-induced liver tumours in rats and mice, whereas methyl-rich diets containing high levels of choline and of the amino acid methionine prevent or reduce these effects. Vitamin B_{12} is involved in the intracellular methyl metabolism and is therefore also considered an anticarcinogen.

Inhibitors of Basement Membrane Degradation

Cancer cells contain several enzymes that digest the basement membrane, allowing invasion through to normal tissues. Examples of these enzymes are the proteases collagenase, hyaluronidase, cathepsin B, elastase and plasminogen activators. Therefore, inhibitors of these proteases may prove to be effective in chemoprevention.

Inhibitors of Arachidonic Acid Metabolism

Among the multiple events that occur during experimentally induced tumour promotion is an increased metabolism of arachidonic acid, which contributes to an overall inflammatory response. Arachidonic acid, a fatty acid, is one of the components of the phospholipids, the normal building blocks of cellular membranes. The cyclooxygenase pathway converts arachidonic acid into prostaglandins, prostacyclins and thromboxanes, whereas lipoxygenase converts arachidonic acid into leukotrienes and hydroxyeicosatetraenic acids. During these processes, reactive oxygen and alkylperoxy species are formed, which may damage cellular structures or molecules such as DNA. Cyclooxygenase inhibitors such as the NSAIDs aspirin, indomethacin, ibuprofen and piroxicam and also certain antioxidants such as flavonoids are effective inhibitors of carcinogenesis, since they prevent the formation or cause the elimination of reactive oxygen species.

CONCLUSIONS AND FUTURE PERSPECTIVES

Although encouraging progress in identifying potential cancer chemopreventive or antigenotoxic compounds has been made, this field is in its earliest stages of development. The overview presented above is intended to give a general idea on what antigenotoxins are, how they can be discovered, selected and tested and how they work. However, one has to realize that (1) many potential antigenotoxins have so far only been tested in animals or animal cells, and care has to be taken in extrapolating these data to the human situation, (2) effects found may be strictly specific for a particular tissue or even for a particular type of cell or type of cancer, and thus an antigenotoxin with proven efficacy for prevention of colon cancer may be of no use at all in preventing gastric or oesophageal cancer, and (3) cancer can develop through a wide variety of mechanisms or by a cascade of events, and therefore chemoprevention may be most effective by using a combination of antigenotoxic substances. It will not be as simple as adding fluoride to drinking water for prevention of caries or adding iodine to bread for prevention of struma.

Figure 8 gives an overview of several aspects of the carcinogenic process, and also shows that antigenotoxins or anticarcinogens may interfere at many different levels, indicating once more that chemoprevention of cancer will be achieved most efficiently when a combination of different antigenotoxins is applied. This is probably why chemopreventive human trials performed until now have shown somewhat disappointing results. Another reason may be that the doses of the antigenotoxins used were inappropriate; for example, it has been shown many times in many different test systems that high doses of a particular compound can have deleterious effects, where low doses were beneficial.

Consuming a diet that contains a wide variety of fruits and vegetables provides a broad range of antigenotoxins and this may be the main reason why these diets have been associated with the prevention of a wide variety of different cancer types.

In conclusion, the search for and testing of chemopreventive compounds is a very promising activity which will lead to further reductions in cancer death rates by prevention of the disease, possibly when combinations of different antigenotoxins at low concentrations, are used. However, such research is time consuming, expensive and requires broad knowledge and intensive cooperation of specialists from several disciplines such as epidemiologists, tumour biologists, biochemists, clinicians and many others.

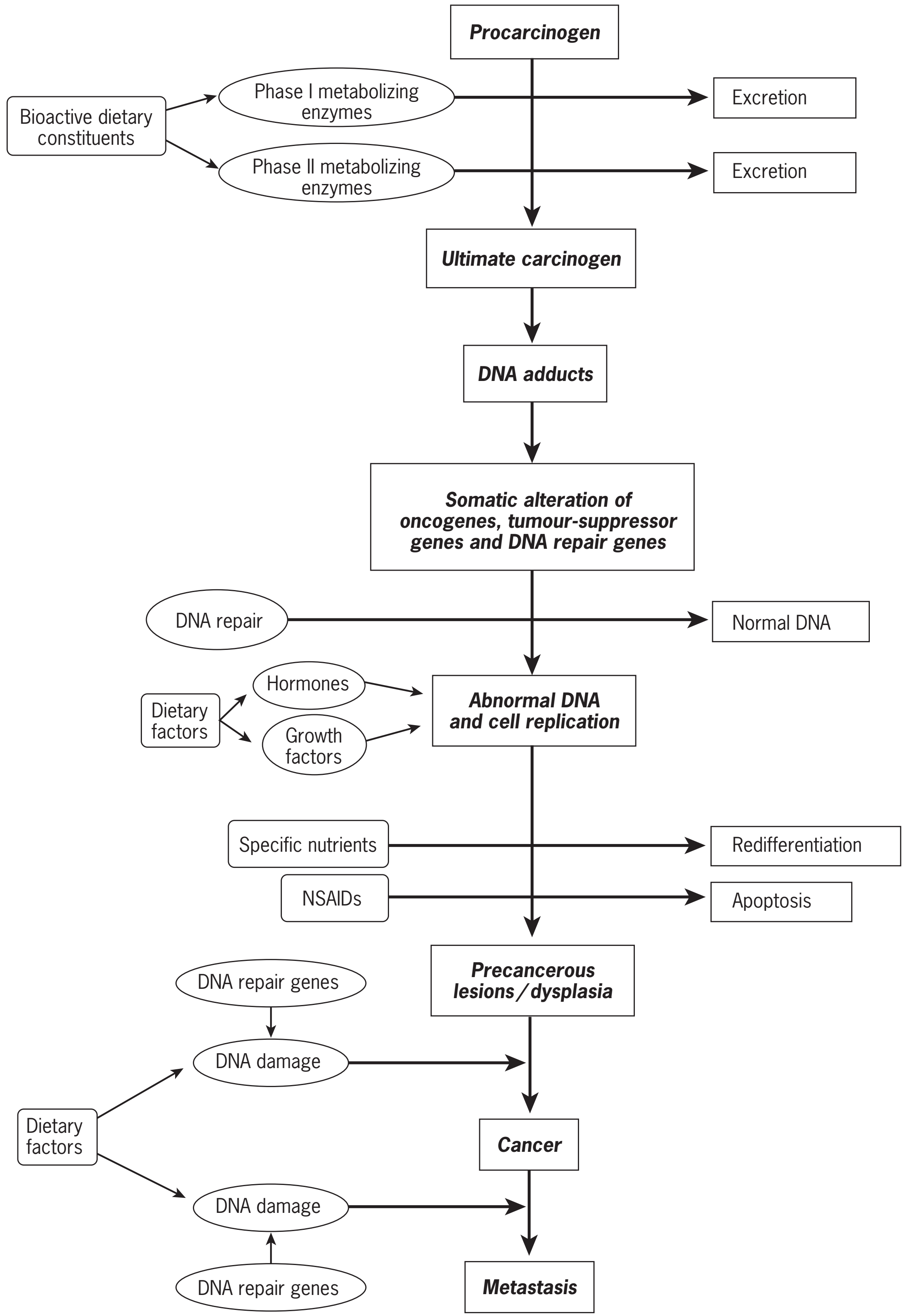


Figure 8 Target sites of antigenotoxins and anticarcinogens during the cancer process. (Adapted from *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. American Institute for Cancer Research / World Cancer Research Fund, Washington, DC, 1997.)

REFERENCES

Ames, B. N., *et al*. (1973). Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. *Proceedings of the National Academy of Sciences of the USA*, **70**, 2281–2285.

Anonymous (1999). Prevention of cancer in the next millennium: report of the chemoprevention working group to the American Association for Cancer Research. *Cancer Research*, **59**, 4743–4758.

Glynn, S. A. and Albanes, D. (1994). Folate and cancer: a review of the literature. *Nutrition Cancer*, **22**, 101–119.

Goelz, S., *et al*. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science*, **288**, 187–190.

Guengerich, F. P. (1988). Roles of cytochrome P-450 in chemical carcinogenesis and cancer chemotherapy. *Cancer Research*, **48**, 2946–2954.

Hayes, J. D. and Pulford, D. J. (1995). The glutathione *S*-transferase supergene family: regulation of GST and contribution of the isoenzymes to cancer chemoprevention and drug resistance. *Critical Reviews in Biochemistry and Molecular Biology*, **31**, 445–600.

Ito, N. and Imaida, K. (1992). Strategy of research for cancer-chemoprevention. *Teratogenesis Carcinogenesis and Mutagenesis*, **12**, 79–95.

Kelloff, G. J., *et al*. (1994). Mechanistic considerations in chemopreventive drug development. *Journal of Cellular Biochemistry* **20**, (Supplement), 1–24.

Mannervik, B. (1985). Glutathione peroxidase. *Methods in Enzymology*, **113**, 490–495.

Miller, J. A. and Miller, E. C. (1977). Ultimate chemical carcinogens as reactive mutagenic electrophiles. Cold Spring Harbor Conference. *Cell Proliferation*, **4**, 604.

Odin, A. P. (1997). Vitamins as antimutagens: advantages and some possible mechanisms of antimutagenic action. *Mutation Research*, **386**, 39–67.

Rendic, S. and Di Carlo, F. J. (1997). Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. *Drug Metabolism Reviews*, **29**, 413–580.

Shin, S. Y., *et al*. (1994). Structure activity relationships of human epidermal growth factor (hEGF). *Life Sciences*, **55**, 131–139.

Singh, J., Hamid, R. and Reddy, B. S. (1998). Dietary fish oil inhibits the expression of farnesyl protein transferase and colon tumor development in rodents. *Carcinogenesis*, **19**, 985–989.

Steinmetz, K. A. and Potter, J. D. (1991). Vegetables fruit and cancer I. Epidemiology. *Cancer Causes and Control*, **2**, 325–357.

Stoner, G. D., *et al*. (1997). Perspectives in cancer chemoprevention. *Environmental Health Perspectives*, **105** (Supplement 4), 945–954.

Van Lieshout, E. M. M., *et al*. (1996). Effects of oltipraz, α-tocopherol, β-carotene and phenyl isothiocyanate on rat oesophageal, gastric, colonic and hepatic glutathione, glutathione *S*-transferase and peroxidase. *Carcinogenesis*, **17**, 1439–1445.

Van Lieshout, E. M. M., *et al*. (1997). Effects of nonsteroidal antiinflammatory drugs on glutathione *S*-transferases of the rat digestive tract. *Carcinogensis*, **18**, 485–490.

Van Lieshout, E. M. M., *et al*. (1998). Effects of sulforaphane analog compound-30, indole-3-carbinol, D-limonene or relafen on glutathione *S*-transferases and glutathione peroxidase activity. *Biochimica Biophysica Acta*, **1379**, 325–336.

Wattenberg, L. W. (1985). Chemoprevention of cancer. *Cancer Research*, **45**, 1–8.

FURTHER READING

Brugge, J., *et al*. (eds) (1991). *Origins of Human Cancer. A Comprehensive Review*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Food, Nutrition and the Prevention of Cancer: A Global Perspective. (1997). (American Institute for Cancer Research/ World Cancer Research Fund, Washington, DC).

Fourth international conference on prevention of human cancer: nutrition and chemoprevention controversies, Tucson, Arizona (1992). *Preventive Medicine*, **22**, 629–811.

Verhoeven, D. T. H., *et al*. (1997). A review of mechanisms underlying anticarcinogenicity by brassica vegetables. *Chemico-Biological Interactions*, **103**, 79–129.

Chapter 32

Intervention and Chemoprevention of Cancer

Gary J. Kelloff
National Cancer Institute, National Institutes of Health, Rockville, MD, USA

Caroline C. Sigman
CCS Associates, Mountain View, CA, USA

CONTENTS

- Intervention and Chemoprevention of Cancer – Treatment of Carcinogenesis
- Molecular Targets and Chemoprevention Drug Discovery and Early Development
- Evaluating Chemopreventive Efficacy
- Chemopreventive Agent Development

INTERVENTION AND CHEMOPREVENTION OF CANCER – TREATMENT OF CARCINOGENESIS

Carcinogenesis is a several years to decades long process expressed in progressive genetic changes and corresponding increasingly severe tissue damage (**Figure 1**). Cancer prevention is intervention in this process before invasive disease develops, when it is potentially reversible, often asymptomatic, easier to control medically and associated with less morbidity. During the last half century, our understanding of carcinogenesis has grown enormously owing largely to technology allowing exploration of molecular pathways, cancer-associated genes and tissue architecture. This knowledge provides the basis for most cancer-preventive intervention strategies and, particularly for one of these strategies, chemoprevention – the use of drugs, biologicals and nutrients to treat precancer (i.e. to inhibit, delay or reverse carcinogenesis) (Sporn, 1976; Wattenberg, 1985; Kelloff *et al*., 1995a, 1996a; Hong and Sporn, 1997; Lippman *et al*., 1998; AACR Chemoprevention Working Group, 1999; Kelloff, 2000; Sporn and Suh, 2000).

Just as for cancer treatment drugs, the discovery and development of chemopreventive agents is mechanistically driven, focusing on modulating molecular targets associated with carcinogenesis (also called carcinogenesis biomarkers). This approach holds even more promise for chemoprevention than for treatment since molecular targets are more available and normal cell function and structure are less disrupted and, hence, easier to control in precancer than in cancer. Clinical cancer is in fact characterized by unregulated proliferation, cellular heterogeneity and, consequently very few sites and processes responsive to therapeutic intervention.

This chapter is an overview of strategies for chemoprevention. First, the principles of molecular target-based chemopreventive drug discovery and early development are described. Then a sequential drug development programme is described, which is translational in nature and builds on the rational mechanism-based and empirical agent discovery approach described, and leads to Food and Drug Administration (FDA) approvals for chemopreventive uses. An important aspect of this programme is identification of precancers (particularly intraepithelial neoplasia (IEN)) and early biomarkers of carcinogenesis, which can serve as surrogate endpoints for cancer incidence in chemoprevention studies. The material covered is drawn from and presented comprehensively in the references cited following the text.

MOLECULAR TARGETS AND CHEMOPREVENTION DRUG DISCOVERY AND EARLY DEVELOPMENT

Mechanisms of Chemopreventive Activity and Potential Chemopreventive Drugs

Basic research in carcinogenesis has identified enzymes, genetic lesions and other cellular components associated with the initiation and progression of precancers to invasive disease. Possible mechanisms for chemoprevention involve interfering with the expression and/or activity of these molecules; examples of the mechanisms, their possible molecular targets and agents that act at these targets

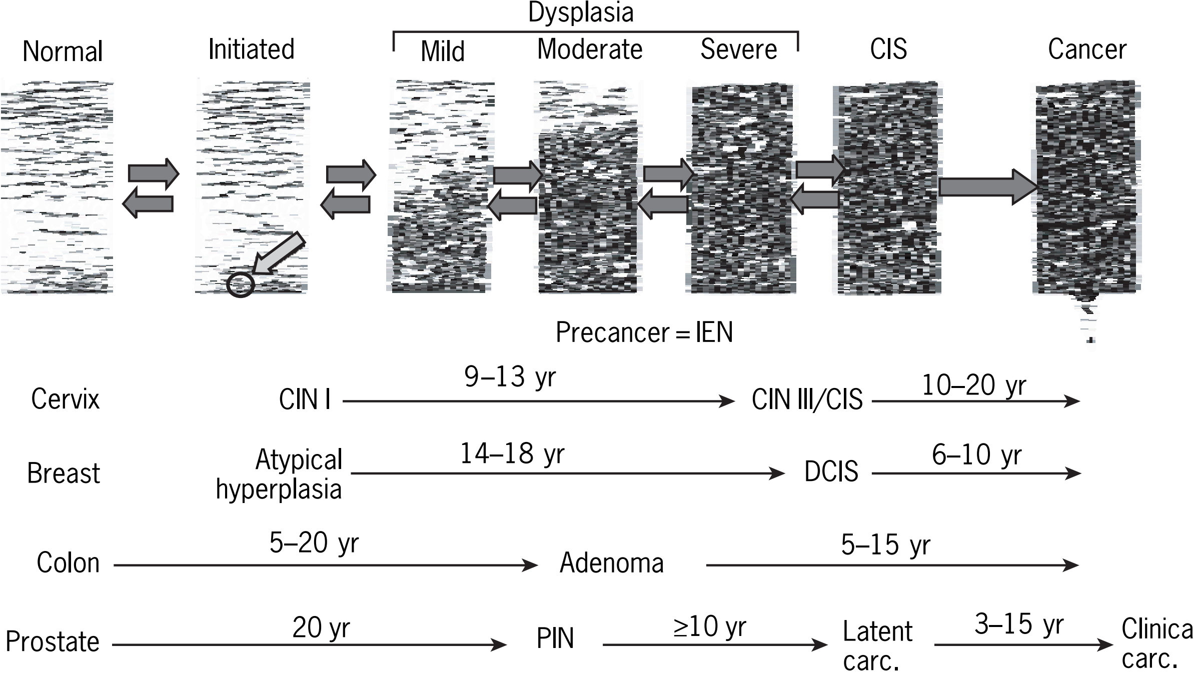


Figure 1 Carcinogenesis is a multiyear and progressive process. Abbreviations: CIN, cervical intraepithelial neoplasia; CIS, carcinoma *in situ*; DCIS, (breast) ductal carcinoma *in situ*; IEN, intraepithelial neoplasia; PIN, prostatic intraepithelial neoplasia. See also Kelloff, 2000, *Advances in Cancer Research*, **278**, 199–334 and Kelloff, *et al.*, 1996, *Oncology*, **10**, 1471–1484.

are listed in **Table 1** (Wattenberg, 1978; De Flora and Ramel, 1998; Hartman and Shankel, 1990; Kelloff *et al.* 1995b; Hong and Sporn, 1997; Singh and Lippman, 1998a,b; Kelloff, 2000).

Genetic progression models have been developed for many cancer sites, including colon, head and neck, bladder, brain, lung and cervix. (See the section on *In Vitro and Animal Models for Human Cancer*.) These models map the appearance of the specific molecular lesions and/or more general genotypic damage in their temporal association with increasingly severe precancer. In many cases early critical steps include inactivation of tumor suppressor genes, such as *APC* or *p53*, activation of oncogenes such as *RAS*, and damage to DNA repair mechanisms, such as by mutations in *MSH* and *MLH* (mismatch repair genes) and in *BRCA* (so named because it is mutated in some women at high risk of breast cancer). (See the section on *The Molecular Basis of Cell and Tissue Organisation*.) These models confirm that changes in specific biomarkers of carcinogenesis may be useful measures of potential chemopreventive activity.

Experimental and epidemiological carcinogenesis studies have associated >90% of cancers with exposure to mutagens and mitogens (agents which enhance cell proliferation). Mutagenesis can damage the cell and disrupt normal growth controls, resulting in loss of programmed cell death (apoptosis) and maturation pathways and increased (hyper)proliferation. Thus, inhibitors of mutagenesis and proliferation and inducers of apoptosis and differentiation are potential chemopreventives. There are also cell and tissue-based processes that are related to carcinogenesis, namely inflammation and oxidation. Hence, anti-inflammatories and antioxidants are possible chemopreventives. In this regard, experimental, epidemiological and clinical studies have demonstrated chemopreventive activity for agents that interfere with either the inflammatory or oxidating activities associated with arachidonic acid (AA) metabolism, for example nonsteroidal anti-inflammatory drugs (NSAIDs), polyphenols (tea, resveratrol) and vitamin A derivatives (retinoids).

Carcinogen-blocking Activities (Antimutagenicity)

Inhibition of carcinogen uptake into cells, inhibition of carcinogen formation or activation, carcinogen deactivation or detoxification, preventing carcinogen binding to DNA, and enhancing the level or fidelity of DNA repair are all carcinogen-blocking activities and potential chemopreventive mechanisms (Wattenberg, 1978; Kelloff *et al.*, 1995b). (See the chapter on *Antigenotoxins and Cancer*.)

Table 1 Mechanisms for chemoprevention with possible molecular targets[a]

Mechanism	Possible molecular targets	Representative agents
Inhibit carcinogen uptake	Bile acids (bind)	Calcium
Inhibit formation/activation of carcinogen	Cytochromes P450 (inhibit)	PEITC, tea, indole-3-carbinol
	PG synthase hydroperoxidase, 5-lipoxygenase (inhibit)	NSAIDs, COX-2 inhibitors, LOX inhibitors, iNOS inhibitors
	Bile acids (inhibit)	Ursodiol
Deactivate/detoxify carcinogen	GSH/GST (enhance)	Oltipraz, NAC
Prevent carcinogen-DNA binding	Cytochromes p450 (inhibit)	Tea
Increase level or fidelity of DNA repair	Poly(ADP–ribosyl)transferase (enhance)	NAC, protease inhibitors
Modulate hormone/growth factor activity	Oestrogen receptor (antagonize)	SERMs, soy isoflavones
	Androgen receptor (antagonize)	Bicalutamide, flutamide
	Steroid aromatase (inhibit)	Exemestane, vorozole, letrozole
	Steroid 5-reductase (inhibit)	Finasteride
	IGF-I (inhibit)	SERMs, retinoids
Inhibit oncogene activity	Farnesyl protein transferase (inhibit)	Perillyl alcohol, limonene, DHEA, FTI-276
Inhibit polyamine metabolism	ODC activity (inhibit)	DFMO
	ODC induction (inhibit)	Retinoids
Induce terminal differentiation	TGF (induce)	Retinoids, vitamin D, SERMs
	PPAR (activate)	GW7845
Restore immune response	COX (inhibit)	NSAIDs
	T, NK lymphocytes (enhance)	Selenium, tea
	Langerhans cells (enhance)	Vitamin E
Increase intercellular communication	Connexin 43 (enhance)	Carotenoids, retinoids
Restore tumour-suppressor function	p53 (stabilize)	CP-31398
Induce apoptosis	TGF (induce)	Retinoids, SERMs, vitamin D
	RAS farnesylation (inhibit)	Perillyl alcohol, limonene, DHEA, FTI-276
	Telomerase (inhibit)	Retinoic acid
	AA (enhance)	NSAIDs, COX-2 inhibitors, LOX inhibitors
	Caspase (activate)	Retinoids
	PPAR (activate)	Phenylacetate
	PPAR (inhibit)	NSAIDs
Inhibit angiogenesis	FGF receptor (inhibit)	Soy isoflavones, COX-2 inhibitors
	Thrombomodulin (inhibit)	Retinoids
Correct DNA methylation Imbalances	CpG island methylation (enhance)	Folic acid
	GSH/GST (enhance)	Oltipraz, NAC
Inhibit basement membrane degradation	Type IV collagenase (inhibit)	Protease inhibitors
Inhibit DNA synthesis	Glucose 6-phosphate dehydrogenase (inhibit)	DHEA, fluasterone

[a]Abbreviations: AA, arachidonic acid; COX, cyclooxygenase; CpG, cytosine–guanosine; DFMO, 2-difluoromethylornithine; DHEA, dehydroepiandrosterone; GSH, glutathione; GST, glutathione-S-transferase; FGF, fibroblast growth factor; IGF, insulin-like growth factor; iNOS, inducible nitric oxide synthase; LOX, lipoxygenase; NAC, *N*-acetyl-L-cysteine; NK, natural killer; NSAID, nonsteroidal anti-inflammatory drug; ODC, ornithine decarboxylase; PEITC, phenylethyl-isothiocyanate; PG, prostaglandin; PPAR, peroxisome proliferator-activated receptor; SERM, selective oestrogen receptor modulator; TGF, transforming growth factor.
See also Kelloff, G. J. (2000). *Advances in Cancer Research*, **278**, 199–334.

Inhibition of Carcinogen Uptake

Agents which inhibit carcinogen uptake appear to react directly with putative carcinogens, both initiators and promoters. For example, calcium inhibits the promotion of carcinogen- and dietary fat-induced colon tumours in rats. It also inhibits carcinogen-induced hyperproliferation induced in rat and mouse colon, including that induced by a Western 'stress' diet, i.e. a diet low in calcium and vitamin D and high in fat and phosphate. A partial explanation of these effects is that calcium binds to excess bile and free fatty acids that irritate the colon lumen and promote the formation of tumours. Such data suggest a potential for other sequestering and chelating agents as chemopreventives, particularly in the colon.

Inhibition of Carcinogen Formation/Activation

Vitamin C prevents the biosynthesis of carcinogenic *N*-nitroso compounds. Other chemopreventive antioxidants such as vitamin E prevent the formation of nitrosamines

from their precursors by scavenging nitrite. Many putative chemopreventive agents interfere with metabolic activation of a procarcinogen (Wattenberg, 1978; De Flora and Ramel, 1998; Hartman and Shankel, 1990; Kelloff *et al.*, 1995b). Examples are allylic sulfides, arylalkyl isothiocyanates, carbamates and flavonoids and other polyphenols. Usually this activity involves inhibition of the cytochrome P-450 enzymes responsible for activating various classes of carcinogens such as polycyclic aromatic hydrocarbons (PAHs).

Steroid aromatase, a cytochrome P-450-dependent enzyme, catalyses the first step in oestrogen biosynthesis in humans: C19 hydroxylation and subsequent oxidative cleavage of the androgens androstenedione and testosterone to oestrone and oestradiol, respectively. Both steroidal (e.g. 4-hydroxyandrostenedione) and nonsteroidal (e.g. vorozole) aromatase inhibitors also inhibit carcinogenesis in oestrogen-sensitive tissues.

Enhancement of Carcinogen Deactivation/Detoxification

Enhancement of carcinogen deactivation/detoxification is potentially a very important strategy for chemoprevention (De Flora and Ramel, 1998; Kelloff *et al.*, 1995b). Two metabolic pathways are critical. The first is the introduction or exposure of polar groups (e.g. hydroxyl groups) on procarcinogens/carcinogens via the phase I metabolic enzymes, which are primarily the microsomal mixed-function oxidases. These polar groups become substrates for conjugation. The second pathway is via the phase II metabolic enzymes responsible for conjugation and the formation of glucuronides, glutathione (GSH) conjugates and sulfates. In both cases, the conjugates are more likely to be excreted from the body than they are to reach sensitive tissues in activated form. Agents that affect phase II enzymes probably hold more promise than those which induce phase I enzymes, since phase I oxidation also can result in carcinogen activation. (See the chapter on *Mechanisms of Chemical Carcinogenesis*.)

GSH is a prototype carcinogen scavenger (see also under the more general mechanism of electrophile scavenging below). It reacts spontaneously or via catalysis of GSH-S-transferases with numerous activated carcinogens including some *N*-nitroso compounds, aflatoxin B_1 (AFB_1) and PAHs. GSH protects against mouse skin tumours induced by PAHs, rat forestomach tumours induced by nitrosamines and rat liver tumours induced by AFB_1. A number of promising chemopreventive agents are potent inducers of GSH and GSH-S-transferases, including allylic sulfides, which are natural products found in onion, garlic and other members of the *Allium* genus as well as the sulfur-containing compounds found in cruciferous vegetables. Oltipraz (a dithiolthione similar to those found in cruciferous vegetables) is a potent GSH-S- transferase inducer with a wide spectrum of chemopreventive activity. Sulforaphane, an isothiocyanate found in broccoli sprouts, induces phase II enzymes and has chemopreventive activity in rat colon (prevents formation of nitrosamine-induced ACF) and mammary gland (PAH-induced carcinoma). *N*-Acetyl-L-cysteine (NAC) is essentially a precursor of GSH. NAC shows inhibitory activity in mouse lung and bladder and rat colon and mammary gland against nitrosamine-induced tumours.

GSH-peroxidases (GSH-Px) catalyse the reduction of hydrogen peroxide (H_2O_2) and organic hydroperoxides; the antioxidant effects of selenium may be related to its function in the enzyme's active site. Although several studies show that the anticarcinogenic activity of selenium in mouse and rat mammary glands is not mediated by GSH-Px, in tissues such as colon, glandular stomach and skin, GSH-Px are thought to play a role.

Another type of carcinogen deactivation is modulation of the mixed-function oxidases involved in the metabolism of oestrogens. Indole-3-carbinol, a compound which occurs naturally in cruciferous vegetables, inhibits the induction of mammary tumours in rats and induces mixed-function oxidases. Particularly, it induces the activity of the enzymes responsible for 2-hydroxylation of oestradiol, leading to increased excretion of oestradiol metabolites.

Inhibition of DNA–Carcinogen Adduct Formation

DNA–carcinogen adduct formation can be considered a biomarker of carcinogen exposure. In most cases, it is probably secondary to other mechanisms of carcinogenesis, such as carcinogen activation and formation. Likewise, inhibition of DNA adduct formation is typically an indirect measure of other mechanisms of chemoprevention, particularly inhibition of carcinogen formation and activation and enhancement of carcinogen detoxification (Hartman and Shankel, 1990; Kelloff *et al.*, 1995b). For example, oltipraz prevents the formation of AFB_1–DNA adducts, an effect which has been attributed to increased rates of aflatoxin detoxification by GST. Nonetheless, inhibition of DNA adduct formation is a convenient assay for screening potential chemopreventive agents which are expected to modulate carcinogen metabolism. There is also limited evidence of chemopreventive agents directly obstructing adduct formation. For example, ellagic acid appears to inhibit the carcinogen adduct formation by itself binding to the duplex form of DNA. (See also chapter on *The Formation of DNA Adducts*.)

Enhancement of DNA Repair

There are three possible chemopreventive mechanisms that involve DNA repair (Kelloff *et al.*, 1995b). First is an increase in the overall level of DNA repair. Second, the enzyme poly(ADP-ribosyl)transferase (ADPRT) is known to be involved in modulation of DNA damage, and the level of this enzyme is decreased in the presence of

carcinogens. The chemopreventive agent NAC prevents carcinogen-induced decreases in ADPRT. The third mechanism is suppression of error-prone DNA repair. It is known that protease inhibitors depress error-prone repair in bacteria, and it has been suggested that they might prevent carcinogenesis by inhibiting an error-prone repair system activated by proteases that, in turn, are induced by tumour promoters. The protease inhibitor best studied as a chemopreventive is Bowman–Birk soybean trypsin inhibitor (BBI), which inhibits nitrosamine-induced tumours in mouse colon and liver and in rat oesophagus.

Antiproliferative Activity via Signal Transduction Pathways

Much is now known about the biochemical control mechanisms involved in regulating cell growth and development. Cells respond to signals from extracellular stimuli via a complicated network of highly regulated events collectively referred to as signal transduction pathways. Stimulation of these pathways results in changes in transcriptional activity. While normal cells respond appropriately to extracellular stimuli, many precancerous and cancerous cells have lost this ability and display aberrant signalling.

Molecular targets on these pathways that allow interference with the deregulated signalling are potential sites for chemopreventive intervention (Powis, 1994; Powis and Workman, 1994). For example, key components of these pathways are growth factors such as epidermal growth factor (EGF), insulin-like growth factor (IGF) and transforming growth factor (TGF) and protein tyrosine kinases which catalyse the transfer of the phosphate of ATP to the hydroxyl group of tyrosine on numerous proteins. Many growth factor receptors (e.g. EGFR) and oncogenes are tyrosine kinases, and loss of tyrosine kinase regulatory mechanisms has been implicated in neoplastic growth.

Invocation of signal transduction pathways also provides a mechanistic rationale for the multiple chemopreventive effects of some classes of agents. For example, chemopreventive agents such as retinoids, antihormones and protein kinase inhibitors which affect activities at the cell membrane level and cytoplasmic and nuclear receptor levels can also affect other connected events. It is evident that many of these activities are interrelated, e.g. effects on the proliferation associated enzyme ornithine decarboxylase (ODC), AA metabolism, protein kinase C (PKC), IGF-I and TGF may be pleiotropic results of activity at single locus on signal transduction pathways. It is also clear that a single activity may not be the most important or the only one required for carcinogenesis. (See section on *The Molecular Basis of Cell and Tissue Organisation*.)

Modulate Hormonal/Growth Factor Activity

Chemicals may inhibit the proliferation associated with carcinogenesis by directly regulating the induction and activity of specific hormones and growth factors that initiate steps in signal transduction (Kelloff *et al.*, 1995b). This regulation may occur at membrane level receptors (for growth factors, peptide hormones and neurotransmitters) or via cytoplasmic and nuclear receptors (for the steroid superfamily consisting of oestrogen, progesterone, retinoid, glucocorticoid, vitamin D and thyroid receptors). For example, antioestrogens such as tamoxifen bind to nuclear oestrogen receptors, preventing the binding and activity of oestrogens. Tamoxifen inhibits carcinogen-induced, oestrogen-sensitive tumours in rat mammary glands and hamster kidney. Most importantly, tamoxifen has been shown to lower the risk of breast cancer in women at high risk. Phyto-oestrogens, such as the isoflavone genistein, have anti-oestrogenic activity. Studies in human breast cancer cells indicate that the anti-oestrogenic effect may result from slowed translocation of genistein-bound receptor from the cytoplasm to the nucleus compared with that of oestradiol-bound receptor.

TGFβ has antiproliferative activity in both normal and cancer cells. These observations suggest that chemicals that activate TGFβ could also control proliferation in carcinogenesis. In this regard, breast cancer cells normally produce only small amounts of activated TGFβ, but treatment with tamoxifen increases production up to 20-fold. Retinoic acid, which inhibits chemical carcinogenesis, particularly tumour promotion in mouse skin, induces TGFβ_2 in mouse skin after topical application. In vitamin A-deficient rats treated with retinoic acid, the level of expression of TGFβ correlates with levels of retinoids in skin, intestine and respiratory tract tissue.

There is also evidence of cross-regulation among membrane and nuclear receptors. For example, IGF-I stimulates cell replication in various tumours. Particularly, human breast cancer cells have membrane receptors for and excrete IGF-I. Tamoxifen lowers blood concentrations of IGF-I in breast cancer patients, suggesting that part of its antitumour activity is inhibition of IGF-I.

Other aspects of receptor activity are possible mechanisms for chemopreventive activity. Generally, receptors are phosphoproteins, and phosphorylation appears to play a role in receptor activation. Thus, chemopreventive agents which inhibit phosphorylation, e.g. inhibit protein kinases, may influence cell proliferation by effects on receptors. An example is the isoflavone genistein, which is a specific inhibitor of tyrosine kinase and other flavonoids.

Deactivation of steroids may prevent hormone-stimulated carcinogenesis. In this regard, aromatase inhibitors and modifiers of steroid hydroxylation have been described above under inhibition of carcinogen formation/activation and carcinogen deactivation/detoxification, respectively.

Inhibit Oncogene Activity

During the course of cell proliferation in carcinogenesis, numerous oncogenes are expressed abnormally, possibly

as intermediates in the signal transduction pathways. The evidence for oncogene activity in signal transduction pathways is based on the similarity of some of their products (protein kinases) to other intermediates in these pathways (Kelloff *et al.*, 1995b). There are several points during activation at which the *ras* oncogene can be inhibited, and there are data relating this inhibition to chemopreventive activity.

First, a membrane receptor-linked tyrosine kinase is involved in Ras activation, and kinase inhibitors would be expected to prevent Ras activation. Particularly interesting are compounds such as genistein which specifically inhibit tyrosine kinase, and thus do not interfere with normal cellular processes mediated by other kinases.

To be activated, the Ras protein must first be farnesylated. *Ras* oncogenes are involved in rat mammary gland carcinogenesis. D-Limonene inhibits the progression of carcinogen-induced mammary tumours induced in rats, and it also inhibits the farnesylation of small G proteins (21–26 kDa); these experimental data suggest that D-limonene could be preventing oncogene activation by inhibiting post-translational farnesylation of the p21 Ras protein. Perillyl alcohol is an even more potent inhibitor of farnesyl–protein transferase. Recently, several specific farnesylation inhibitors have been described which are structural analogues of the C-terminal tetrapeptide of farnesyl–protein transferase and inhibit the growth of *ras*-dependent tumours.

Further, farnesyl pyrophosphate, the substrate for farnesyl–protein transferase, is an intermediate in the synthetic pathway from hydroxymethylglutaryl coenzyme A (HMG CoA) reductase to cholesterol. Inhibitors of HMG CoA reductase, e.g. lovastatin, and probably inhibitors of other enzymes along the synthetic route to cholesterol, have been shown to inhibit Ras farnesylation.

Cyclooxygenase (COX) catalyses the synthesis of prostaglandins (PGs) from AA. COX inhibitors also might inhibit proliferation in carcinogenesis by inhibition of oncogene expression, although the evidence is less direct than for other effects of AA metabolism inhibitors. Expression of the oncogene c-*myc* occurs early in EGF-induced cell proliferation. PGs are required but not sufficient for c-*myc* expression and DNA synthesis stimulated by EGF. The NSAID indomethacin inhibits both EGF-induced DNA synthesis and oncogene expression; this inhibition is reversed by addition of PGG_2.

Studies *in vitro* indicate inhibition of oncogene expression as a mechanism for chemopreventive activity of protease inhibitors and retinoids. For example, the protease inhibitors inhibit transformation of cells transfected with activated H-*ras* oncogene, suppress c-*myc* expression in mouse fibroblasts and inhibit carcinogen-induced tumours in rat colon, mouse lung and mouse skin. Retinoic acid also inhibits H-*ras*-induced transformation in cancer cells and mouse skin carcinogenesis.

Inhibit Polyamine Metabolism

Polyamines play a significant role in cell proliferation, differentiation and malignant transformation. The mode of action is not yet known, but it has been suggested that polyamines stabilize DNA structures; they have been shown to affect DNA and protein synthesis. A critical step in polyamine biosynthesis is the synthesis of putrescine from ornithine that is catalysed by ODC. There is ample evidence that ODC participates in carcinogenesis – the enzyme is induced during cell transformation by chemical carcinogens, viruses and oncogenes.

Association with cell proliferation during carcinogenesis is also well established. TPA and other tumour promoters increase ODC activity in skin, colon, bladder and liver. In mouse skin, topically applied TPA causes an approximately 200-fold increase in ODC activity within 4.5 h after treatment. The increase is dose dependent and correlates with the ability of the TPA dose to promote skin tumours. Also, the increased ODC activity has been proposed to be specific to tumour promotion, since most carcinogens that are not tumour promoters do not induce ODC.

Likewise, chemicals that inhibit induction of or deactivate ODC also inhibit carcinogenesis. Some of the most convincing results demonstrating that inhibition of ODC prevents cancers come from studies with DFMO. DFMO is a specific, mechanism-based irreversible inhibitor of ODC – that is, DFMO is activated by ODC into a form that reacts with the enzyme to inactivate it. DFMO inhibits carcinogen-induced tumours in mouse and rat colon and bladder, rat mammary glands and mouse skin.

ODC induction by TPA is regulated at the transcription level. Regulation occurs in part via signal transduction events at the membrane. For example, PKC appears to be involved, as are diverse signal transduction intermediates induced by TPA, including PGs, other products of AA metabolism and free radicals. Chemicals that inhibit PKC and AA metabolism and those that scavenge free radicals also may inhibit the induction of ODC, hence they may be chemopreventives by this mechanism. In this regard, several of the PKC inhibitors, including glycyrrhetinic acid, inhibit ODC induction and tumour promotion in mouse skin. AA metabolism inhibitors also inhibit both ODC induction and TPA-promoted mouse skin tumorigenesis, as do free radical scavengers such as GSH, flavonoids and green tea polyphenols.

Vitamin A (retinol) and its derivatives (i.e. retinoids) inhibit carcinogenesis specifically during promotion. There is evidence that the cancer inhibitory activity of these compounds may be mediated partially by regulation of ODC induction. One of the most active retinoids is fenretinide. This compound is a potent inhibitor of ODC induction as well as TPA promotion in mouse skin. It also

inhibits carcinogen-induced mammary gland tumours in rats and bladder tumours in mice.

Inhibition of *S*-adenosyl-L-methionine (SAM) decarboxylase is another mechanism for inhibiting polyamine biosynthesis that may prove useful in chemoprevention. This enzyme, like ODC, is highly regulated in mammalian cells and catalyses the formation of the polyamines spermidine and spermine from putrescine.

Induce Terminal Differentiation

Terminal differentiation is one of the steps in normal, regulated cell proliferation in epithelial tissues. Proliferating cancer cells often have lost the ability to differentiate. These cancer cells are either deficient in or incapable of responding to differentiation signals. Abundant evidence demonstrates that restoring the ability of abnormally proliferating cells to differentiate suppresses carcinogenesis. Several classes of chemopreventives also induce differentiation. Retinoids are the best-studied example (Singh and Lippman, 1998a). For many years it has been known that vitamin A deficiency causes squamous metaplasia and hyperkeratinization – both are signs of excessive tissue. Studies in hamster trachea and various cancer cells show that the differentiated phenotype can be restored by treatment with retinoids. Evidence indicates that retinoids control differentiation via intracellular binding proteins (cellular retinol-binding protein and cellular retinoic acid-binding protein) and nuclear receptors.

Calcium and vitamin D_3 are well-known differentiating agents that also inhibit carcinogenesis. Calcium induces differentiation in epithelial tissues including rat oesophagus, mouse skin and human mammary gland and colon. Vitamin D_3 induces differentiation in human colon, human and mouse myeloid leukaemia cells, mouse skin cells, mouse melanoma cells and other cells. It has been suggested that the effects of the two chemicals on differentiation may be mediated by the same signal transduction pathway, involving the vitamin D_3 nuclear receptor with calcium as the messenger.

Restore Immune Response

Antibodies to oncogene products are important in the inhibition of cell transformation and tumour growth. PGE_2 is known to suppress immune response in certain tumour cells. COX inhibitors diminish the immune suppression, and it has been suggested that this effect on immune suppression may be part of the mechanism by which COX inhibitors reduce tumour growth, as seen in several animal tumour models including colon and Lewis lung carcinoma.

Retinoids also are known to be immunostimulants. Retinoic acid increases cell-mediated and natural killer (NK) cell cytotoxicity; retinoids also cause some leukaemia cells to differentiate to mature granulocytes comparable to mature neutrophils. These effects might be partially responsible for the activity of retinoids against established tumours.

Pharmacological doses of vitamin E fed with normal, well-balanced animal diets increase humoural antibody production, especially IgG; this effect has been observed repeatedly in chickens, mice, turkeys, guinea pigs and rabbits. Vitamin E also stimulates cell-mediated immunity, as evidenced by enhanced mitogenesis and mixed lymphocyte response in spleen cells from mice fed the vitamin. In particular, vitamin E prevents the carcinogen-induced decrease in the density of macrophage-equivalent cells (Langerhans cells) in the buccal pouch of carcinogen-treated hamsters. Likewise, vitamin E inhibits the induction and causes regression of tumours in hamster buccal pouch.

The role of selenium in mediating immune responses suggests that the broad spectrum activity of selenium in inhibiting chemical carcinogenesis may be attributed partially to stimulation of the immune system. In general, selenium deficiency causes immunosuppression, while supplementation with low doses of selenium restores and increases immune response. Perhaps most important in inhibiting tumorigenesis is the effect of selenium on the cytotoxicity of immune system cells. Compared with normal cells, both T and NK lymphocytes from selenium-deficient mice have decreased ability to destroy tumour cells *in vitro*. Supplementation with selenium enhances the ability of rat NK cells to kill tumour cells. The role of immunostimulation in carcinogenesis inhibition by selenium has been studied to only a limited extent and has not been confirmed. However, the potent inhibitory activity of selenium compounds against DMBA-induced tumors in rat mammary glands is suggestive, since the immunosuppressive effects of DMBA are well documented.

Increase Intercellular Communication

Gap junctions are the cell components that coordinate intercellular communication. They are composed of pores, or channels, in the cellular membranes that join channels of adjacent cells; these pores are regulated and, when open, allow passage of molecules up to about 1000 D in size. Gap junctions may allow growth regulatory signals to move between cells. There is evidence from studies *in vitro* that inhibition of gap junctional intercellular communication occurs in the proliferative phase of carcinogenesis. In *in vitro* studies, enhancement of communication correlates to inhibition of cellular transformation.

To date, only limited data suggest the potential for inhibiting chemical carcinogenesis by the other antiproliferative/antiprogression mechanisms listed in **Table 1**, but the possibilities exist and warrant consideration here.

Restore Tumour-suppressor Function

Many so-called 'tumour-suppressor' genes have been found that may be involved in controlling proliferation and

differentiation in cells. Particularly, their function is associated with control of abnormal growth in carcinogenesis. Several of these genes have been identified and implicated in pathogenesis by the presence of mutated or otherwise dysfunctional forms in specific cancers. For example, the tumour suppressor *Rb* is involved in retinoblastoma, osteosarcoma and tumours in lung, bladder, prostate and breast; *p53* in adenocarcinomas in colon and breast, human T cell leukaemias, glioblastomas, sarcomas, and tumours in lung and liver; *WT* in Wilm tumour; and *DCC* (Deleted in Colorectal Cancer) in colon tumours. There is potential for treating cancer patients with exogenous functional tumour-suppressor genes to inhibit tumour growth and spread. Possibly, it also will be found that chemicals can modulate the expression and activity of tumour suppressors and inhibit carcinogenesis by this mechanism. CP-31398 stabilizes the DNA binding domain of both normal and mutant p53 in an active conformation, induces the $p21^{WAF1}$ cell cycle regulatory protein in the absence of normal p53 and inhibits growth of human tumours with p53 mutated tumours in a mouse model.

Induce Programmed Cell Death (Apoptosis)

Apoptosis is a well-regulated function of the normal cell cycle requiring gene transcription and translation. Tumour suppressors, such as p53 and certain regulatory growth factors, particularly TGFβ1, have been implicated as inducers of apoptosis. Programmed cell death has been described as the complement to mitosis in the maintenance, growth and involution of tissues; it is the process by which damaged and excessive cells are eliminated. Apoptosis is inhibited by tumour promoters such as TPA and phenobarbital and other chemicals that stimulate cell proliferation such as hormones. These data suggest that induction of apoptosis may inhibit tumour formation and that agents which inhibit tumour promotion may act by inducing or preventing inhibition of apoptosis through any one of several signal transduction pathways. For example, hamster pancreatic cancers regress when apoptosis is induced, and many potential chemopreventive agents (e.g. tamoxifen, NSAIDs, retinoids) induce programmed cell death in precancerous and cancer cells.

Correct DNA Methylation Imbalances

Changes in DNA methylation patterns appear to be involved in carcinogenesis. (See the chapter on *Non-Genotoxic Causes of Cancer*.) Methyl-deficient diets cause fatty livers, increased cell turnover and promote the development of carcinogen-induced liver tumours in rats and mice. Conversely, methyl-rich (fortified with choline and methionine) diets prevent or reduce these effects. Changes in gene expression, such as increased expression of oncogenes, appear in animals on methyl-deficient diets. These effects are similar to those seen in rodents given tumour-promoting chemicals and they are reversible on methyl replacement. Hypomethylation is also associated with hyperproliferation in colon tissue. Methionine, which is involved with choline, folic acid and vitamin B_{12} in regulating intracellular methyl metabolism inhibits carcinogen-induced mammary gland cancers in rats. Also, folic acid inhibits carcinogen-induced lung tumours in mice. Conversely, methylation of CpG islands in the promoter regions of tumour-suppressor and *GST* genes has been seen in cancers of several major target organs including colon, prostate, breast and lung. This methylation prevents gene expression and provides a rationale for the chemopreventive activity of agents which induce GST and tumour-suppressor activity.

Inhibit Angiogenesis

Angiogenesis is the process leading to the formation of new blood vessels. In normal tissue, it is a highly regulated process essential to reproduction, development and wound repair. In carcinogenesis, it is required in tumour growth and involved in metastasis, and there is evidence that angiogenesis also may occur early in carcinogenesis. There is indirect evidence that certain chemicals that inhibit carcinogenesis may inhibit angiogenesis. For example, PGs E_1 and E_2 are angiogenic. Therefore, agents that inhibit PG synthesis may inhibit carcinogenesis by inhibiting angiogenesis. Similarly various growth factors and, particularly, vascular endothelial growth factor (VEGF), increase angiogenesis by activating signal transduction pathways. Inhibition of angiogenesis may be a chemopreventive mechanism for agents which affect these pathways.

Inhibit Basement Membrane Degradation

Tumour cells produce various enzymes that destroy the basement membrane which acts as a barrier against malignant cancer cells and prevents cancer spread. These enzymes include, among others, the proteases collagenase, cathepsin B, plasminogen activators and prostate-specific antigen (PSA). Protease inhibitors are known to act against thrombin and type IV collagenase, which are among the proteases hypothesized to participate in the destruction of basement membranes during cancer invasion. Proteases are also involved in angiogenesis. Thus, protease inhibitors that slow carcinogenesis may derive their effects, in part, by inhibiting basement membrane degradation or by inhibiting angiogenesis.

Antioxidant/Anti-inflammatory Activity

Many classes of antioxidants and anti-inflammatories have shown chemopreventive activity in animal models (Wattenberg, 1978; Kelloff *et al.*, 1995b; Singh and Lippman, 1998a,b). Those with activity derived by inhibition of AA metabolism are among the most promising

chemopreventive agents. AA is metabolized to PGs, thromboxanes, leukotrienes and hydroxyeicosatetraenoic acids (HETEs) via oxidative enzymes. Activated oxygen species and alkylperoxy species are formed throughout this process; AA metabolism is increased during inflammation. Two aspects of AA metabolism are associated strongly with carcinogenicity

The first is the PG synthetic pathway, involving the enzyme PG H synthase (PHS). This enzyme has two activities – COX, which catalyses the formation of PGG_2 from arachidonic acid, and hydroperoxidase, which catalyses the reduction of PGG_2 to PGH_2. To return to its native state, the hydroperoxidase requires a reducing cosubstrate; procarcinogens, e.g. arylamino and arylnitro compounds, are such substrates. According to the model proposed, the carcinogens are activated (oxidized) during catalysis to free radicals and electrophiles that can form adducts with DNA and initiate carcinogenesis. This process can be stopped in four ways: (1) at formation of PGG_2 via inhibition of COX, (2) by inhibition of peroxidase activity, (3) by preventing formation of reactive intermediates and (4) by scavenging reactive intermediates (e.g. by GSH conjugation). Relevant to these potential mechanisms, COX inhibitors such as NSAIDs and certain antioxidants (e.g. flavonoids) are effective inhibitors of carcinogenesis. Additionally, PGH_2 itself breaks down to form a known direct-acting mutagen, malondialdehyde. Thus, inhibition of COX may directly prevent the formation of a potential carcinogen.

NSAIDs, which are COX inhibitors have demonstrated potent activity against colon cancer in epidemiological and animal studies and against bladder cancer in animal models. They have also shown chemopreventive activity in skin, oesophagus, lung and breast in animal studies. However, PGs and other PHS products such as thromboxanes have multiple activities, some of which are beneficial and tissue-specific. For example, PGE_2 in the gut promotes protective mucosal secretions; lowered gut PG levels resulting from NSAID administration are associated with one of the major side effects of long-term NSAID treatment, gastrointestinal ulceration and bleeding. Likewise, PGs in the kidney and thromboxanes in platelets are important to normal physiological function. Their inhibition is associated with renal tubule toxicity and excessive bleeding, respectively. The development of chemopreventive agents which retain the ability to inhibit the carcinogenesis-associated activities of PGs without depressing protective effects is an attractive strategy. The discovery of an inducible form of COX, COX-2, which is predominant at inflammation sites, in macrophages and in synoviocytes, suggested that such an approach is feasible. In contrast to COX-2, constitutive COX-1 predominates in the stomach, gastrointestinal tract, platelets and kidney. Traditional NSAIDs inhibit both forms of the enzymes, but other compounds inhibit COX-2 selectively – glucocorticoids such as dexamethasone (which, for example, has chemopreventive activity on topical application to mouse skin) and newer NSAIDs such as celecoxib and rofecoxib. Moreover, COX-2 regulates or participates in a number of different cellular pathways and biological processes important in carcinogenesis. Expression of COX-2 is increased in many types of human cancers (e.g. colon, oesophagus, lung, breast, pancreas, prostate) and precancers. COX-2 inhibitors have potent chemopreventive activity in animal models of colon, bladder and skin carcinogenesis. In a recent clinical study the COX-2 inhibitor celecoxib significantly reduced the number of colorectal adenomas (which are precancers) in patients with familial adenomatous polyposis (FAP), supporting the concept that COX-2 inhibitors may have therapeutic utility against colorectal cancer and gaining FDA approval for celecoxib use in treatment of FAP (see the chapter *Chemoprevention*).

The second aspect of AA metabolism associated with carcinogenesis is the burst of PHS and lipoxygenase activity that is seen during inflammation and is stimulated by the tumour promoter TPA. The available evidence suggests that the immediate products of lipoxygenase activity, the HETEs and their hydroperoxy precursors (HPETEs), are more important to tumour promotion than are PGs. Compounds that inhibit lipoxygenase, such as aesculetin, inhibit carcinogenesis in animal models.

Evidence of Chemopreventive Potential Derived from Studies at Molecular Targets

In evaluating the potential efficacy of chemopreventive agents, several molecular target parameters are weighed: (1) number of chemoprevention-related pharmacological activities, (2) impact of the agent on likely carcinogenesis pathways to the targeted cancer, (3) pharmacodynamics and (4) specificity for chemopreventive activity compared with effects on normal cellular function. These data are important throughout the development process for chemopreventive agents, and they are particularly important in the earlier phases of identifying promising candidate agents and characterizing efficacy. *In vitro* molecular target assays are a first step in evaluating chemopreventive potential.

However, given today's state of knowledge, activity at molecular targets alone may be inconclusive. Carcinogenesis can take multiple paths and be multifocal (not all cancers in a given tissue or all cells in a given cancer may ultimately contain the same molecular lesions); only a few early molecular lesions will progress. Progression can be influenced by factors specific to the host tissue's environment, such as the action of hormones. Further, although the progression models suggest that there is a rough order in which the various types of changes appear; until the time comes (which may be relatively soon) when the progression models are worked out in fine detail supported by functional genomic/proteomic analyses, the accumulation of multiple effects as evidenced in phenotypic changes at the cellular and tissue level can be more reliable measures of

carcinogenesis and their inhibition or reversal, of chemopreventive potential. Hence, the early development of chemopreventive drugs also uses short-term *in vitro* assays that measure changes in cell proliferation and cell proliferation kinetics, other parameters of malignant transformation and even animal models of carcinogenesis in which tissue characteristics associated with precancer are measured. Nonetheless, effects on molecular targets are also useful in defining the most appropriate animal efficacy models and in interpreting the results of assays in these models.

EVALUATING CHEMOPREVENTIVE EFFICACY

In developing chemopreventive agents a sequential approach is used to evaluate potential efficacy – starting with *in vitro* and cell-based mechanistic assays and efficacy screening tests, then screens *in vivo* in animal carcinogenesis models with cancers and precancerous lesions as endpoints, and finally the most promising agents are characterized more fully in animal carcinogenesis models (e.g. dose–response and dosing regimens are evaluated, combinations with other agents are tested).

Mechanistic Assays

Agents not previously tested are put first into mechanistic assays to determine their potential range of chemopreventive activities (**Table 2**). Many of the mechanistic endpoints are described above in relation to characterizing various classes of chemopreventive agents. The battery of assays used is continually evolving and is designed to address various specific activities associated with general

Table 2 Representative assays of chemopreventive mechanisms[a]

Assays	Cell substrate	Mechanism measured
Antimutagenesis		
B(*a*)P-DNA adduct formation (inhibition)	Human bronchial epithelial cells (BEAS2-B)	DNA damage inhibition
NAD(P)H:quinone reductase (induction)	Human (Chang) liver cells	Carcinogen detoxification
GSH *S*-transferase (induction)	Human (Chang) liver cells	Carcinogen detoxification
GSH synthesis and GSSG reduction (induction)	Buffalo rat liver (BRL-3A) cells	Carcinogen detoxification
Antiproliferation		
TPA-induced ODC (inhibition)	Rat tracheal epithelial cells (2C5 cell line)	Antiproliferative activity
Normal epithelial cell proliferation (inhibition)	Primary human keratinocytes	Antiproliferative activity
Poly(ADP–ribose)polymerase (inhibition)	Primary human fibroblasts	Error-prone DNA repair inhibition (DNA damage inhibition)
Calmodulin-regulated phosphodiesterase (inhibition)	Human leukaemia (HL60) cells	Signal transduction regulation
TPA-induced tyrosine kinase (inhibition)	Human leukaemia (HL60) cells	Signal transduction regulation
EGFR (inhibition)	Human A431 and mouse 3T3 cells	Signal transduction regulation
ras farnesylation (inhibition)	Rat brain farnesyl transferase	Signal transduction regulation
HMG–CoA reductase (inhibition)	Rat liver HMG–CoA reductase	Signal transduction regulation
Steroid aromatase (inhibition)	PMSG-stimulated rat ovary aromatase	Antioestrogenic activity
Oestrogen receptor (antagonism of binding and expression)	MCF-7 cells	Antioestrogenic activity
5-reductase (inhibition)	Rat prostate 5-reductase	Antiandrogenic activity
Cellular differentiation characteristics (modulation)	Human leukaemia (HL60) cells	Differentiation
DNA fragmentation (induction)	Human leukaemia (HL60) or U937 cells	Apoptosis
Antioxidant/anti inflammatory activity		
AA metabolism: micronuclei in keratinocytes (inhibition)	P388 macrophages/human keratinocytes	Anti-inflammatory activity
TPA-induced active oxygen (inhibition)	Human leukaemia (HL60) cells	Free radical scavenging
COX-2 (inhibition)	Sheep placenta COX-2	Anti-inflammatory activity
LOX (inhibition)	Rat RBL-1 cell LOX (for 5-LOX)	Anti-inflammatory activity

[a]Abbreviations: AA, arachidonic acid; B(*a*)P, benzo[*a*]pyrene; COX, cyclooxygenase; EGFR, epidermal growth factor receptor; GSH, glutathione; GSSG, glutathione disulfide; GST, glutathione-*S*-transferase; HMG-CoA, hydroxymethylglutaryl coenzyme A; LOX, lipoxygenase; ODC, ornithine decarboxylase; TPA, tetradecanoyl phorbol acetate.

See also Kelloff, 2000, *Advances in Cancer Research*, **278**, 199–334.

Table 3 *In Vitro* Assays of Chemopreventive Efficacy[a]

Assay	Cell system	Carcinogen
Morphological transformation (inhibit)	Rat tracheal epithelial cells	Benzo[*a*]pyrene
Hyperplastic nodules (inhibit)	Mouse mammary gland organ culture	7,12-dimethylbenz[*a*]anthracene tetradecanoylphorbol acetate
Anchorage independence	Human lung tumour (A427) cells	None

[a]See also Steele *et al.*, 1996, *Journal of Cellular Biochemistry*, Suppl. 26, 29–53.

categories of chemopreventive activity – carcinogen blocking (antimutagenicity), antiproliferation and antioxidation/antiinflammation. As the molecular bases of carcinogenesis become better known, additional mechanistic activities appropriate to chemoprevention are identified. Examples of those in the early stages of investigation are inhibition of cell cyclins, telomerase and angiogenesis, as well as binding to peroxisome proliferator-activated receptors (PPARs) (Kelloff, 2000).

Cell-based Assays

Selected cell-based assays have been used routinely to screen the efficacy of potential chemopreventive agents (**Table 3**) (Steele *et al.*, 1996). Initial criteria for selecting the *in vitro* tests include (1) efficiency in terms of time and cost, (2) sensitivity and ease of quantitation, (3) controlled test conditions, (4) relevance to organ systems of interest, (5) use of epithelial cells and (6) if possible, use of human cells. In each assay, the agents are tested over a wide range of concentrations, and IC_{50}s are determined. New cell and organ culture technology is being used to design assays of chemopreventive efficacy, e.g. raft cultures (allowing evaluation of stromal–epithelial interactions), cells from transgenic mice and human cells carrying known cancer-predisposing genes (e.g. *p53* mutations characterizing Li–Fraumeni syndrome).

Inhibition of Carcinogenesis in Animal Models

Prevention of Cancers in Carcinogen-induced Animals

Numerous animal models are used to study inhibition of chemical carcinogenesis (**Table 4**). (See the section on *In Vitro and Animal Models for Human Cancer*.) Typically, a carcinogen is administered to the animal at a dose level high enough to induce a significant incidence of tumours in a specific target tissue. The carcinogen dose and treatment schedule are selected to ensure that the tumour incidence is not so high as to mask the potential of the inhibitor to reduce tumour incidence. The inhibitor is administered before, at the same time, after or in any combination of these times relative to the administration of the carcinogen. The relative timing of the administration of the carcinogen and the inhibitor is useful in interpreting the mechanism of inhibition. For example, a compound that inhibits when it is administered before the carcinogen, but not when it is given after carcinogen treatment is completed, most likely affects carcinogen blocking activities.

Studies typically last as long as required for the carcinogen to induce a high tumour incidence. Because the activity of most of the carcinogens used is well known, these tests are usually shorter than chronic carcinogenicity studies, i.e. they last 6–12 months. Inhibition is usually measured as the percentage by which the inhibitor lowers the incidence, multiplicity or total number of tumours, or increases the latency of tumour induction. Sometimes such factors as tumour size and degree of invasiveness are considered. Results are usually based on microscopic evaluation of the target tissues, although gross pathology also may be used. For example, rat mammary gland tumours often are detected by palpation, and mouse skin tumours are determined visually. Some general guidelines have been suggested for interpreting the results of testing a potential inhibitor of carcinogenesis (**Table 5**) (Steele *et al.*, 1994).

Inhibition of Carcinogenesis in Transgenic and Gene Knockout Mice

Animal models which mimic specific characteristics of human carcinogenesis are valuable for fully evaluating chemopreventive efficacy and for determining appropriate carcinogenesis biomarkers for measuring chemopreventive activity. Transgenic and gene knockout mice which carry one or more well-characterized gene mutations predisposing to carcinogenesis are appropriate models (**Table 6**) (Kelloff, 2000).

CHEMOPREVENTIVE AGENT DEVELOPMENT

A collaboration between the USA National Cancer Institute (NCI) and the FDA resulted in conceptual and practical guidelines for developing cancer chemopreventive drugs (Kelloff *et al.*, 1995a; Kelloff *et al.*, 1997). These strategies are outlined in **Table 7**. First the efficacy of candidate drugs is characterized using epidemiological

Table 4 Carcinogen-induced animal models for chemoprevention efficacy studies[a]

Organ model	Species	Carcinogen	Endpoint: inhibition of
Buccal pouch	Hamster	DMBA	Squamous cell carcinoma, papilloma
Colon	Mouse	AOM, DMH, MAM	Adenocarcinoma, adenoma, aberrant crypt foci
	Rat	AOM, DMH, MAM, MNU	Adenocarcinoma, adenoma
Oesophagus	Rat	Nitrosamines	Squamous cell carcinoma, papilloma
Forestomach	Mouse	B(*a*)P	Squamous cell carcinoma, papilloma
Intestines[b]	Rat	AOM, DMH	Adenocarcinoma, adenoma
Liver	Mouse	Various	Hepatocellular carcinoma, adenoma
	Rat	AAF, DEN, DMN, me-DAB	Hepatocellular carcinoma, adenoma
Lung	Mouse	B(*a*)P, DMBA, NNK, urethane	Adenoma
	Hamster	DEN, MNU (trachea)	Squamous cell carcinoma, adenosquamouscarcinoma
Mammary glands	Mouse	DMBA	Adenocarcinoma, adenoma
	Rat	DMBA. MNU	Adenocarcinoma, adenoma
Pancreas	Hamster	BOP	Ductal adenocarcinoma, adenoma
	Rat	L-Azaserine	Acinar cell carcinoma
Skin	Mouse	UV radiation, B(*a*)P, B(*a*)P/TPA, DMBA, DMBA/TPA,	Carcinoma, papilloma
Stomach (and glandular stomach)	Rat	MNNG	Adenocarcinoma
Urinary bladder	Mouse	OH-BBN	Transitional cell carcinoma
	Rat	MNU, OH-BBN	Transitional cell carcinoma

[a]Abbreviations: AAF, 2-acetylaminofluorene; AOM, azoxymethane; B(*a*)P, benzo[*a*]pyrene; BOP, *N*-nitrosobis(2-oxopropyl)amine; DEN, *N,N*-diethylnitrosamine; DMBA, 7,12-dimethylbenz[*a*]anthracene; DMH, dimethylhydrazine; DMN, *N,N*-diethylnitrosamine; MAM, methylazoxymethyl acetate; me-DAB, 3′- methyl-*N,N*-dimethylaminoazobenzene; MNNG, *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine; MNU, *N*-methyl-*N*-nitrosourea; NNK, *N*-nitrosonornicotine; OH-BBN, *N*-butyl-*N*-(4-hydroxybutyl)nitrosamine;TPA, tetradecanoylphorbol acetate.
[b]Small or large intestine not specified.
See also Steele *et al.*, 1994, *Journal of Cellular Biochemistry*, Suppl. 20, 32–54.

Table 5 Criteria for evaluating chemopreventive efficacy in animal carcinogenesis models

Conclusive *evidence of chemopreventive effect:*

Agent must cause a statistically significant ($p<0.05$) decrease in tumour incidence, multiplicity, size or invasiveness, or a statistically significant increase in tumour latency (time to appearance of first tumour or time to 50% incidence) compared with carcinogen controls

In the absence of statistics, a twofold decrease in incidence, multiplicity, size, or invasiveness, or a similar increase in latency, is required to confirm a chemopreventive effect

The criteria above are for a single-dose level of agent. If at least three doses are tested and a dose–response inhibition or increase in latency is observed, the result may be considered chemopreventive, even if the effect is not statistically significant or of twofold magnitude at any dose tested

Suggestive *evidence of chemopreventive effect:*

No statistical analyses are performed, but the inhibition ranges from 35 to 50%

Other factors considered in determining adequacy of the test, regardless of the magnitude of the effect:

The numbers of animals in the treatment and control groups are sufficient to demonstrate statistical significance

Survival in both test and control groups is adequate to allow statistical evaluation, i.e. the toxicities observed due to carcinogen or chemopreventive agent treatment should not be so severe as to compromise the results of the study

Evidence of carcinogenicity is established in concurrent carcinogen-treated control animals

Body weight in animals treated with chemopreventive agent are not statistically significantly lower than those of carcinogen controls (body weight is a particularly meaningful and confounding factor in inhibition experiments, since decreased or delayed weight gain can be a measure of slowed growth. Slowed growth alone can depress tumorigenicity without other specific effects of a chemopreventive agent)

Tumour incidence in the carcinogen control group is significantly lower than expected

Table 6 Representative transgenic/gene knockout mouse models for chemoprevention studies[a]

Transgenic Mouse model	Target	Genetic lesions	Histological lesions
Min	Colon	Heterozygous *Apc*2549	Adenomas, adenocarcinomas, some CIS
Apc	Colon	Heterozygous *Apc*1638	Adenomas, adenocarcinomas
MLH1/*Apc*1638	Colon	Heterozygous MLH1 and *Apc*1638	Adenomas, carcinomas
MSH2/*Min*	Colon	Heterozygous MSH2 and *Apc*2549	Adenomas, carcinomas
pim	Lymphatic system	Amplified *pim*-1	T-cell lymphomas
TG.AC	Skin	Ha-*ras* mutation	Papillomas, possible carcinomas
TSG-p53	Skin	Heterozygous p53 deficient	Papillomas, possible carcinomas
A/JxTSG-p53	Lung	Heterozygous p53 deficient	Adenomas
A/JxUL53	Lung	Heterozygous p53 mutant	Adenomas
TGFβ 1	Liver, lung	Heterozygous TGFβ 1 mutant	Adenomas, carcinomas
v-Ha-*ras*	Skin	Ha-*ras* + human keratin K-1	Hyperplasia, hyperkeratoses, squamous papillomas
K14-HPV16	Skin	HPV-infected (K14-HPV16 heterozygote), oestradiol-treated + SV40T-antigen	Papillomas, condylomas
K14-HPV16	Cervix	HPV-infected (K14-HPV16 heterozygote), oestradiol-treated + SV40T-antigen	Cervical dysplasia
C3(1)-SV40	Prostate	Heterozygous rat prostatic steroid binding gene [C3(1)] + SV40 T-antigen	Dysplasia, adenoma, adenocarcinoma
C3(1)-SV40	Mammary glands	Heterozygous rat prostatic steroid binding gene [C3(1)] + SV40 T-antigen	Adenocarcinomas

[a]Abbreviations: HPV, human papilloma virus; TGF, transforming growth factor
See also Kelloff, 2000, *Advances in Cancer Research*, **278**, 199–334.

Table 7 FDA/NCI guidance on approaches to the development of chemopreventive drugs[a]

I. Preclinical efficacy studies recommended for initiation of phase I/II clinical trials for chemopreventive investigational drugs are (1) + (7), (2) + (4 or 5) + (8), (3) + (4 or 5) + (8), (6):
   1. *In vivo* tumour modulation with (statistically significant) reduced tumour incidence or multiplicity or increased latency
   2. *In vivo* tumour modulation with (statistically nonsignificant but dose-associated positive trend) reduced tumour incidence or multiplicity or increased tumor latency
   3. *In vivo* surrogate endpoint modulation (statistically significant)
   4. *In vitro* transformation modulation
   5. *In vitro* chemoprevention-related mechanistic studies
   6. Epidemiological study demonstrating a cancer-inhibitory effect of the specific agent in the target tissue
   7. *In vivo* concentration–effect relationship
   8. *In vitro* concentration–effect relationship

II. Preclinical safety studies for initiation of phase I/II clinical trials for chemopreventive investigational drugs are:
   1. General toxicity studies conducted in two species, rodent and nonrodent, of equal or greater duration than the proposed clinical trial or up to 6 months in rodents and 12 months in dogs. Route of administration should be equivalent to the intended clinical route, and drug substance should be that prepared for clinical trials (preferably the clinical formulation)
   2. Genotoxicity assessed in a battery of assays [Ames *Salmonella typhimurium*, gene mutation in mammalian cells *in vitro* (either L5178Y mouse $TK^{+/-}$ lymphoma cells or another cell line with an autosomal locus with documented sensitivity to mutagenic chemicals, such as Chinese hamster ovary AS52 cells), and cytogenetic damage *in vivo* (mouse bone marrow micronucleus and/or mouse or rat chromosomal aberration tests)]
   3. Segment I reproductive performance/fertility in rat and Segment II teratology in rat and rabbit should be conducted as early as possible, prior to large clinical trials or trials of long duration, and in accordance with the ICH and the Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs
   4. Combinations of chemopreventive drugs should be evaluated in at least one general toxicity study of appropriate duration in the most appropriate species for interactions in pharmacokinetics, toxicity, enzyme effects or other relevant parameters
   (Pharmacokinetics and metabolite profiles should be examined in conjunction with toxicity studies to aid in interpretation of findings and evaluation of relevance to humans)

(*Continued*)

Table 7 (*Continued*)

Required before clinical studies >1 year's duration (large phase II and phase III)

1. Completion of general chronic toxicity studies in two species (6 months in rodent, 12 months in non-rodent)
2. All special toxicity studies (assessing neurotoxicity, cardiotoxicity, etc., as appropriate) before phase III
3. Initiation, preferably completion, of at least one of the rodent carcinogenicity bioassay prior to initiation of large phase III studies

Required for new drug application

1. Segment III perinatal and postnatal development study in rats
2. Completion of two rodent carcinogenicity bioassays.

III. Phase I–III clinical studies for chemopreventive investigational drugs

Required phase I

1. Single-dose studies in both fasting and nonfasting normal subjects to characterize single-dose pharmacokinetics (i.e. absorption, distribution, metabolism, elimination) and acute toxicity
2. Repeated daily dose studies using multiple dose levels for a period of 1–3 months in normal subjects **or** up to 12 months in subjects at increased risk to cancer(s), for which the drug demonstrates efficacy in preclinical evaluation, to assess multiple dose pharmacokinetics and chronic toxicity. Participation of normal subjects for more than 1 month is considered based on available information (toxicity, clinical experience, etc.) for each drug on a case-by-case basis

Recommended phase I

Under III 1 above, include placebo control and pharmacodynamic evaluation of dose–response for modulation of selected drug effect or surrogate endpoint biomarkers. Subject follow-up upon completion of treatment will include evaluation of modulation of marker status

Phase II

1. *Phase IIa*. In the event that a clearly defined and standardized surrogate endpoint biomarker is not identified, then a randomized, blinded, parallel dose–response chronic dosing study will be conducted for 3 months or more in subjects at high risk of cancer at the site of investigation using dosing levels shown to be safe in prior Phase I studies. As a basis for the phase IIb study, the objectives are to evaluate measurements of candidate biomarkers (drug effect and/or surrogate endpoint) and the dose–response relationship of biomarker modulation and tolerance to modulation, to standardize assays and quality control procedures and to characterize chronic dosing toxicity
2. *Phase IIb*. Randomized, blinded, placebo-controlled chronic dosing study for 3 months or more in subjects at high risk of cancer at the site of investigation at one or more dosing levels shown to be safe and effective in modulating biomarkers. Study objectives are to establish dose–surrogate endpoint marker response and chronic dosing toxicity and to select a safe and effective dose based on surrogate endpoint marker response and chronic dosing toxicity

Phase III

Randomized, blinded, placebo-controlled clinical trials with the following objectives:

1. Demonstrate a significant reduction in incidence or delay in occurrence of cancer
2. Validate surrogate endpoints
3. Assess drug toxicity
4. Characterize the relationship of dose and/or pharmacokinetics to efficacy and toxicity
5. In case of formulation differences, establish the bioequivalence between the to-be-marketed formulation and the formulation used in pivotal clinical trials

[a]See also Kelloff *et al*., 1995, *Cancer Epidemiological Biomarkers and Prevention*, **4**, 1–10.

data, and the mechanistic assays, *in vitro* efficacy tests and animal models of carcinogenesis just described. The most promising candidates are evaluated for preclinical toxicity and pharmacokinetics as needed. Clinical development is then planned and implemented for those agents that meet the criteria for acceptable toxicity as well as efficacy. Often, additional efficacy and toxicity testing are done to test alternative routes of agent delivery, dosage regimens, new target tissues, combinations of agents for increased efficacy and decreased toxicity, and to evaluate toxicities seen in early clinical studies.

Clinical development of chemopreventive agents, as for other pharmaceuticals, is carried out primarily in phases I, II, and III trials. Phase I clinical trials are safety and pharmacokinetics studies. These trials include single-dose studies in both fasting and nonfasting normal subjects to characterize single dose pharmacokinetics and acute toxicity. Also, repeated daily dose studies to assess multiple dose pharmacokinetics and chronic toxicity are conducted using multiple dose levels for a period of 1–3 months in normal subjects or up to 12 months in subjects at increased risk of cancer(s), for which the drug demonstrates efficacy in preclinical evaluation. Participation of normal subjects for more than 1 month is considered based on available information (toxicity, clinical experience, etc.) for each drug on a case-by-case basis. In most cases, the phase I studies evaluate agent effects as well as agent serum (and, sometimes, tissue) levels. Agent effects measured are those

believed to be potentially associated with chemopreventive activity. For example, in studies of NSAIDs serum and tissue levels of PGE_2 would be measured. In studies with the ODC inhibitor, DFMO, tissue levels of polyamines are measured.

Considerations and current progress in clinical development of chemopreventive agents are described fully in the chapter Chemoprevention. (See also the rest of the section on *The Treatment of Human Cancer*.) Briefly, phase II trials are initial efficacy studies. These randomized, double-blind, placebo-controlled trials emphasize the evaluation of biomarkers of carcinogenesis that are highly correlated to cancer incidence and that may serve as surrogate endpoints for cancer incidence reduction. Phase III studies are randomized, blinded, placebo-controlled clinical efficacy trials. These studies are typically large and have the objectives of demonstrating a significant reduction in incidence or delay in occurrence of cancer, validating surrogate endpoints, further assessing drug toxicity and further characterizing the relationship of dose and/or pharmacokinetics to efficacy and toxicity.

Role of Biomarkers of Carcinogenesis

Because of cancer's long latency, reduced cancer incidence is an impractical endpoint for clinical evaluation of chemopreventive agents. Thus, biomarkers of carcinogenesis are being evaluated and validated as surrogate endpoints for chemoprevention trials. These biomarkers are addressed in both preclinical and clinical studies. The criteria for surrogate endpoint biomarkers are that they fit expected biological mechanisms (i.e. differential expression in normal and high-risk tissue, on or closely linked to the causal pathway for the cancer, modulated by chemopreventive agents and short latency compared with cancer), may be assayed reliably and quantitatively, may be measured easily and correlate with decreased cancer incidence). They must occur in sufficient incidence to allow their biological and statistical evaluation relevant to cancer.

The rationale for the testing done during chemopreventive agent development is described below (Kelloff, 2000).

Preclinical Efficacy Development

Efficacy testing starts with a battery of mechanistic assays representing a wide range of chemopreventive activities. Positive agents may then be screened in a battery of *in vitro* cell-based efficacy assays. Positive agents are then moved on to animal models – typically, the models used are selected based on clues provided by the mechanistic and cell-based assays. Translational research is carried out to prepare agents for clinical development. These studies include evaluating new animal carcinogenesis models, agent-delivery mechanisms, potential synergy of efficacy and safety by the use of agent combinations and rational approaches to the development of defined mixtures as the best means to further test and verify hypotheses being generated by epidemiology data on diet and cancer.

A recent example confirmed the chemopreventive potential of aerosolized steroids in lung and the possible advantage of topical drug delivery. In this study, carcinogen-induced mice were treated with aerosol budesonide. This approach is particularly promising for preventing lung carcinogenesis, but is applicable to several target organs, the primary advantage being to improve therapeutic index. That is, relatively low doses of drug are required to reach the cancer target directly, and the toxicity in non-target organs seen on systemic absorption after oral doses is reduced. A phase II clinical study of aerosol budesonide is now in progress in patients with precancerous lesions in the bronchus (see also the chapter *Chemoprevention*).

Agent Combinations

Another strategy to improve efficacy and lessen toxicity is combinations of agents. In some combinations of two agents with different presumed mechanisms of activity, synergistic or additive efficacy is seen. Such improved activity may allow either or both the agents to be administered at lower doses, thereby reducing potential toxicity. For example, synergistic activity has been observed in rat colon studies with combinations of DFMO and the NSAID piroxicam and in rat mammary with combinations of retinoids and antioestrogens, and these strategies are now being tested clinically. Another combination strategy uses a second agent to counter the toxicity of a known effective chemopreventive agent. An example is coadministration of the PGE_2 analogue misoprostol to counter the gastrointestinal toxicity associated with administration of NSAIDs.

Development of Dietary Components

Dietary components with chemopreventive activity typically start as complex mixtures from which optimal standardized mixtures and purified active substance are prepared and characterized. For example, two soy isoflavone mixtures containing genistein (the presumed active substance), other isoflavones (primarily daidzein), fat and carbohydrate are being developed in parallel. One is nearly 'pure,' containing 90% genistein; the second more closely resembles a natural soy product, containing less than 50% genistein. Both mixtures will continue in development until one proves to have superior efficacy or practicality. Alternatively, the two may be developed for different chemoprevention uses. Similarly, well-characterized tea polyphenol extracts have been evaluated in preclinical studies, and epigallocatechin gallate (EGCG),

which appears to be a primary active component, is being developed in parallel. The effort to confirm dietary leads is expected to burgeon over the next few years. For example, the FDA has proposed guidelines for the identification and evaluation of heterogeneous botanicals such as the tea and isoflavone mixtures, and the number of publications on chemopreventive effects of characterized dietary components is increasing (e.g. many on tea polyphenols, curcuminoids, selenized garlic/selenomethylcysteine, conjugated fatty acids, and broccoli compounds (sulforaphane)). Increasingly sophisticated analyses of epidemiological dietary data could produce many more new chemopreventive hypotheses regarding dietary components.

Toxicology and Pharmacology

As for other pharmaceuticals, the FDA requires sufficient preclinical toxicity and phase I clinical safety and pharmacokinetics testing to ensure that an investigational chemopreventive agent will not jeopardize the health of patients in efficacy trials. Because they are intended for chronic use in relatively well subjects, the safety criteria are more stringent for chemopreventives than for many other classes of pharmaceuticals (Kelloff *et al.*, 1995a).

Preclinical Toxicity and Pharmacokinetics

Preclinical safety studies for chemopreventive drugs are generally the same as for other drugs and include acute and subchronic toxicity (incorporating pharmacokinetic measurements), reproductive performance and genotoxicity. Generally included are single-dose, acute toxicity study and an absorption–elimination study in rats, and subchronic repeated daily dosing studies in rodents and dogs. Combinations of chemopreventive drugs are evaluated in the species most closely related to humans in terms of metabolism in at least one study of appropriate duration (generally studies greater than 90 days are not needed) to determine interactions in pharmacokinetics, toxicity, enzyme effects, or other relevant parameters.

Preclinical efficacy studies also incorporate limited toxicity evaluations that may help identify appropriate doses for the formal toxicity studies. For example, most animal efficacy screens include a preliminary 2–6-week study to determine the maximum tolerated dose (MTD) of the test agents. Blood levels of test agent are usually obtained during animal studies designed to characterize more fully the efficacy of an agent or agent combination.

Pharmacokinetic data help in the development of a phase I clinical dose escalation strategy. Absorption–elimination studies in rats are used to develop analytical methods for drug monitoring, which can be standardized and used in the clinic. These studies also provide other information on agent behaviour (e.g. protein binding). Single-dose pharmacokinetics are also assessed at the initiation of the repeated daily dosing studies in dogs, and measurements of plasma drug levels at steady state are performed in these studies in rodents and dogs (pharmacokinetic studies using radioactive drug to quantify tissue distribution and metabolism are performed later in development). The information developed at this stage (e.g. maximum and minimum blood levels, elimination time) is evaluated with information from efficacy studies in order to provide a dose–concentration–effect profile of the test agent and to estimate a margin of safety; the relationship of dose to effectiveness and toxicity is then used to refine dosing strategies and regimens. For example, if a promising chemopreventive agent has slight toxicity with daily dosing it might be evaluated in further phase I trials using intermittent dosing schedules chosen to induce or inhibit a drug-effect enzyme over the whole treatment period while not reaching blood concentrations that may cause side effects.

As needed, a battery of three genotoxicity tests is performed. Chronic toxicity, carcinogenicity and reproductive toxicity tests are undertaken later in development, prior to or during phase III development. Special toxicity studies are also undertaken, as appropriate, in response to safety issues arising on clinical use of the agent. One example of special studies recently pursued is for DFMO which is now in phase II/III clinical trials. In previous clinical studies, this agent had shown significant ototoxicity. The mechanism appeared to be destruction of cochlear cilia. Thirteen-week studies in dogs were undertaken to evaluate the effect quantitatively. At clinical doses no effects were seen on cochlear hair cell measurements, brainstem auditory evoked responses (including histology of auditory nuclei) or observed response to auditory stimuli such as clapping and calling.

For most investigational drugs used chronically to treat disease states, carcinogenicity studies are required prior to petitioning the FDA for approval to market the drug. Generally one rodent carcinogenicity study is initiated prior to initiation of large phase III clinical studies. However, for drugs under development for cancer prevention, where the agent is to be used prophylactically in essentially well people, completion of one carcinogenicity study prior to conduct of sizable long-term trials is considered based on the expected toxicity of the drug, the population, the planned clinical trial duration, the trial design and other factors.

Phase I Clinical Safety and Pharmacokinetics

Phase I single-dose studies are designed to characterize agent pharmacokinetics and tolerability. The dose and

schedule of administration are based initially on preclinical toxicity and efficacy and are selected to achieve safe and effective plasma agent levels in humans. As is typical for other pharmaceuticals, the maximum initial dose in humans is a mg kg^{-1} dose that is the lower of one-tenth the highest no observed adverse effect dose (NOAEL) in rodents or one-sixth the highest NOAEL in mg kg^{-1} in nonrodents. The NOAEL is based on toxicity studies of equal or greater duration than the proposed clinical trial. The *in vitro* inhibitory concentration and *in vivo* plasma drug levels from efficacy testing may be used as a relative guide to the needed concentration, i.e. within an order of magnitude, but are not easily quantitatively extrapolated due to the conditions used in screening tests (i.e. the high dose of carcinogen). Ordinarily the human dose, usually in mg kg^{-1}, is not escalated above the animal NOAEL, but this could depend on the nature of the adverse effect. Higher human doses may be justified based on pharmacokinetic or pharmacodynamic differences between humans and animals or clinical experience at lower doses. Where possible the dose escalation strategy uses pharmacokinetic parameters across species. After a cautious initial dose, further escalation is based on blood levels compared with those associated with toxicity in animals. Differences in the pharmacokinetic profile after acute and chronic dosing are also evaluated. Dose selection is, of course, ultimately controlled by emphasis on empirical clinical safety and toxicity observations.

Consistent with current FDA regulatory practice, normal subjects are used in studies 1–3 months in duration; participation of normal subjects for more than 1 month is considered based on available information (toxicity, clinical experience, etc.) for each drug on a case-by-case basis. When longer phase I studies are undertaken (up to 12 months), subjects at increased risk for cancer(s) are enrolled. Longer studies are designed not only to obtain pharmacokinetic and safety information after chronic administration but also to develop and evaluate effects on drug activity and carcinogenesis biomarkers.

Table 8 Requirements for successful phase II clinical chemoprevention trials

Agent
Experimental and/or epidemiological data supporting chemopreventive activity (efficacy)
Safety on chronic administration at multiple of efficacious dose
Mechanistic rationale for chemopreventive activity
Cohort
Suitable for chemopreventive activity of agent
Suitable for measurement of biomarkers
Risk/benefit analysis acceptable
Biomarkers
Fits expected biological mechanism
– differentially expressed in normal and high-risk tissue
– on or closely linked to causal pathway for cancer (e.g. expression increases/decreases with severity of precancer; intraepithelial neoplasia is the most promising in this regard)
– modulated by chemopreventive agents
– latency is (relatively) short compared with cancer
Biomarker and assay provide acceptable sensitivity, specificity, and accuracy
– assay for biomarker is standardized and validated
Sampling is reliable
– dose-related response to the chemopreventive agent is observed
– statistically significant difference is seen between levels in treatment groups and controls
Biomarker is easily measured
– biomarker can be obtained by noninvasive or relatively noninvasive techniques
– assay for biomarker is not technically difficult
Biomarker modulation correlates to decreased cancer incidence (i.e. the biomarker can be validated as a surrogate endpoint for cancer incidence)

Clinical Efficacy

Clinical trials that support claims of chemopreventive efficacy can be designed using biomarkers as surrogate endpoints for cancer incidence. At each major target site, three components govern the design and conduct of these trials – well-characterized agents, reliable biomarkers for measuring efficacy and suitable cohorts (**Table 8**) (Kelloff, 1996b; Kelloff, 2000; Kelloff *et al.*, 2000).

The promise of chemoprevention is evidenced by the increasing number of clinical strategies and studies in the major cancer target organs – prostate, breast, colon, lung, head and neck, bladder, oesophagus, cervix, ovary, skin, liver (**Table 9**). In these organ systems cancer is associated with earlier, well-defined precancers (usually intraepithelial neoplasia) that may serve as biomarkers of carcinogenesis and surrogate endpoints for cancer incidence, as well as targets for treatment in their own right. Some examples are colon adenomas, bronchial dysplasia (lung), prostatic intraepithelial neoplasia (PIN), cervical intraepithelial neoplasia (CIN), actinic keratoses (skin), oral leucoplakia and ductal carcinoma *in situ* (breast). The remarkable progress that has been made in chemoprevention of cancer is evidenced by FDA approvals for the use of tamoxifen to reduce the risk of breast cancer in high-risk women and of celecoxib in the treatment of colorectal adenomas in patients with FAP. This progress is described fully in the chapter Chemoprevention. (See also the rest of the section on *The Treatment of Human Cancer*.)

Table 9 Aspects of chemoprevention at major cancer target sites[a]

	Prostate	Breast	Colon	Lung	Head and neck
Risk factors/ markers	Age >50 years; familial history of prostate cancer; high serum testosterone; high-fat diet/high red meat consumption; population/geographical background (highest incidences in Canada and northwest Europe); prostatitis, genetic polymorphisms (e.g. in SRD5A2, gene for steroid 5-reductase); low micronutrient levels (e.g. selenium, carotenoids, vitamin D)	Age >50 years, familial history of breast cancer or genetic syndrome (e.g. Li–Fraumeni, BRCA1), previous breast, endometrial or ovarian cancer, atypical hyperplasia, DCIS, LCIS; oestrogen exposure (e.g. early menarche, late menopause, late age at first full-term pregnancy); life-style factors (e.g. diet)	High-fat/low-fibre diet, low fresh fruit, vegetable intake, low calcium and vitamin D intake; familial history of genetic syndrome (e.g. FAP, HNPCC); familial/past history of colorectal cancer or adenomatous polyps; past history of breast or endometrial cancer; inflammatory bowel disease	Tobacco use (smoking, chewing); alcohol consumption, especially combined with tobacco use; occupational exposure (e.g. asbestos, nickel, copper); cytochrome P-450 genetic polymorphisms (e.g. CYP1A1, GSTM2); low fruit, vegetable consumption; previous oral, laryngeal, lung cancer	Tobacco use (smoking, chewing); alcohol, especially combined with tobacco use; males, 50–70 years
Promising agents	Steroid 5-reductase inhibitors (e.g. finasteride); retinoids (e.g. 9-*cis*-retinoic acid); RAMBA; antiproliferatives (e.g. DFMO, DHEA analogues); differentiating agents (e.g. vitamin D analogues); antioxidants (e.g. vitamin E, selenium, lycopene); GSH-enhancing agents (e.g. oltipraz); antioestrogens (e.g. toremifene, tamoxifen, raloxifene and other SERMs); aromatase inhibitors (e.g. vorozole); antiandrogens (e.g. leuprolide, flutamide); angiogenesis inhibitors (e.g. linomide); signal transduction regulators (e.g. soy isoflavones), Anti-inflammatories (e.g. lipoxygenase inhibitors, selective COX-2 inhibitors)	Antioestrogens (e.g. tamoxifen, raloxifene and other SERMs); aromatase inhibitors; antiproliferatives (e.g. DFMO); soy isoflavones; fluasterone; retinoids (e.g. fenretinide, 9-*cis*-retinoic acid); monoterpenes (e.g. limonene, perillyl alcohol)	Anti-inflammatories (e.g. sulindac, piroxicam, aspirin, selective COX-2 inhibitors, curcumin, iNOS inhibitors, ASA derivatives); antiproliferatives (e.g. calcium, DFMO, ursodiol)	Retinoids, (e.g. vitamin A, 13-*cis*-retinoic acid, fenretinide, all-*trans*-retinoic acid); antimutagens (e.g. oltipraz, anethole trithione, PEITC); anti-inflammatories (e.g. aerosolized corticosteroids, lipoxygenase inhibitors, COX-2 inhibitors)	Retinoids/carotenoids (e.g. vitamin A, 13-*cis*-retinoic acid, fenretinide, β-carotene); anti-inflammatories (e.g. tea, curcumin)

Intermediate biomarkers	Histological: PIN (nuclear morphometry, nucleolar morphometry, nuclear texture, DNA ploidy); proliferation: loss of high molecular weight cytokeratins (50–64 kDa), altered blood group antigens (e.g. Lewis^{Y} antigen), vimentin; genetic/regulatory: c-*erb*B-2, TGF, P53, *bcl*-2/*bax*, pc-1 chromosomal loss or gain (e.g. 8p, 9p AND 16q), TGF, IGF-1; biochemical: PSA levels, PAP levels; angiogenesis: microvessel density, VEGF	Atypical hyperplasia, DCIS, nuclear morphometry, ploidy, c-*erb*B-2 amplification, p53 mutation, IGF-I	Adenomas (recurrence, regression); ACF; nuclear and nucleolar morphometry; apoptosis; proliferation indices (PCNA, Ki-67); crypt proliferation kinetics; differentiation indices (Lewis blood group antigens, sialyl-Tn antigen)	Cellular atypia in sputum, bronchial atypical metaplasia/dysplasia, increased cytokeratin 19 expression, PCNA, blood group-related antigens, p53 mutation; RAR induction	Leucoplakia with dysplasia, erythroplakia, GGT, keratins, c-*erb*B-1 amplification; LOH; proliferation indices (PCNA, Ki-67)
Clinical cohorts: Phase II	Patients scheduled for radical prostatectomy; patients with PIN; patients with cancer on biopsy, treated by watchful waiting; patients at high risk of biochemical failure or rising PSA postradical prostatectomy; subjects with positive family history	Patients scheduled for breast cancer surgery, patients with LCIS or mammographically detected calcifications/DCIS, high risk with multiple biomarker abnormalities	Patients with previous colon cancer or adenomatous polyps, FAP patients; HNPCC patients/carriers	Patients with recently resected stage I lung or laryngeal cancer, chronic smokers with squamous metaplasia/dysplasia	Patients with dysplastic leucoplakia; patients with previous head and neck cancers
Clinical cohorts: phase III	HGPIN; men at high risk (e.g. PSA >4 ng ml^{-1} and negative biopsy); men from general population, age $\geq$55 years, normal PSA and DRE	Women $\geq$60, or 35–59 years old with risk factors for 60 years old, patients with previous breast cancer	Patients with previous colon cancer or adenomatous polyps, FAP patients; HNPCC patients/carriers	Men exposed to asbestos or patients with asbestosis, chronic or heavy cigarette smokers, patients with previous lung, head or neck cancers	Patients with previously treated head and neck cancer; subjects at high risk (e.g. smokers, tobacco chewers)

(*Continued*)

Table 9 (*Continued*)

	Bladder	Oesophagus	Cervix	Skin	Liver
Risk factors/ markers	Males, age>55 years, cigarette, pipe, and cigar smoking; occupational exposures to aromatic amines; metabolic polymorphism: slow *N*-acetyltransferase phenotype; chronic cystitis or urinary tract infections; coffee drinking; chlorinated tap water	Alcohol and tobacco use; poor diet (e.g. lacking fresh fruit and vegetables); chronic gastro-oesophageal reflux disease (GERD) for the cancer, as well as for Barrett oesophagus; genetic syndrome (e.g. tylosis); Barrett oesophagus	HPV infection; early age at first intercourse; multiple sexual partners; oral contraceptive use; immunodepression; smoking	Dermatological factors (e.g. fair skin, freckling); genetic susceptibility (e.g. xeroderma pigmentosum, basal cell nevus syndrome, albinism, epidermodysplasia verruciformis); environmental exposures (e.g. UV radiation, cigarette smoke, tanning booths, PAH); lupus; immunosuppression	Alcohol consumption; smoking; HBV and HCV infection
Promising agents	Anti-inflammatories (e.g. sulindac, piroxicam, aspirin, ibuprofen); antiproliferatives (e.g. DFMO); retinoids (e.g. fenretinide)	Antiproliferatives (e.g. DFMO, BBI, selenium); antioxidants (e.g. tea polyphenols, PEITC, diallyl sulfide); anti-inflammatories (e.g. NSAIDs, selective COX-2 inhibitors); vitamins (e.g. vitamin E, nicotinic acid, riboflavin)	Retinoids (e.g. vitamin A, fenretinide, 9-*cis*-retinoic acid); antiproliferatives (e.g. DFMO), folic acid	Anti-inflammatories (e.g. piroxicam, curcumin, selective COX-2 inhibitors); antimutagens (e.g. oltipraz, diallyl sulfide); antiproliferatives (e.g. DFMO); retinoids (e.g. fenretinide, 13-*cis*-retinoic acid, retinyl palmitate, vitamin A); antioxidants (e.g. tea polyphenols, selenium, carotenoids)	Antimutagens (e.g. oltipraz); retinoids (e.g. polyprenoic acid)
Intermediate biomarkers	TIS, dysplasia, DNA content, LOH, Rb, blood group-related antigens, F- and G-actins, integrins	Barrett's oesophagus (area and grade of dysplasia); nuclear/nucleolar polymorphism; DNA ploidy; proliferation indices (Ki-67); apoptosis; p53; EGFR, EGF, TGFα, LOH (e.g. chromosome 17); microsatellite instability; iNOS expression	CIN (grade); aneuploidy; nuclear polymorphism; proliferation indices (e.g. PCNA, EGFR, TGFα, TGFβ) differentiation markers (e.g. involucrin) *ras* oncogene expression	Actinic keratosis; proliferation indices (e.g. PCNA, IGF-I/IGFR, EGFR cyclin D1, ODC); TGFβ; differentiation indices (e.g. integrins); genetic/regulatory biomarkers (e.g. c-*fos*, c-*myc*, c-*jun*)	Carcinogen-DNA adducts

Clinical cohorts: phase II	Patients with previous resected superficial transitional cell carcinoma (TA/T1 with or without TIS), patients with previous resected superficial transitional cell carcinoma treated with BCG	Patients with low-grade, intestinal-type Barrett's oesophagus with or without dysplasia	HPV-negative patients with CIN III	Patients with actinic keratosis	Subjects with environmental exposure (e.g. carcinogen or HBV)
Clinical cohorts: phase III	Subjects at high risk (e.g. occupational exposure to aromatic amines)	Patients at high risk of oesophageal cancer (e.g. GERD, smokers, geographic/ethnic, such as Linxian, China)	Patients with CIN I, II; patients with HPV infection	Patients with previous BCC or SCC; subjects with previous intense chronic or episodic sun exposure; patients with pre-existing dermatological disorders; patients with actinic keratosis; subjects with dysplastic naevi	Patients with previous hepatoma

[a]Abbreviations: ASA, acetylsalicylic acid; BCC, basal cell carcinoma; BBI, Bowman–Birk protease inhibitor; BCG, *Bacillus Calmette Guerin*; CIN, cervical intraepithelial neoplasia; COX, cyclooxygenase; DCIS, (breast) ductal carcinoma *in situ*; DFMO, 2-dimethylfluorornithine; DHEA, dehydroepiandrosterone; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; FAP, familial adenomatous polyposis; GERD, gastro-oesophageal reflux disease; GGT, γ-glutamyltransaminase; HBV, hepatitis B virus; HCV, hepatitis C virus; HNPCC, hereditary nonpolyposis colorectal cancer (syndrome); HPV, human papilloma virus; IGF, insulin-like growth factor; iNOS, inducible nitric oxide synthase; LCIS, (breast) lobular carcinoma *in situ*; LOH, loss of heterozygosity; NSAID, nonsteroidal anti-inflammatory drug; ODC, ornithine decarboxylase; PAHs, polycyclic aromatic hydrocarbons; PAP, prostatic acid phosphatase; PCNA, proliferating cell nuclear antigen; PEITC, phenylethyl isothiocyanate; PIN, prostatic intraepithelial neoplasia; PSA, prostate-specific antigen; RAMBA, retinoic acid metabolism blocking agent; RAR, retinoic acid receptor; SCC, squamous cell carcinoma; SERM, selective oestrogen receptor modulator; TGF, transforming growth factor; TIS, transitional cell carcinoma *in situ*; VEGF, vascular endothelial growth factor.
See also Kelloff, 2000, *Advances in Cancer Research*, **278**, 199–334.

REFERENCES

AACR Chemoprevention Working Group (1999). Prevention of cancer in the next millenium: report of the chemoprevention working group to the American Association of Cancer Research. *Cancer Research*, **59**, 4743–4758.

De Flora, S and Ramel, C. (1988). Mechanisms of inhibitors of mutagenesis and carcinogenesis. Classification and overview. *Mutation Research*, **202**, 285–306.

Hartman, P. E. and Shankel, D. M. (1990). Antimutagens and anticarcinogens: a survey of putative interceptor molecules. *Environmental Molecular Mutagenesis*, **15**, 145–182.

Hong, W. K. and Sporn, M. B. (1997). Recent advances in chemoprevention of cancer. *Science*, **278**, 1073–1077.

Kelloff, G. J. (2000). Perspectives on cancer chemoprevention research and drug development. *Advances in Cancer Research*, **278**, 199–334.

Kelloff, G. J., *et al*. (1995a). Approaches to the development and marketing approval of drugs that prevent cancer. *Cancer Epidemiology Biomarkers and Prevention*, **4**, 1–10.

Kelloff, G. J., *et al*. (1995b). Inhibition of chemical carcinogenesis. In: Arcos, J., *et al*. (eds), *Chemical Induction of Cancer: Modulation and Combination Effects*. 73–122 (Birkhäuser Boston, Boston).

Kelloff, G. J., *et al*. (1996a). Strategy and planning for chemopreventive drug development: clinical development plans II. *Journal of Cellular Biochemistry*, Suppl. 26, 54–71.

Kelloff, G. J., *et al*. (1996b). Strategies for identification and clinical evaluation of promising chemopreventive agents. *Oncology*, **10**, 1471–1484.

Kelloff, G. J., *et al*. (1997). Progress in clinical chemoprevention. *Seminars in Oncology*, **24**, 241–252.

Kelloff, G. J., *et al*. (2000). Perspectives on surrogate endpoints in the development of drugs that reduce the risk of cancer. *Cancer Epidemiology Biomarkers and Prevention*, **9**, 127–134.

Lippman, S. M., *et al*. (1998). Cancer chemoprevention: progress and promise. *Journal of the National Cancer Institute*, **90**, 1514–1528.

Powis, G. (1994). Recent advances in the development of anticancer drugs that act against signalling pathways. *Tumori*, **80**, 69–87.

Powis, G. and Workman, P. (1994). Signalling targets for the development of cancer drugs. *Anticancer Drug Design*, **9**, 263–277.

Singh, D. K. and Lippman, S. M. (1998a). Cancer chemoprevention part 1: retinoids and carotenoids and other classic antioxidants. *Oncology*, **12**, 1643–1660.

Singh, D. K. and Lippman, S. M. (1998b). Cancer chemoprevention part 2: hormones, nonclassic antioxidant natural agents, NSAIDs and other agents. *Oncology*, **12**, 1787–1803.

Sporn, M. B. (1976). Approaches to prevention of epithelial cancer during the preneoplastic period. *Cancer Research*, **36**, 2699–2702.

Sporn, M. B. and Suh, N. (2000). Chemoprevention of cancer. *Carcinogenesis*, **21**, 525–530.

Steele, V. E., *et al*. (1994). Preclinical efficacy evaluation of potential chemopreventive agents in animal carcinogenesis models: methods and results from the NCI chemoprevention testing program. *Journal of Cellular Biochemistry*, Suppl. 20, 32–54.

Steele, V. E., *et al*. (1996). Use of *in vitro* assays to predict the efficacy of chemopreventive agents in whole animals. *Journal of Cellular Biochemistry*, Suppl. 26, 29–53.

Wattenberg, L. W. (1978). Inhibition of chemical carcinogenesis. *Journal of the National Cancer Institute*, **60**, 11–18.

Wattenberg, L. W. (1985). Chemoprevention of cancer. *Cancer Research*, **45**, 1–8.

FURTHER READING

Crowell, J. A., *et al*. (1994). Chronic toxicity studies of the potential cancer preventive 2-(difluoromethyl)-*dl*-ornithine. *Fundamentals of Applied Toxicology*, **22**, 341–354.

Fabian, C. J., *et al*. (2000). Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. *Journal of the National Cancer Institute*, **92**, 1217–1227.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759–767.

Fisher, B., *et al*. (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute*, **90**, 1371–1388.

Harris, C. C. (1991). Chemical and physical carcinogenesis: advances and perspectives for the 1990s. *Cancer Research*, **51**, 5023S–5044S.

Hawk, E., *et al*. (1999). Chemoprevention in hereditary colorectal cancer syndromes. *Cancer (Supplement)*, **86**, 2551–2563.

Henderson, B. E., *et al*. (1992). Environmental carcinogens and anticarcinogens. In: Wattenberg, L., *et al*. (eds), *Cancer Chemoprevention*. 3–17 (CRC Press, Boca Raton, FL).

Ilyas, M., *et al*. (1999). Genetic pathways in colorectal and other cancers. *European Journal of Cancer*, **35**, 335–351.

Kelloff, G. J., *et al*. (1996). Clinical development plan: tea extracts, green tea polyphenols, epigallocatechin gallate. *Journal of Cellular Biochemistry*, Suppl. 26, 236–257.

Kelloff, G. J., *et al*. (1996). Epidermal growth factor receptor tyrosine kinase inhibitors as potential cancer chemopreventives. *Cancer Epidemiology Biomarkers and Prevention*, **5**, 657–666.

Kelloff, G. J., *et al*. (1997). Farnesyl protein transferase inhibitors as potential cancer chemopreventives. *Cancer Epidemiology Biomarkers and Prevention*, **6**, 267–282.

Kelloff, G. J., *et al*. (1998). Aromatase inhibitors as potential cancer chemopreventives. *Cancer Epidemiology Biomarkers and Prevention*, **7**, 65–78.

Kelloff, G. J., *et al*. (1999). Cancer chemoprevention: progress and promise. *European Journal of Cancer*, **35**, 1755–1762.

Lipkin, M. (1992). Prototypic applications of intermediate endpoints in chemoprevention. *Journal of Cellular Biochemistry*, **16** (Suppl. G), 1–13.

Marnett, L. J. (1992). Aspirin and the potential role of prostaglandins in colon cancer. *Cancer Research*, **52**, 5575–5589.

Moon, R. C., *et al*. (1992). Chemoprevention of MNU-induced mammary tumors in the mature rat by 4-HPR and tamoxifen. *Anticancer Research*, **12**, 1147–1153.

Moon, R. C., *et al*. (1993). Chemoprevention of OH-BBN-induced bladder cancer in mice by piroxicam. *Carcinogenesis*, **14**, 1487–1489.

Oshima, M., *et al*. (1996). Suppression of intestinal polyposis in *Apc* 716 knockout mice by inhibition of cyclooxygenase 2 (COX-2). *Cell*, **87**, 803–809.

Rao, K. V. N., *et al*. (1996). Differential activity of aspirin, ketoprofen, and sulindac as cancer chemopreventive agents in the mouse urinary bladder. *Carcinogenesis*, **17**, 1435–1438.

Reddy, B. S., *et al*. (1990). Chemoprevention of colon carcinogenesis by concurrent administration of piroxicam, a nonsteroidal antiinflammatory drug, with D,L-α-difluoromethylornithine, an ornithine decarboxylase inhibitor, in diet. *Cancer Research*, **50**, 2562–2568.

Reddy, B. S., *et al*. (1996). Evaluation of cyclooxygenase-2 inhibitor for potential chemopreventive properties in colon carcinogenesis. *Cancer Research*, **56**, 4566–4569.

Schipper, H., *et al*. (1996). A new biological framework for cancer research. *Lancet*, **348**, 1149–1151.

Smalley, W. E. and DuBois, R. N. (1997). Colorectal cancer and nonsteroidal anti-inflammatory drugs. *Advances in Pharmacology*, **39**, 1–20.

Steele, V. E., *et al*. (1999). Lipoxygenase inhibitors as potential cancer chemopreventives. *Cancer Epidemiology Biomarkers and Prevention*, **8**, 467–483.

Steinbach, G., *et al*. (2000). The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *New England Journal of Medicine*, **342**, 1946–1952.

Taketo, M. M. (1998). Cyclooxygenase-2 inhibitors in tumorigenesis (Part I). *Journal of the National Cancer Institute*, **90**, 1529–1536.

Taketo, M. M. (1998). Cyclooxygenase-2 inhibitors in tumorigenesis (Part II). *Journal of the National Cancer Institute*, **90**, 1609–1620.

Wattenberg, L. W., *et al*. (1997). Chemoprevention of pulmonary carcinogenesis by aerosolized budesonide in female A/J mice. *Cancer Research*, **57**, 5489–5492.

Zenser, T. V. and Davis, B. B. (1992). Arachidonic acid metabolism. In: Steele, V. E., *et al*. (eds), *Cellular and Molecular Targets for Chemoprevention*. 225–243 (CRC Press, Boca Raton, FL).

Chapter 33

Introduction to the Diagnosis of Cancer

Sharon Poltis Wilczynski
City of Hope National Medical Center, Duarte, CA, USA

CONTENTS

- Introduction to the Diagnosis of Cancer
- Screening and Early Detection
- Routine Diagnosis
- Molecular Diagnostics
- Minimal Residual Disease (MRD)
- The Future

INTRODUCTION TO THE DIAGNOSIS OF CANCER

One of life's most horrifying experiences is to be told that you or someone you love has cancer. Unfortunately, each year over 1.2 million individuals in the United States alone are diagnosed with serious invasive cancers. The life of the individual diagnosed with cancer is profoundly altered spiritually and emotionally as well as physically by both the disease and the treatment, which is itself often life threatening. The impact of a cancer diagnosis is so great that it must be made with absolute certainty. For most, the road to the diagnosis starts with unexplained bleeding, pain or the presence of a lump somewhere in or on the body. After numerous medical tests, a piece of the abnormal tissue or some cells are removed by biopsy or fine needle aspiration. In the laboratory the tissue/cells are examined by a pathologist, a physician who specializes in the diagnosis of disease. The interpretation of the morphology along with incorporation of pertinent clinical and laboratory data are integrated into a final diagnosis, a task that is both art and science. Simply rendering a diagnosis of benign or malignant is far from sufficient. The tumour must be given an exact histological classification and if the tumour is removed surgically, the extent of spread is determined by a detailed examination of the specimen. Based on accumulated experience with previous patients having similar tumours with similar spread, the natural course of the disease is predicted and appropriate therapy initiated.

The word tumour originally meant 'swelling' caused by inflammation but now generally means a 'new growth' or neoplasm. A variety of reactive processes can simulate neoplasms but these are usually polyclonal and eventually respond to endogenous homeostatic mechanisms and cease to proliferate. By definition a neoplastic process is an abnormal growth of cells and implies a clonal proliferation in which all the tumour cells are descendants of a single cell that gained the ability to replicate autonomously. A benign tumour will grow at the site of origin usually as an expansile mass without infiltration into adjacent normal tissue and does not have the capability to spread to distant sites. Generally, benign tumours can be surgically removed and do not kill the patient, but if they originate in critical areas such as the brain, blood vessels or in the airway they too can be lethal. Malignant tumours or cancers, on the other hand, infiltrate and destroy surrounding tissue and have the ability to disseminate to distant organs, often even when the primary tumour is small. In general, the pattern of spread is somewhat predictable for each given tumour type, e.g. most epithelial cancers first spread to the regional lymph nodes via lymphatics. Most sarcomas (cancers arising in mesenchymal tissue), however, spread by the vascular system to distant sites and less commonly involve the lymph nodes. This ability of malignant cells to metastasize is the greatest obstacle to the successful treatment of cancer.

There are some tumours that are difficult to pigeonhole neatly as benign or malignant. For example, basal cell carcinomas of the skin aggressively infiltrate surrounding tissue but rarely metastasize to distant sites. Borderline ovarian tumours (also called ovarian tumours of low malignant potential) will widely disseminate in the peritoneal cavity but have limited ability to invade into the abdominal organs. Unfortunately, there are no markers (morphological, molecular or otherwise) that denote malignant cell populations regardless of morphology, clinical setting or anatomic location. It is unlikely that such markers exist but if one should ever be identified it would be equivalent to finding the 'holy grail' of pathology.

Pathology is the branch of medicine that studies the mechanisms of cell/tissue/organ injury and the structural changes that underlie disease processes. The surgical

pathologist is a physician who specializes in the examination of tissue from living patients, diagnosing diseases and guiding management of the patient. The surgical pathologist usually examines tissue sections while the cytopathologist specializes in the study of cells in smears, aspirates and fluids. This chapter concentrates on the process by which a pathological diagnosis is made in an academic medical centre in the United States with a discussion of the basic histological and immunological methods that are commonly available for routine diagnosis and clinical management. Experimental techniques that are likely to be in routine practice in the next few years are also briefly included. It is hoped that the reader will get a flavour of how surgical pathologists and cytopathologists interpret histological images and incorporate clinical and molecular findings to arrive at cancer diagnoses.

SCREENING AND EARLY DETECTION

Unfortunately, only a few cancers are detected by screening programmes and most are still found only after patients seek medical attention for symptoms related to their tumours. The underlying premise for successful cancer screening is that treatment improves outcome if the malignancy is detected before it is clinically evident. This is an assumption and must be validated for each tumour type that is targetted. Slow-growing tumours that metastasize late in their course and tumours which are resistant to any treatment are not good candidates for a screening programme. In the first case, the late-metastasizing tumours may be effectively treated when they become clinically obvious and in the second case, if there is no effective treatment for a specific type of cancer, even when small, screening serves no useful purpose.

There are three primary types of screening programmes, based on the stage of malignancy that is targetted. Some screening programmes are designed to detect precursor lesions, before they become invasive cancers and are very effective when removal of the preinvasive lesion is curative. The Papanicolaou (PAP) test for cervical cancer is the model for this type of cancer screening and has significantly contributed to reduction in death from cervical cancer in developed countries by identification and treatment of the preinvasive lesion (cervical intraepithelial neoplasia) (Miller *et al.*, 2000). ‘Organ-confined’ invasive cancer screening is exemplified by detection of prostate cancer by elevated levels of serum prostate-specific antigen (PSA). Screening for breast cancer by mammography detects both small invasive breast cancers that cannot be identified by physical examination as well as the preinvasive ductal carcinoma *in situ*. As our understanding of cancer on the molecular level increases, a third type of screening programme is becoming available that identifies predispositions to certain types of cancer by genetic testing. For example, individuals with strong family histories of cancer can be tested for genetic mutations in relevant genes such as *BRCA 1* (breast and ovarian cancer) or *DCC* (colon cancer). To be beneficial, therapeutic interventions or prevention strategies that are inexpensive and have little if any morbidity in a currently healthy population must be available for patients with such predispositions. Clinical testing for predisposition to cancer has huge emotional costs to patients and their families and it is essential that high-quality genetic counselling be given before the tests are ordered as well as at the time the results are discussed with the patient. Finally, the ethical and legal implications (such as effects on insurability) of identifying at risk populations have yet to be resolved.

Screening tests are not diagnostic, but rather identify asymptomatic, apparently healthy individuals that may have the disease from those that probably do not have it. The qualifiers ‘may’ and ‘probably’ reflect the sensitivity and specificity of the screening test. The sensitivity is defined as the percentage of individuals with the disease that the test was able to identify correctly while the specificity is defined as the percentage of individuals who are free of disease that are identified correctly (**Table 1**). Some other independent method must be used to make the final determination of who has disease and who does not in order to calculate these values. Ideally a screening test should be 100% specific and 100% sensitive, but in reality, the sensitivity and specificity are usually inversely related. A screening test that is highly sensitive identifies most of the individuals with disease but will often have a high percentage of false positives that lowers its specificity. Another critical factor in evaluating the effectiveness of a screening test is the prevalence of the disease in the population screened. If the disease is rare, even a screening test with excellent specificity and sensitivity will give a large number of false positives because of the many nondiseased individuals tested. This is usually expressed as ‘low predictive value’ for a positive test. In general, the positive predictive value of a screening test is high only if the disease is common in the population screened. Thus, limiting screening to a population at high risk due to age, family history or some other clinical variable is often necessary.

Further evaluation of all individuals who test positive in a screening programme is mandatory. Facilities for diagnosis and treatment must be available as well as resources to follow individuals that test positive until it is clear that they are cured or are truly without disease and not just in a period in which the cancer is so small that it is undetectable with current technology. Detection of presymptomatic cancers is often very difficult and requires expensive imaging tests and uncomfortable procedures. The monetary costs of screening programmes are substantial particularly when diagnostic medical procedures, lost work time and follow-up for individuals without disease are included in the calculations. The anxiety and stress of being told that a cancer screening

Table 1 Basic terms for evaluating screening tests

	Disease	No disease
Screening test positive	True positive (TP)	False positive (FP)
Screening test negative	False negative (FN)	True negative (TN)

Term	Definition	Calculation
Sensitivity	Percentage of individuals with the disease who are correctly identified by the test	$100 \times TP/(TP + FN)$
Specificity	Percentage of individuals without disease who are correctly identified by the test	$100 \times TN/(TN + FP)$
Per cent false negative	Percentage of individuals with the disease who were not detected by the test	$100 \times FN/(TP + FN)$
Per cent false positive	Percentage of individuals without disease who tested positive	$100 \times FP/(TN + FP)$
Positive predictive value	The likelihood that the individual who tests positive has the disease	$TP/(TP + FP)$
Negative predictive value	The likelihood that the individual who tests negative does not have the disease	$TN/(TN + FN)$

Table 2 Example of the results of a screening test with 95% sensitivity and 95% specificity for a cancer that is present in 1% of a population[a]

	Individuals with cancer	Individuals without cancer
Screening test positive	950 (true positive)	4950 (false positive)
Screening test negative	50 (false negative)	94050 (true negative)
Total	1000	99000

[a]In this example, in a population of 100 000 individuals, 5900 individuals will test positive, but only 950 have the cancer, a positive predictive value of 16%, while 50 of those with the cancer will be missed. This example is a 'best case' as most cancers will be less prevalent in the population screened and few tests are this sensitive and specific.

test is positive are considerable even when ultimately the individual is found to be free of disease. It is also inevitable that any screening test for cancer will miss a few individuals that actually have cancer and in our legalistic environment that can be a tragedy for all involved. As demonstrated in the example presented in **Table 2**, the evaluation of any cancer screening test must consider the disadvantages of applying additional medical tests to a large number of people who never had or never will develop the cancer for which they are screened.

With all these issues, it is no wonder that the concept of cancer screening is widely embraced but that there are only a few cancers for which it is actively pursued and shown to be effective in lowering the incidence of cancer death. The cost and infrastructure necessary to maintain screening programmes has been beyond the economic resources of many countries and even in the developed world the expense is staggering. It is beyond the scope of this chapter to discuss in detail specific cancer screening programmes and for a discussion the reader is referred to the published literature. (See also chapter on *Inherited Predispositions to Cancer*.)

ROUTINE DIAGNOSIS

To make a diagnosis is to recognize a disease process, distinguish if from all others and assign it a name. This requires, as King stated over 30 years ago (King, 1967), both knowledge and judgement. Knowledge is information acquired from text books, personal experience or journal articles. Knowledge is dynamic and always in a state of flux both for the profession as new information is acquired and for individuals as we learn or forget. Yet application of that knowledge to a specific patient biopsy requires inference and judgment, qualities that reflect both reality and a perception of reality. It is easy to determine knowledge by testing, but judgment is a much more elusive and subjective quality yet essential for the intellectual process of rendering a diagnosis.

The initial pathological assessment of any tumour in the body is to differentiate a reactive process from benign or malignant growths. In most cases this can be accomplished by morphological evaluation of a tissue section stained with haematoxylin and eosin (H&E). As we enter the twenty-first century, the visual interpretation of the H&E slide by a surgical pathologist is still the primary means by which almost all patients are diagnosed with cancer. In this age of incredible technology the method is archaic but in most cases it is efficient, economical and conclusive. Furthermore, the easily performed H&E stain is universally available in hospitals worldwide, so that a biopsy prepared in a remote corner of the globe can be easily transported and interpreted by experts in distant cities or across oceans. The knowledge obtained through interpretation of the H&E slide by generations of pathologists

and clinicians is amazingly concordant with the molecular abnormalities that underlie the malignant process being deciphered in laboratories today. Although this may seem surprising, it should not be as the visual interpretation of the architectural and cytological features in the H&E slide is based on the interactions of the chemical dyes with thousands of gene products that are a reflection of the distinctive differentiation state of the tissue.

H&E stands for haematoxylin and eosin, the most universally used histological stain. Haematoxylin is a dye extracted from the heartwood of a Central American legum tree, *Haematoxylin campechianum*, and was widely used in the 1800s in the textile industry for dying calico and woolen goods in shades of lavender and purple, and it was also used to colour wine. In the 1860–1870s it was adapted as a histological stain along with the xanthene dye eosin. Haematoxylin itself has little affinity for tissue, but with oxidation to haematein and in the presence of metal mordants (usually aluminium or iron) it binds to nucleoproteins, probably though electrostatic interactions. Eosin, a potassium salt of tetrabromofluorescein, was first synthesized by Baeyer in 1871. It stains cytoplasm and connective tissue a variety of shades of red and pink by chemical reactions with proteins, particularly those with basic amino acids such as arginine, lysine and histidine. Haematoxylin and eosin were combined by Busch in 1878 and over a century later the combination is still the most common means by which tissue sections are examined. A myriad of cellular substances interact with these two chemical compounds to produce patterns that convey an enormous amount of information that reflects the underlying biochemical and molecular processes occurring in the cells and tissue. With the aid of a microscope and this simple cellular stain, the surgical pathologist can diagnose almost all pathological processes from infectious diseases to cancer.

Most cancers are detected when they have grown large enough to interfere with the function of a specific organ or to cause pain or other physical symptoms. The patient, after reporting symptoms to a physician, usually will have numerous radiological and other studies to localize and characterize the process, but eventually a tissue biopsy (either surgical or fine needle aspiration) is required before therapy for cancer can be initiated. To prevent autolysis and to keep the cells in as close to their living state as possible, the tissue is 'fixed.' Although a large number of substances have been and are used, 10% neutral buffered formalin (with an actual working concentration of about 4% formaldehyde) is still the most widely used fixative. The chemistry of formalin fixation is complex and still not completely understood, but proteins and nucleic acids are reversibly cross-linked through the addition of methylene groups. Before microscopic sections can be prepared, thin pieces of the fixed tissue must be permeated by a supporting medium such as paraffin. First most of the water is removed from the fixed tissue by dehydration through graded alcohols. The tissue is then 'cleared' by removal of the dehydrant with a substance that is miscible with paraffin. Traditionally xylene was used but newer, less toxic and more environmentally friendly agents are now available. Finally, the tissue is permeated by the embedding agent, almost always paraffin. This process is automated and computer-controlled instruments move the tissue in cassettes from one solution to the next at preset times and often with applications of heat and vacuum to speed the process. Before the tissue can be sectioned it is removed from the cassette and must be oriented in the final block so that tissue relationships such as the junction between epithelium and stroma are preserved. This is a critical step that requires skill and a great deal of manual dexterity from the technician. After hardening, the tissue is now ready to be sectioned using a microtome, an instrument that has a very sharp knife and can advance the tissue block precisely to give sections 6–8 μm thick. Sections are floated on warm water to remove wrinkles and then picked up on glass microscope slides. After baking to adhere the tissue to the glass, the slides are now ready for staining. Since most histological stains are aqueous, the embedding process must be reversed to rid the tissue of paraffin. After running through xylene (or substitute), graded alcohols to water the section is finally ready for the H&E stain.

After staining, the H&E slide is examined under the microscope by a surgical pathologist and the tumour is characterized by its morphological phenotype as expressed in tissue architecture and cytological appearance. The process by which the slide is evaluated and the findings translated into a diagnosis is subjective and no two pathologists approach or 'see' the slide in exactly the same way. Usually it is examined first at low power (4×) to discern the overall pattern and to detect areas that have abnormalities. Once the overall pattern is appreciated and specific areas of the slide identified as worthy of additional study, the pathologists will look at the cells on medium and high power (20 or 40×). Any identified abnormal cell population is further examined for uniform or variable nuclear features, the size and shape of the nuclei and features of the nuclear membrane, the chromatin pattern, the degree of staining (hyperchromatic or hypochromatic), mitoses (number and abnormalities), the quantity and quality of cytoplasm, and the shape and pigmentation of the cells and their relationship to each other and to stromal cells. The type and extent of inflammatory or stromal reaction is also noted. A wide variety of 'special stains' can be performed in the histology laboratory to identify substances produced by the malignant cells, such as mucin (for epithelial mucins), PAS (distinguishes neutral polysaccharides and glycoproteins from sialic acid containing muco-substances) or Alcian blue (acid muco-substances). Other histochemical stains such as reticulum or trichrome are useful for evaluating the relationship of the tumour to the stroma. Although many of these substances are now defined antigenically and are

detected by immunohistochemistry, these special staining techniques are still useful, inexpensive and can be performed in most histology laboratories on the same day as the initial H&E section.

The experienced histopathologist collects the visual information from the H&E slide almost subliminally, analysing and comparing with a set of internalized patterns that have been accumulated in memory from personal experience. The diversity of the histological features in human tumours is great and rarely does a pattern conform to the idealized description and no two tumours are ever exactly alike. Before a final diagnosis is reached, the pathologist must be methodical in eliminating histological mimics and consider unusual entities that may show similar morphological features. They also must determine if the biopsy is representative of the clinical lesion and if the diagnosis 'fits' with the clinical setting. No diagnosis can be made without knowledge of the clinical setting, including the age and sex of the patient, the anatomical site, previous histological and/or radiological findings and past surgical, chemotherapy or other medical interventions. Changes that would indicate malignancy in a brain biopsy from a 40-year-old man may be normal if from a 2-week-old baby.

In surgical specimens, the extent of the tumour is determined by the involvement of specific anatomical structures leading to a pathological stage. Particularly important is the presence or absence of tumour metastasis in regional lymph nodes. Margins of resection are carefully evaluated and prognostic markers are assessed. In some settings the decision for additional therapy is based on the findings in the resected specimen. For example, osteosarcomas are often treated with chemotherapy prior to resection and the presence and amount of viable tumour in the specimen dictates whether or not additional chemotherapy is given. All of this information is communicated by the pathologist to the clinician in a formal written report. There is growing support for the utilization of standardized surgical pathology reports so that the pathological characteristics of a tumour resection specimen are recorded in a complete and consistent manner (Rosai *et al*., 1993).

The pathologist is ever more called on to render precise diagnoses on smaller and smaller fragments of tissue. Still, in most cases, the distinction between benign and malignant, and the classification and grade of the tumour can be determined solely by morphological evaluation of the H&E slide. Ancillary studies on the tissue such as immunohistochemistry and molecular analysis as discussed below can refine the diagnosis and contribute clinically useful information.

Tumour Classification

All tumour classification schemes are by their nature artificial and arbitrary as biological processes are generally continua. Yet specific landmarks can often be used to define boundaries not only in the spectrum from benign to malignant but also through the grades of malignancy and among the types of tumours. For example, the presence of stromal invasion is a critical parameter that separates carcinoma *in situ* of the cervix, a lesion with no metastatic potential, from an invasive cervical cancer that has the potential to metastasize and kill the patient. However, many of the boundaries are less clear and they can be very difficult to interpret in an individual case whether from inherent properties of the lesion or from inadequate tissue sampling. It is not surprising that pathologists will often vary in their thresholds for recognizing some specific diagnostic entities. However, meaningful reproducible criteria have been identified and are generally recognized for the diagnosis of most human cancers. With some exceptions, the most universally accepted classification schemes for human tumours are published in a series of monographs by the World Heath Organisation and in the systemic series of fascicles developed for the United States Armed Forces Institute of Pathology.

Currently tumours are classified based on their histogenesis, or the cell of origin for the neoplastic proliferation. Tumours that arise from the glandular epithelial cells of the body are called adenomas if benign or carcinomas if malignant. Malignant tumours from the supporting tissues derived from the embryonic mesodermal layer are sarcomas. Haematopoietic cells give rise to leukaemias and lymphomas. Tumours are thought to arise from stem cells, those cells that have retained the ability to replicate and differentiate into specialized tissues. Neoplasms are classified based on the type of differentiation the specific tumour cells display. Carcinomas with a glandular growth pattern are adenocarcinomas whereas those that have recognizable squamous differentiation are squamous cell carcinomas. Angiosarcomas have vascular channels whereas a chondrosarcoma will show cartilaginous differentiation. Identification of specific cellular products either morphologically by light or electron microscopy or by immunohistochemical detection is frequently helpful in the determination of the lineage of a tumour. For example, a squamous cell carcinoma can be recognized by the presence of intercellular bridges, an adenocarcinoma by the presence of acinar formation and/or mucin and a rhabdomyosarcoma by immunoreactivity with desmin. Although the classification of tumours is still primarily morphological, it has become more sophisticated and biologically more meaningful by incorporating immunohistochemical techniques that recognize specific proteins expressed by the tumour cells and by molecular studies that identify tumour-specific genetic changes.

By assigning a name to a tumour the probable behaviour and clinical implications are communicated. A diagnosis of cancer of the testis is essentially meaningless, but to call a testicular mass a seminoma indicates that the tumour is

probably localized and cured by surgery or if spread to lymph nodes it is treatable by radiation therapy. On the other hand, a diagnosis of embryonal or choriocarcinoma of the testis implies a more ominous prognosis and the likelihood of distant spread is high and aggressive chemotherapy is needed to control the disease.

Electron Microscopy in Tumour Classification

Ultrastructural analysis of a tumour was for many years the primary means by which a poorly differentiated tumour was classified. For example, using transmission electron microscopy, malignant nerve sheath tumours could be recognized by their long cytoplasmic processes, primitive cell junctions and fragmented external basal lamina. Rhabdomyosarcomas had 6-nm (actin) and 15-nm (myosin) filaments associated with Z discs. Identification of subcellular structures such as desmosomes (epithelial tumours), neurosecretory granules (neuroendocrine tumours) or melanosomes (malignant melanoma) are also examples of how the electron microscope was applied in diagnostic pathology. Although immunohistochemistry has replaced electron microscopy for many of these applications, it is still useful in some cases and ultrastructural localization of specific antigenic substances may be a powerful tool in the future (Herrera *et al*., 2000).

Immunohistochemistry in Tumour Classification

The diagnosis of tumours as benign or malignant is primarily based on morphological features, but immunohistochemistry is a powerful and complementary aid in classifying tumours and in identifying markers that have prognostic significance. The basic premise of classification by immunoreactivity is that malignant cells, even when very undifferentiated, continue to make antigenic substances that characterize the cell or tissue from which they arose. The same antigenic substance is present in normal, benign and malignant cells so that the expression of an antigen by itself is not usually diagnostic for cancer. It is absolutely essential that immunohistochemical analysis of a tumour be interpreted in the context of a carefully selected differential diagnosis based on the clinical setting and morphologic features of the tumour. To do otherwise is to invite diagnostic disaster.

One of the most challenging problems to a surgical pathologist is the diagnosis of a poorly differentiated tumour when no primary site is clinically evident. From a therapeutic perspective, it is most critical to correctly identify those tumours for which there is effective specific systemic chemotherapy or hormonal therapy, such as for lymphomas and metastatic carcinomas of the breast, prostate and ovary. Although extensive radiological and endoscopic work-ups can be done, these are often uninformative, expensive and may have significant morbidity. Frequently the use of antigenic markers can be helpful in narrowing the possibilities and a likely site can be suggested based on clinical setting, subtle morphological features of the tumour combined with a limited immunohistochemical panel of antibodies. Among the most useful antigens in this setting are the intermediate filaments that composed the cell's cytoskeletal framework (**Table 3**).

There is lineage-dependent expression of these proteins that is maintained even in very undifferentiated tumours. Immunohistochemical typing with appropriate antibodies can be very helpful in differentiating epithelial (expressing keratin filaments) from mesenchymal tumours (expressing vimentin filaments). Metastatic melanoma and lymphoma are also usually in the differential of a poorly differentiated tumour and the addition of antibodies to S100 and CD45 can be useful in recognition of these tumours. S100 is a calcium-binding protein that is strongly expressed in melanomas although it can also be seen in neural tumours, histiocytic lesions and some carcinomas. This demonstrates that no single antibody can be interpreted in isolation and evaluation by a panel of antibodies is necessary. A melanoma should also stain strongly with vimentin and will be negative for keratin and lymphoid

Table 3 Intermediate filaments

Intermediate filament	Location	Tumours
Keratin	Epithelial cells	All carcinomas, adenocarcinomas, squamous cell carcinomas, mesotheliomas
Desmin	Muscle tissue, smooth, skeletal and cardiac	Tumours derived from muscle such as leiomyosarcoma, rhabdomyosarcomas
Vimentin	Mesenchymal cells, fibroblasts, endothelial cells, muscle cells	Sarcomas including fibrosarcoma, liposarcomas, angiosarcomas, also lymphomas, melanomas
Neurofilament	Neural tissue, both central and peripheral, cells derived from neural crest	Adrenal and extra-adrenal pheochromocytomas, neuroblastomas
Glial fibrillary acidic protein	Glial cells, astrocytes, ependymal cells	Astrocytomas, ependymomas, gliomas

markers. Melanocytic differentiation can be confirmed by reactivity to premelanosome markers such as HMB45 or Mel-A.

Immunophenotyping of lymphomas is essential for their classification. A family of membrane protein tyrosine phosphatases, recognized by the antibody cluster CD45, is present on most haematolymphoid cells and their tumours. Membrane positivity for CD45 in an undifferentiated malignant neoplasm is virtually diagnostic of lymphoma. Most non-Hodgkin lymphomas have immunophenotypes that correspond to stages in the normal development of lymphoid cells. The current classification of lymphomas (Revised European–American Classification of Lymphoid Neoplasms (REAL)) depends heavily on the integration of morphological features with clinical, immunophenotypic and genetic features (Jaffe, 1999).

The cytokeratins have a diverse and unique expression pattern that has been found to be useful in identification of the site of origin for many epithelial tumours. The cytokeratins consist of a family of at least 20 different polypeptide chains and two-dimensional gel electrophoresis studies showed that these are more or less distributed in a tissue-specific manner. Tumours tend to retain the cytokeratin profile of the epithelial tissue from which they arose (Chu *et al*., 2000). About 95% of colonic carcinomas are keratin 7 negative and keratin 20 positive whereas the majority of adenocarcinomas in the differential such as endometrioid ovarian adenocarcinoma and most adenocarcinomas of the lung are keratin 7 positive and keratin 20 negative. **Table 4** lists the keratin 7/20 profile of some common epithelial cancers that are frequently in the differential of metastatic carcinoma of unknown origin.

Table 4 Use of keratin 7 and keratin 20 in the differential diagnosis of common epithelial cancers

Keratin 7	Keratin 20	Tumours
Positive	Positive	Transitional cell Ovarian cancer (mucinous) Pancreatic cancer
Positive	Negative	Breast (lobular and ductal) Lung (adenocancer, non-small cell) Ovarian cancer (serous) Endometrial adenocancer Epithelial mesothelioma Thymoma
Negative	Positive	Colorectal adenocancer
Negative	Negative	Hepatocellular cancer Renal cell cancer Prostatic cancer Squamous cell cancer Small cell (neuroendocrine)

A large number of antibodies that are helpful in the characterization of tumours include a few that are relatively tissue specific. For example, the antibody to PSA is relatively sensitive and specific for prostate tissue as is thyroglobulin for thyroid. Neuroendocrine tumours will react with markers for neurosecretory granules such as synaptophysin and chromogranin. Immunoreactivity for desmin is strong evidence that the tumour is of muscle derivation and only glial tumours (or some germ cell tumours differentiating into glial tissue) will express GFAP. For further reading on the role of immunohistochemistry in tumour diagnosis, the readers are referred to an excellent recent issue of *Seminars in Diagnostic Pathology* (Suster *et al*., 2000).

Although expression of a specific antigen is usually not diagnostic for malignancy, finding aberrant expression at a site that does not normally have immunoreactive cells can be diagnostically significant. For example, finding malignant glandular cells in an abdominal lymph node that express PSA is good evidence for metastatic prostate cancer. Similarly, keratin-positive cells are not usually found in bone marrow, and in a patient with breast cancer the presence of cells that react with keratin antibodies is evidence for metastatic disease. Morphologically atypical cells can be found in ascites and pleural fluids as a result of a reactive or malignant process. If the atypical cells in the fluids express epithelial antigens such as CEA, B72.3, CD15 or BerEP4, a malignant process is likely as these markers are not found on reactive mesothelial or inflammatory cells. Also, aberrant expression of antigens can be extremely helpful in establishing the diagnosis of lymphoma. Demonstration of a lymphoid population that co-expresses CD20 (a B cell marker) as well as CD43 (a T cell marker) is very strong evidence for malignant lymphoma.

Immunohistochemical identification of specific gene products in a patient's tumour is increasingly being used as an aid in the selection of therapy. Breast cancers expressing oestrogen and/or progesterone receptors respond to hormonal therapy and for the last decade immunohistochemistry has been the standard for determination of steroid receptor status. With the development of Herceptin® treatment for metastatic breast cancer, the immunodetermination of Her2/neu overexpression is also routinely performed.

Immunohistochemistry is based on a series of biochemical reactions that start with an antibody recognizing a specific tissue epitope and end with the visual detection of the antibody–antigen complex. Antibodies generally recognize a specific shape of an epitope formed either by continuous amino acid sequences or discontinuous residues that are conformationally folded into the recognized epitope. Detection of specific proteins/polypeptides in denaturing gels such as Westerns or sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) often employs antibodies raised against polypeptides and

which are usually continuous epitopes. However, many of these antibodies will not recognize the same antigen in tissue sections and may actually have unexpected reactivity with unrelated molecules. In tissue sections, diagnostically useful antibodies often recognize spatially related conformational epitopes that are lost when proteins denature. Most primary antibodies for diagnostic immunohistochemistry are murine monoclonals, but several monoclonal antibodies recognizing different epitopes on the same antigen can be combined into a cocktail.

For the detection of the antigen–antibody complex, hetero-antisera are raised against species-specific epitopes on the immunoglobulin protein. The oldest immunopathological method was to visualize the antigen using a secondary antibody labelled with a fluorescent dye that emits visible light after exposure to ultraviolet (UV) radiation. Although this method is suitable for virtually any antigen, it requires a specialized microscope and the immunostain is not permanent. More useful for tumour pathology are the detection systems in which a bridging antibody links the primary antibody to an antibody that reacts with a enzyme that precipitates a chromogen at the site such as the peroxidase–antiperoxidase (PAP) system (**Figure 1**). Another common detection system is the avidin–biotin peroxidase complex (ABC). Here the biotinylated secondary antibody links the primary antibody to a large preformed complex consisting of avidin, biotin and peroxidase. Large lattice-like complexes are formed so that several active peroxidase molecules are linked to each biotin binding site increasing the sensitivity of the system (**Figure 2**). With this system, endogenous biotin must be blocked or nonspecific background staining can be high, particularly in kidney, stomach or liver. The visual detection of specific antigen–antibody complexes in tissue sections depends on many factors including the nature of the antigenic determinants recognized, how well they survived tissue fixation and processing, the type of antibody used and the sensitivity of the detection system. There are now numerous commercial systems available for automation of immunohistochemical staining, but careful attention to technical details is still necessary for optimal results.

Loss of antigenicity during the fixation process is a major source of false negatives, but with the development of antigen retrieval methods, this is less of a problem than

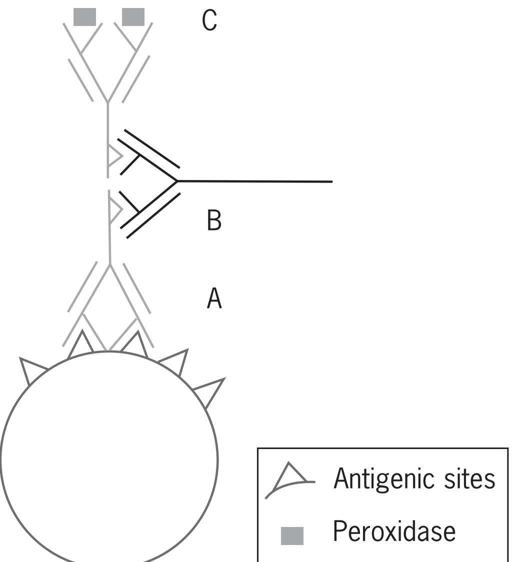


Figure 1 A secondary antibody (B) is directed against the immunoglobulins of the species producing the primary antibody which interacts with the antigen (A) and the antibody of the enzyme immune complex (C). The secondary antibody must be added in excess so that the FAB portions can bind to both antibodies. The substrate for the peroxidase enzyme – a chromogen (often 3,3′-diaminobenzidine tetrahydrochloride (DAB) – is added and an insoluble coloured precipitate is deposited at the site of the complex. This is referred to as the peroxidase-antiperoxidase (PAP) method.

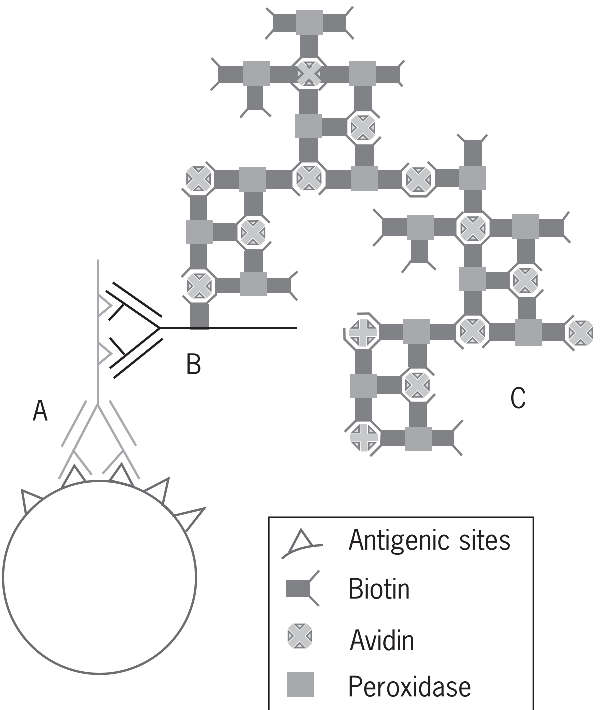


Figure 2 The avidin–biotin complex (ABC) method consists of a primary antibody (A) that interacts with the antigen. Its immunoglobulin portion is recognized by a biotinylated secondary antibody (B). A preformed avidin–biotin–peroxidase complex (C) binds to the biotin on the linked antibody. The peroxidase enzyme reacts with a chromogen (such as DAB) and a visual precipitate is produced. The high affinity of avidin for biotin (dissociation constant of 10^{-19}) contributes to the sensitivity of this method.

it once was. Enzymatic digestion of the tissue with proteinases was first found to be helpful in detecting a few antigens in formalin fixed tissue, but heat-induced epitope retrieval (HIER) has been shown to be more effective for a wider spectrum of antigens. The mechanism by which exposure of formalin-fixed sections in an aqueous medium to high temperatures reconstitutes epitopes is only partially understood. Fixatives such as formalin probably cross-link adjacent proteins in association with calcium. This cage-like complex is disrupted by heat and hydrolysis, aided by calcium chelation and/or precipitation by agents in the buffers. During the cooling phase, antigenic sites can be reformed. Small solubilized proteins that sterically inhibit the binding of antibodies may also be released during the process. A variety of means have been utilized to supply heat, including microwave, hot water baths, autoclaves, pressure cookers and rice steamers. Other variables include the temperature, duration of heating, pH, concentration and the buffers (citrate, borate, EDTA and several proprietary solutions). The exact conditions to release a specific antigenic site are variable and must be empirically determined for each antibody. However, using antigen retrieval methods, a wide variety of antibodies that were once restricted to analysis of only frozen sections are now determined in routinely processed formalin-fixed material.

When interpreting immunostains, it is critical that the pattern of reactivity be consistent with the known distribution of the cellular antigen. Antibodies to the intermediate filaments should give a fibrillary cytoplasmic staining pattern and if the staining is granular or muddy, or is present only in the Golgi region, artifactual staining should be suspected. S-100 is relatively unique in that it has both nuclear and diffusely cytoplasmic staining. p53 and the hormone receptors oestrogen and progesterone should be nuclear. Many of the lymphoma markers are membrane specific, such as CD45 and CD20. A common source of false-positive errors is the failure to recognize an aberrant cellular distribution pattern.

Both positive and negative controls are critical in preventing interpretation errors in histological material. Known positive and negative tissue samples must be analysed with every run, and if at all possible, internal controls in the same tissue should also be scrutinized. Immunohistochemical studies have a number of potential problems with both false negatives and false positives that may lead to diagnostic errors. Tumours are heterogeneous and irregular expression is a major source of false negatives, especially on small biopsies that are not representative. It is becoming increasingly clear that tumours can have aberrant antigen expression and few if any markers are specific to one or to a small group of tumours. Sarcomas such as leiomyosarcoma can express keratins and for some sarcomas such as synovial sarcomas and epithelioid sarcomas keratin expression is a diagnostic feature. Vimentin expression in carcinomas is not at all unusual either and it is actually expected in some such as endometrial, renal and most carcinomas growing in body fluids. A few years ago CD99 was thought to be relatively specific for Ewing sarcoma/PNET, but reactivity is now appreciated in a wide variety of sarcomas and lymphoid malignancies. Although reactivity to the leucocyte common antigen (CD45) is characteristic of haemato-lymphoid cells, some lymphomas with plasmacytoid differentiation or large cell anaplastic lymphomas are negative. False positives due to aberrant expression of antigens can also be a source of errors, emphasizing the need to interpret immunohistochemistry results in the context of the clinical and morphological features.

Molecular Studies in Tumour Classification

At the present time, the majority of cancers can be routinely classified in a clinically useful manner by morphology with immunohistochemistry as an important accessory technique. The current classification schemes for tumours based on histogenesis and differentiation, however, have limitations and depend on subjective characteristics interpreted by individual pathologists. At least theoretically, a classification scheme based on molecular characteristics should be more objective, measurable and reproducible with the biological behaviour of an individual tumour better predicted. The application of molecular and genetic techniques to the classification schemes based on a comprehensive molecular profile of tumours is just beginning but will be a major thrust for translational research in the future.

Although studies on the molecular classification of tumours are in their infancy, perhaps the most reassuring finding has been that the molecular and/or cytogenetic alterations identified in research laboratories substantially agree with the traditional morphological classifications. This has been particularly true for mesenchymal and haematopoietic tumours that have specific cytogenetic and chromosomal translocations. For example, fatty tumours are among the more common mesenchymal tumours in adults and most show cytogenetic abnormalities that segregate them into categories similar to standard histopathology (**Table 5**) (Fletcher *et al*., 1996). Myxoid and round cell liposarcomas have been suspected to be closely related tumours and this is supported by the cytogenetic observation that both are characterized by the same translocation, t(12;16)(q13p11), involving a breakpoint in the *CHOP* gene on chromosome 12. Ring or giant marker chromosomes with consistent involvement of chromosome 12 are found in well-differentiated liposarcomas. The more aggressive dedifferentiated liposarcomas that mimic malignant fibrous histiocytoma are probably related, as they often have additional complex chromosomal alterations superimposed on ring and giant chromosomes. Benign fatty tumours also have specific molecular changes and the ordinary and innocuous lipoma has numerous cytogenetic abnormalities. These too tend to correlate well

Table 5 Cytogenetic and molecular alterations in fatty tumours

Histological type	Translocation	Genes
Myxoid liposarcoma	t(12;16)(q13;p11) t(12;22)(q13;p12) (rare)	TLS/CHOP EWS/CHOP
Round cell liposarcoma	t(12;16)	TLS/CHOP
Atypical lipomatous tumours	Ring and giant chromosomes, often with abnormalities of 12q13–q15	
Lipoma	Abnormalities of 12q, 6p,13q	HMGIC/LPP
Lipoblastoma	8q rearrangements	
Spindle cell and pleomorphic lipomas	Aberrations of 16q and/or 13q	
Angiolipoma	Normal	

Table 6 Translocations used in the diagnosis of sarcomas

Sarcoma	Translocation	Genes involved	Available molecular test
Ewing sarcoma	t(11;22)	Fli1/EWS	PCR, FISH
Clear cell sarcoma	t(12;22)	ATF1/EWS	PCR, FISH
Myxoid chondrosarcoma	t(9;22)	CHN/EWS	PCR, FISH
Desmoplastic round cell tumour	t(11;22)	WT1/EWS	PCR, FISH
Synovial sarcoma	t(X;18)	SSX1/SYT	PCR, FISH
Alveolar rhabdomyosarcoma	t(2;13) t(1:13)	PAX3/FKHR PAX7/FKHR	PCR

with the previous morphological classification schemes. Of the common types of lipomas only angiolipomas have a normal karyotype. The common fatty lipomas of adults often have rearrangements involving 12q13–15. The breakpoint on chromosome 12 involves a gene, *HMGIC*, that codes for a member of the high-mobility group of proteins. These are small, acidic, nonhistone chromatin associated proteins that bind to AT-rich regions of the DNA. They have no inherent transcriptional activity but function by altering the nuclear chromatin, probably through interactions with other proteins. This alters the DNA structure and facilitates the assembly of transcriptional complexes. Spindle cell/pleomorphic lipomas usually have abnormalities of chromosome 16q or 13q, validating their histological separation from common lipomas and atypical lipomatous tumours.

Characteristic translocations have also been found by cytogenetic and molecular methods in a number of sarcomas (Ladanyi and Bridge, 2000) that are concordant with morphological and immunohistological categorization (**Table 6**). In some cases, a better understanding of the molecular defects found in a tumour have led to more accurate recognition of these rare tumours. For example, identification of the characteristic translocation t(X;18) in poorly differentiated spindle cell lesions by FISH analysis helps to discriminate monophasic and poorly differentiated synovial sarcomas from other spindle cell sarcomas.

Unfortunately, none of the common epithelial tumours of adults have simple genetic or cytogenetic changes underlying the malignant transformation. These tumours often have multiple changes in multiple pathways with no simple association with the current histological classifications. Microarray technology permits the expression of thousands of genes to be analysed simultaneously in a single tumour. The beginning of a molecular nonmorphological classification of tumours is seen in the pioneering work of Alizadeh *et al*. (Alizadeh *et al*., 2000). Diffuse large cell lymphoma is a common type of non-Hodgkin lymphoma of adults and is clinically a heterogeneous group with some patients responding well to chemotherapy and others rapidly dying of their disease. Using a DNA array constructed primarily from B cell libraries, they demonstrated that two distinct subgroups could be identified by differential expression patterns. One group expressed genes characteristic of germinal centre B cells and had a 76% survival after 5 years whereas the other group had an expression pattern resembling activated B cells with only 16% survival after 5 years. Golub and colleagues (Golub *et al*., 1999) developed an expression-based microarray technology that discriminates and correctly classifies acute lymphoid and acute myeloid leukaemia without morphological evaluation. These are among the first but undoubtedly not the last examples of tumour classifications by gene expression profiles and the

application of this technology to the common epithelial malignancies is to be expected in the near future.

Tumour Grading

The grade of a cancer is a reflection of how closely the tumour cells resemble their normal counterparts both morphologically and functionally. Malignant cells in well differentiated cancers can usually be readily recognized as having a histiogenic relationship to their cell of origin, i.e., they look like the tissues from which they arose. In some cases, e.g. very well-differentiated follicular cancers of the thyroid, the cells cannot be recognized as malignant until they demonstrate invasion. At the other extreme, some very undifferentiated tumours can only be assigned a lineage after exhaustive immunohistochemical or molecular studies and occasionally the cell of origin is never identified. In general, poorly differentiated cancers have a more aggressive behaviour than well-differentiated tumours of the same histological type.

For most tumours there are specific morphological criteria for assignment of grade and these have been shown to correlate with clinical behaviour. By evaluating architecture, presence or absence of specialized structures, the degree of cellular anaplasia and mitotic activity, many tumours are graded I–III with grade I tumours well differentiated and grade III poorly differentiated. This rather simplistic approach conveniently separates those tumours that are expected to do better (grade I) from those that are more deadly (grade III) than the majority of tumours in the middle. However, for some organs, more formalized grading systems are widely used such as the combined histological grade (modified Bloom–Richardson criteria) for breast carcinomas and Gleason's grading system for prostate cancers. These assign a numerical value to specific features. For example with breast cancer, a value of 1–3 is assigned for each of the degrees of gland formation, the nuclear features of anaplasia and the mitotic activity. The sum of these values (between 3 and 9) determines the tumour grade. In the prostate, architectural features are graded between 1 and 5 and the sum of the predominate and secondary patterns gives a score. Both of these systems have been shown to be reasonably reproducible and to correlate strongly with prognosis. Such grading systems are necessary for comparing results in therapeutic studies.

Tumour Staging

It is essential to have a uniform standardized system for classifying the extent of disease (staging) in order to compare therapeutic intervention and estimate outcome. One of the most widely used staging systems to describe the anatomic extent of neoplastic diseases has been developed by The American Joint Committee on Cancer (AJCC) (American Joint Committee on Cancer, 1997) in cooperation with the TNM Committee of the International Union Against Cancer. For most organs, the size of the tumour at its primary site and/or the involvement of local structures describe the tumour topography (T). The presence and extent of regional lymph node involvement (N) and whether or not there is documented evidence of distant metastasis (M) indicate spread of the tumour. These variables are incorporated into a shorthand notation called the TNM staging system. For most anatomical sites there are at least four categories for T, three for N and two for M, so that for some purposes the TMN groupings are condensed into stages (0, I–IV). Carcinoma *in situ* at all sites is Stage 0 while metastasis to distant sites is Stage IV and Stages I–III indicate increasingly greater extent of tumour spread. Staging can be clinical (cTMN), determined prior to initial therapy from the physical examination and imaging studies. The pathologic stage (pTMN) is assigned with information obtained from the surgical pathologist's assessment of the initial surgical resection of the primary tumour and regional lymph nodes. If the tumour recurs after treatment, the patient can be assigned a recurrent tumour stage (rTMN) and finally if a post mortem examination is performed, an autopsy stage (aTMN) may be assigned.

For common epithelial malignancies of adults (breast, colon, lung, prostate), lymph node status is probably the single most important standard risk factor for recurrence and a few additional comments should be made about their assessment. Extensive lymph node dissection has been the norm for cancer surgery and is responsible for much of the morbidity of the operations. Particularly for breast cancer and melanoma, sentinel lymph node biopsy is emerging as a technique to separate patients who have clinically occult lymph node metastasis from those whose tumours have not spread, sparing the latter group from the complications associated with radical lymph node dissection.

The first lymph node to receive lymphatic drainage from a tumour bed is termed the 'sentinel lymph node' (although it is often two or three nodes) and should be the first lymphatic tissue to be colonized by metastatic carcinoma spreading by lymphatics. A few hours before surgery, dyes or weak radioactive compounds are injected into the tumour bed, enter the lymphatics and are transported to the draining lymph nodes. At surgery these are identified visually or with the use of a hand-held Geiger counter and are selectively removed. Usually a frozen section is done during surgery and if micrometastases are found, a full regional lymph node dissection is performed. If no tumour is seen during intraoperative examination, the tissue is subjected to an extensive histological and immunohistochemical study to find tiny (even single cell) metastasis. For epithelial tumours and melanoma, immunohistochemistry for keratins and S100, respectively, is often used for identification of metastatic foci that could be easily missed by morphological examination only. Identification of tumour cell DNA or mRNA by PCR-mediated

amplification can also be done but is still primarily a research technique reserved for patients on clinical protocols. If occult tumour is identified in these lymph nodes, the patient can then be taken back to the operating room for removal of the regional lymph nodes. For breast cancer, if the sentinel lymph node(s) is negative, there is less than a 2% chance that a full axillary dissection would reveal a positive node (reviewed by Beechey-Newman, 1998). The procedure, however, can be technically challenging for both the surgeons and the pathologist and considerable experience is necessary to achieve this success rate.

Tumour Markers and Prognostic Markers

In a broad sense, tumour markers are molecules or substances that are produced by tumours or in response to tumours that can be used for early detection, diagnosis, prediction of disease progression and monitoring response to therapy. Many tumour markers are substances normally produced by the tissue of origin such as prostate specific antigen (prostate) or α-fetoprotein (liver and germ cell tumours) and can be elevated in both benign and malignant processes. A great number of tumour markers have been proposed as clinically useful but only a few are in routine clinical practice (**Table 7**). Ultimately a tumour marker is routinely evaluated if the results alter therapy and this improves clinical outcome. One of the most useful prognostic indicators for breast cancer is expression of oestrogen and progesterone receptors. Not only is the expression a favourable prognostic marker, effective antioestrogenic therapy is available for women whose tumours express the receptors. In the early 1990s immunohistochemical detection of oestrogen and progesterone receptors in breast cancer essentially replaced the dextran-coated charcoal biochemical analysis. This eliminated the need for frozen tissue and now the status can be obtained for very small tumours or even in fine needle aspiration specimens.

Markers are most useful when linked to specific pharmacological therapy or other forms of intervention based on the results. Her2/neu amplification in breast cancer predicting responsiveness to therapy with Herceptin® has already been mentioned and is now also being evaluated in other cancers such as lung and prostate. Similarly for lymphomas that express CD20 antigen, antibody therapy with Rituximab® is effective. On an investigational basis,

Table 7 Commonly used tumour markers

Tumour marker	Elevated	Comments
α-Fetoprotein (AFP)	Germ cell tumours Hepatocellular carcinomas	Glycoprotein synthesized by yolk sac and fetal liver. Usually measured in serum but can be detected by immunohistochemistry in the tumour tissue
CA125	Ovarian Cancer	Glycoprotein is elevated in sera of ovarian cancer patients, but also in many benign and reactive processes affecting the peritoneal lining
Carcinoembryonic antigen (CEA)	Gastrointestinal tumours, lung cancers, some breast cancers	Family of cell-surface glycoprotein that is elevated in a wide variety of carcinomas
Calcitonin	Medullary carcinoma of the thyroid	Approximately 80–90% of medullary carcinomas of the thyroid secrete calcitonin and serum elevation can be used to identify individuals at risk in families with an inherited form of multiple endocrine neoplasias
Her2/neu	Breast cancer	Amplification of gene associated with response to Herceptin therapy
Human chorionic gonadotropin (HCG)	Trophoblastic tumours	Glycoprotein secreted by placenta and is elevated in germ cell tumours with a trophoblastic component. Usually measured in serum but can be detected by immunohistochemistry in the tumour tissue
n-Myc	Neuroblastoma	Amplification of the gene is a poor prognostic indicator
Prostate-specific antigen (PSA)	Prostate cancer, some hindgut carcinoids	Produced by epithelial cells of the prostate gland and can be elevated in prostate cancer as well as by some benign prostate conditions

many markers are associated with biological behaviour but they are not routinely assessed as the marker result does not lead to changes in therapy or outcome. Small blood vessel density has prognostic value in breast cancer, but it is not routinely evaluated. Drugs that block angiogenesis first have to be proved effective in clinical trials and then the method for assessment of neovascularization also has to be proved to be predictive of response to therapy.

To be useful in determining successful treatment of cancer, a marker must be elevated prior to the initiation of therapy and then has to fall to a normal range when the tumour is surgically removed or is shrunken by therapy. Any elevation in subsequent measurements is evidence that the tumour has returned. Measurement of serum α-fetoprotein and/or β-HCG has been used very successfully in this manner for assessment of response to chemotherapy in germ cell tumours. A rise in these markers after a patient has been rendered free of disease by surgery or chemotherapy is an indication for treatment regardless of whether the recurrence can be detected by radiographic or other means.

Many markers show good predictive value in univariant analysis, but each new prognostic marker must be evaluated against well-established markers using appropriate multivariant statistical analysis. Many have shown strong correlation with traditional prognostic factors such as tumour stage, lymph node status, histological type of tumour, mitotic activity and histological grade. For example, p53 mutations occur in about 50% of breast cancers and these appear to be more aggressive than those with normal *p53* genes. However, the predictive value was much weaker in multivariant studies as the tumours with p53 mutations tended to be hormone receptor negative and had a high nuclear grade with more lymph node metastasis (Barbareschi, 1996). It is relatively easy by routine histology to separate well-differentiated tumours with an excellent prognosis from the aggressive poorly differentiated tumours, but there remains a large group of women who have tumours with intermediate histological features. Markers are needed that can segregate these tumours into those that are likely to have a good outcome from those that will do badly, so that aggressive therapy is given only to patients who are likely to benefit and the rest of the patients can be spared the toxicity and unpleasant side effects. Although a number of markers such as p53 mutational status and DNA ploidy are often considered in the clinical decision-making process, none has been shown to predict outcome reliably enough to be the sole or even a major criterion for management decisions.

Therapy

Unfortunately, most cancers are currently still treated by the relatively crude methods aimed at removal or ablation of malignant cells by surgery, chemotherapy or radiation therapy. Hormonal therapy is of value for the few cancers that are responsive. However, despite the tremendous amount of research, numerous press releases and scientific publications, immunotherapy and gene therapy have had little impact on routine clinical cancer care. Both of these therapies are likely eventually to find a niche and will be helpful for treatment of restricted tumours, but neither is likely to revolutionize cancer therapy in the near future.

Molecular pharmacogenomics may improve cancer therapy by predicting disease response to specific drug regimens and in reducing side effects and toxicity. The customization of therapy based on molecular targets is exemplified by the treatment of patients with acute promyelocytic leukaemia with all-*trans*-retinoic acid. This form of acute myeloid leukaemia is characterized by a t(15;17) translocation that involves the retinoic acid receptor-α and treatment with *trans*-retinoic acid induces differentiation with the majority of patients cured of their leukaemia (Slack and Rusiniak, 2000). Analysis of patient's enzyme systems or genes for drug metabolism and detoxification as well as tumour-specific factors such as the presence or absence of multiple drug-resistant genes may lead to improvements in selection of active drugs, their dosage and timing of therapy. However, the molecular foundations for determination of a tumour's resistance or sensitivity to specific chemotherapies is still not yet understood on either an empirical or a scientific basis and unfortunately there are still few, if any, applications in routine cancer care. (See the section on *The Treatment of Human Cancer*.)

MOLECULAR DIAGNOSTICS

The first use of molecular techniques in diagnostic pathology was the determination of clonality of haematolymphoid proliferations and this is still a potent tool for the diagnosis of lymphomas. The immunoglobulin genes and T cell receptor genes undergo unique somatic rearrangements during lymphocyte maturation. In reactive processes there are expansions of many clones of lymphocytes but in neoplastic proliferations the tumour is descendent from a single lymphoid cell. Originally, Southern blot analysis was used to detect these gene rearrangements, but now robust polymerase chain reaction (PCR) assays are available. Many lymphomas and leukaemias have specific characteristic translocations that can be identified by cytogenetics and/or molecular methods (Medeiros and Carr, 1999; Willman, 1999). **Table 8** lists some of the more common molecular abnormalities used in the diagnosis of lymphomas and leukaemias.

Identification of specific translocations by molecular tests in some of the sarcomas can also be diagnostically useful, especially in cases that are difficult to classify on morphological criteria, have aberrant immunohistochemical profiles or that occur at unusual sites (**Table 6**). These

Table 8 Some common molecular abnormalities of lymphomas and leukaemias

Lymphoma/leukaemia	Molecular defect	Genes involved
Follicular lymphoma	t(14;18)	IgH/bcl-2
Mantle cell lymphoma	t(11;14)	CCND-1 (bcl-1)/IgH
Burkitt lymphoma	t(8;14)	c-myc/IgH
Anaplastic large cell lymphoma	t(2;5)	npm/alk
Lymphoplasmacytoid lymphoma	t(9;14)	pax-5/IgH
B cell chronic lymphocytic lymphoma	t(14;19)	IgH/bcl-3
Acute lymphoblastic leukaemia		
Pre B	t(1;19)	E2A/PBX-1
B-cell	t(9;22) and others	BCR/ABL
Acute myeloid leukaemia		
M0 AML with minimal differentiation	−5/del, −7/del, +8	
M1 AML without differentiation	−5/del, −7 del, t(3;3), t(9;22)	
M2 AML with differentiation	t(8;21)	AML/ETO
M3 acute promyelocytic leukaemia	t(15;17)	PML-Rar-α
M4 acute myelomonocytic leukaemia	inv (16)	CBDF-β
M5 acute monocytic leukaemia	t(9;11)	MLL/AF9

are for the most part rare tumours and most are reliably diagnosed using morphological and immunohistological criteria, so that molecular testing is rarely required.

The difficulty in obtaining relatively pure tumour samples for analysis has hindered the widespread application of molecular techniques to solid tumours. Unlike the lymphomas and many sarcomas that grow as relatively pure tumour masses, for most epithelial cancers there is an intimate intermingling of tumour cells with stromal tissue and inflammatory cells. Separation of the malignant cell population by laser-capture microdissection is one method for the isolation of a relatively pure tumour cell population for study (Simone *et al*., 1998). However, this technique still depends on the morphological recognition of the malignant cells and is very labour intensive, which will limit its use in routine diagnostic practice.

The inherent increased cost of sophisticated technology is justified if it can reduce total disease management by eliminating unnecessary diagnostic procedures and ineffective treatment. However, billing and reimbursement issues as well as the costs of required licensure are major obstacles for implementation of the current technology in many hospital settings. In addition, any one assay is required for clinical care relatively infrequently and the expense of training and maintaining personnel and equipment in a small laboratory is prohibitive. For most institutions, selected cases that have specific indications for molecular testing will be referred to a reference laboratory.

MINIMAL RESIDUAL DISEASE (MRD)

Failure of primary surgical treatment of cancer is usually attributed to the presence of undetected metastatic disease (minimal residual disease). The goal of adjuvant chemotherapy or radiation therapy is the elimination of occult micrometastatic disease before it becomes clinically evident. For most oncology patients, the decision to add systemic therapy to surgical treatment is based on the staging parameters of the patient's tumours (tumour size, lymph node status and, to some extent, the histological type and grade of the tumour). A statistical analysis of accumulated data from similar patients with the same type and extent of disease is in most cases used to assess the likelihood that the disease will recur and whether additional therapy is given or withheld. However, conventional histological and clinical staging systems are limited. For example, with breast cancer patients who have small tumours and no lymph node metastasis, there is still a 15–25% chance of distant metastasis and no current histological or molecular marker identifies this population. This is far from ideal and there is a real need for markers that will unequivocally segregate patients into those whose tumours have already metastasized from those whose tumours have not metastasized. If a patient does not have metastatic disease, surgical removal is curative and cytotoxic therapy and all its negative side effects can then be reserved for those patients who may actually benefit.

Bone marrow is an accessible and frequent site for tumour metastasis and it is becoming increasingly common to examine bone marrow for the presence of metastatic tumour cells from patients with solid tumours. Immunohistochemical detection of micrometastasis is currently the standard method for the detection of the disseminated tumour cells. The identification of minimal residual disease depends on recognizing specific characteristics of the tumour cells that are not found in the surrounding tissue to which it has presumably metastasized. For example,

haematopoietic cells rarely express epithelial antigens such as keratins and occult epithelial tumour cells in the bone marrow can be assessed using immunohistochemistry for cytokeratins. As few as one or two tumour cells in 10^6 bone marrow cells can be detected by cytokeratin immunohistochemistry and enrichment techniques can increase that sensitivity by a factor of 10 (Osborne *et al.*, 1991).

Increasingly, molecular methods are used to assess the presence or absence of micrometastatic disease. PCR amplification of tumour-specific abnormalities present either in the DNA or mRNA of the tumour cells or, alternatively, amplification of tissue-specific mRNA can identify individual patients who have metastatic disease at the time of diagnosis and are most or least likely to benefit from the added therapy. In haematological malignancies, the identification of tumour-specific abnormalities at metastatic sites has been very useful for the detection of MRD. For example, primers can be made that span the t(14;18) translocation in the majority of follicular lymphomas, and amplification occurs only if DNA with the translocation is present in the sample. A few solid tumours, primarily sarcomas, have similar characteristic translocations, and similar methods have been used to identify MRD in patient with Ewing sarcoma or alveolar rhabdomyosarcoma.

Unfortunately, the genetic abnormalities are much more heterogeneous in the common malignancies of adults such as breast, prostate, lung and colon cancers. However, these malignant cells continue to express markers characteristic of their cell of origin and identification of a tissue-specific mRNA can be used to recognize MRD. For example, even very poorly differentiated metastatic prostate cancer to lymph nodes often express PSA mRNA, sometimes even when no protein is detectable by antigenic methods. Detection of micrometastasis by molecular methods is likely to have a profound impact on how cancer patients are treated and monitored in the future. For further information, the reader is referred to a recent review (Ghossein *et al.*, 1999).

The process of metastasis is poorly understood, but malignant tumour cells leave the primary site and gain access to the lymphatic or vascular system. After transportation to a new organ they adhere to the walls of the vessels, transverse them, implant and eventually grow. Even less understood is the observation that malignant cells can become dormant for months, years or even decades at a distant site only to wake up and renew growth with a vengeance. In this dormant or nondividing state, malignant cells are unlikely to be killed by cytotoxic chemotherapy directed against replicating cells. Understanding and developing strategies either to eliminate the dormant tumour cells by novel immunotherapies or to prevent their regrowth by inhibiting angiogenesis are just two means under intense investigation.

THE FUTURE

Most of us who have spent our professional lives diagnosing cancer recognize that we are at the beginning of a revolution that is not just going to have a technical impact but that is also fundamentally challenging our traditional concepts of tumour classification by histogenesis and morphology. It is likely that the H&E section will be the cornerstone of diagnosis for the near future, but it is being and increasingly will be challenged. There will be reevaluation and redefinition of many diagnostic entities as our knowledge and understanding of the underlying alterations in DNA and RNA expression becomes increasingly sophisticated. The transfer to clinical medicine of this new technology and the information it generates will not be quick, easy or painless and diagnostic anatomical pathology will be in a transition state for many years. It will take time to determine what information is therapeutically applicable and cost effective and how it should be obtained. Application of this knowledge to individual patients will still require physicians, whether traditional pathologists or specialists in the new molecular pathology, to exercise judgment and insight. However, in the end the hope is that the process will open up new areas of treatment options that can be individualized for each patient based on the inherent characteristics of the tumour.

REFERENCES

Alizadeh, A. A., *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, **403**, 503–511.

American Joint Committee on Cancer (1997). *AJCC Cancer Staging Manual*, 5th edn. (Lippincott Williams and Wilkins, Philadelphia, PA).

Barbareschi, M. (1996). Prognostic value of the immunohistochemical expression of p53 in breast carcinomas: a review of the literature involving over 9,000 patients. *Applied Immunohistochemistry*, **4**, 106–116.

Beechey-Newman, N. (1998). Sentinel node biopsy: a revolution in the surgical management of breast cancer? *Cancer Treatment Reviews*, **24**, 185–203.

Chu, P., *et al.* (2000). Cytokeratin 7 and cytokeratin 20 expression in epithelial neoplasms: a survey of 435 cases. *Modern Pathology*, **13**, 962–972.

Fletcher, C. D. M., *et al.* (1996). Correlation between clinicopathological features and karyotype in lipomatous tumors: a report of 178 cases from the chromosomes and morphology (CHAMP) collaborative study group. *American Journal of Pathology*, **148**, 623–630.

Ghossein, R. A., *et al.* (1999). Molecular detection of micrometastases and circulating tumor cells in solid tumors. *Clinical Cancer Research*, **5**, 1950–1960.

Golub, T. R., *et al*. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531–537.

Herrera, G. A., *et al*. (2000). Immunoelectron microscopy in the age of molecular pathology. *Applied Immunohistochemistry and Molecular Morphology*, **8**, 87–97.

Jaffe, E. S. (1999). Hematopathology: integration of morphologic features and biologic markers for diagnosis. *Modern Pathology*, **12**, 109–115.

King, L. S. (1967). What is a diagnosis? *Journal of the American Medical Association*, **202**, 154–157.

Ladanyi, M. and Bridge, J. A. (2000). Contribution of molecular genetic data to the classification of sarcomas. *Human Pathology*, **31**, 532–538.

Medeiros, L. J. and Carr, J. (1999). Overview of the role of molecular methods in the diagnosis of malignant lymphomas. *Archives of Pathology and Laboratory Medicine*, **123**, 1189–1207.

Miller, A. B., *et al*. (2000). Report on consensus conference on cervical cancer screening and management. *International Journal of Cancer*, **86**, 440–447.

Osborne, M. P., *et al*. (1991). Sensitivity of immunocytochemical detection of breast cancer cells in human bone marrow. *Cancer Research*, **15**, 2706–2709.

Rosai, J., *et al*. (1993). Standardized reporting of surgical pathology diagnoses for the major tumor types. *American Journal of Clinical Pathology*, **100**, 240–255.

Simone, N. L., *et al*. (1998). Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends in Genetics*, **14**, 272–276.

Slack, J. L. and Rusiniak, M. E. (2000). Current issues in the management of acute promyelocytic leukemia. *Annals of Hematology*, **79**, 227–238.

Suster, S., *et al*. (guest eds) (2000). Immunohistochemistry in tumor diagnosis. *Seminars in Diagnostic Pathology*, **17**, 169–256.

Willman, C. L. (1999). Acute leukemias: a paradigm for the integration of new technologies in diagnosis and classification. *Modern Pathology*, **12**, 218–228.

FURTHER READING

Cancer Screening

Smith, R. A., *et al*. (2000). American Cancer Society guidelines for the early detection of cancer. *Cancer*, **50**, 34–49.

Cuzick, J. (1999). Screening for cancer: future potential. *European Journal Cancer*, **35**, 685–692.

Diagnostic Surgical Pathology

Silverberg, S. G. (ed.) (1997). *Principles and Practice of Surgical Pathology and Cytopathology*, 3rd edn (Churchill Livingston, New York).

Classification of Cancer

Kleihues, P. and Sobin, L. H. (series eds) (2000). *World Health Organization Classification of Tumours* (IARC Press, Lyon).

Rosai, J. and Sobin, L. H. (eds) (2000). *Atlas of Tumor Pathology, Third Series* (American Registry of Pathology, Washington, DC).

Histotechnology

Prophet, E. B., *et al.* (eds) (1992). *Laboratory Methods in Histotechnology* (American Registry of Pathology, Washington, DC).

Sheehan, D. C. and Hrapchak, B. B. (1980). *Theory and Practice of Histotechnology*, 2nd edn. (C. V. Mosby, St. Louis, MO).

Immunohistochemistry

Taylor, C. R. (1986). *Immunomicroscopy: A Diagnostic Tool for the Surgical Pathologist* (W. B. Saunders, Philadelphia, PA).

Molecular pathology

Hammond, M. E., *et al*. (2000). College of American Pathologists Conference XXXV: solid tumor prognostic factors – which, how and so what? Summary document and recommendations for implementation. Cancer Committee and Conference Participants. *Archives of Pathology and Laboratory Medicine,* **124**, 958–965.

Wilczynski, S. P. (2001). Molecular biology. In: Weidner, N., *et al.* (eds), *Modern Surgical Pathology* (W. B. Saunders, Philadelphia, PA).

Websites

http://www.oncolink.upenn.edu. Oncology site at the University of Pennsylvania. Has a large number of links for information on cancers at specific sites and also for general oncology. Designed for both the medical professional and patients.

http://www-medlib.med.utah.edu/WebPath/webpath.html. Large number of digital images for basic pathology and neoplastic diseases. It also has mini-tutorials on specific topics.

http://edcenter.med.cornell.edu/CUMC_PathNotes/Neoplasia/Neoplasia. Descriptions of the pathology, aetiology and clinical aspects of a variety of neoplastic processes with links to images and further information.

Chapter 34

Skin

Vania P. Rudolf and George F. Murphy
Jefferson Medical College, Philadelphia, PA, USA

CONTENTS

- Normal Development and Stucture
- Tumour Pathology
- Aetiology
- Screening and Prevention
- Gross and Microscopic Pathology
- Molecular Genetic Findings
- Prognostic Factors
- Overview of Clinical Management
- Acknowledgement

NORMAL DEVELOPMENT AND STRUCTURE

The human integument is an extraordinarily versatile organ composed of tissue that is in a state of constant self-renewal and differentiation. In adults, the skin is the largest and most massive organ of the body, accounting for a 1.2–2.3 m^2 of surface, 16% of body weight and a plethora of protective functions. Human skin provides protection, sensation, thermoregulation, biochemical, metabolic, and immune functions. It is composed of tissue that grows, differentiates and renews itself constantly. The skin is a complex organ system and there are both benign and malignant tumours described for each and every component. The benign tumours usually present as isolated, relatively stable skin nodules; the malignant lesions, on the other hand, form a complete spectrum from slowly growing tumours, which tend to invade only locally, to aggressive, rapidly growing and metastasizing lesions.

In utero, the skin forms from ectoderm and mesoderm, giving rise to the most superficial epidermal layer and the underlying dermal layer, respectively. Appendages results from down growths of specialized epidermal cells, which differentiate to become hair follicles, eccrine sweat glands and apocrine sweat glands. The epidermis is composed primarily of a continuum of four layers which represent a maturation sequence which produces tough, flattened cells filled with the structural protein keratin at the skin surface (**Figure 1; see colour plate section**). The basal cell layer, a reservoir for proliferating cells, takes part in maintaining a rate of epidermal turnover that is sufficient to maintain an effective environmental interface. Basal cell carcinoma (BCC) is an example of a tumour composed of cells with morphological resemblance to basal keratinocytes and basaloid hair matrix epithelial cells. The spinous layer contains larger polyhedral cells and is a site for active synthesis of keratin proteins. Usually, squamous cell carcinoma (SCC) recapitulates the spinous layer. The uppermost granular and cornified layers, representing either nucleated cells with keratohyaline granules or anucleated cells, are sites of terminal differentiation of the epidermal cells. Therefore, tumours, like some squamous cell carcinomas, that form keratin in association with granular layer differentiation show primarily epidermal-type differentiation, e.g. invasive squamous cell carcinoma. In addition, the epidermis contains two minority populations of dendritic cells, one (melanocytes) producing brown photoprotective melanin pigment, and the other (Langerhans cells) partly responsible for immune surveillance against environmental pathogens and proteins perceived as foreign. The dermis contains nutrient blood vessels, sensory and afferent nerve fibres, immune cells, dermal dendritic cells and fibroblasts which produce a tough leathery enveloping matrix termed collagen and more resilient elastin fibres which impart stretch and recoil upon exposure to mechanical stress. Tumours of the dermis may arise from any one of its cellular components e.g. the blood vessels, nerve cells, fibriblasts or smooth muscle. Usually, such tumours tend to be nonspecific cutaneous nodules and histological examination determines their origin. Metastatic skin cancer is relatively rare, but again diagnosis depends on histological analysis.

Specific anatomical features of skin structure and related function are described in **Figure 1**. As will be described below, understanding of the basic oncology of the skin depends heavily on appreciation of normal

structure, since most tumours recapitulate to varying degrees their normal cellular counterparts.

TUMOUR PATHOLOGY

Tumour pathology of the skin involves literally hundreds of benign and malignant neoplasms derived from epidermal cells (true epithelial cells, termed keratinocytes, and melanocytes) and dermal cells (endothelial cells, nerve cells, immune cells and fibroblasts). The primary goals of diagnostic tumour pathology are to utilize gross and microscopic examination (1) to classify neoplasms according to histogenesis, (2) to make some prediction concerning anticipated biological behaviour and (3) to assess adequacy of surgical treatment. Classification generally depends on how closely given neoplasms recapitulate their 'cell of origin.' For example, the two most common forms of epidermal malignancy, the basal cell carcinoma and the squamous cell carcinoma, show differentiation features that most resemble cells of the normal stratum basalis and stratum spinosum, respectively. Tumours anticipated to behave in a benign manner possess a number of features, including slow growth, architectural symmetry and cellular uniformity. By contrast, aggressive tumours (i.e. those prone to produce damage to normal structures or to seed distant sites through metastasis) often grow in a rapid and destructive manner, demonstrate asymmetry of architecture and show considerable variability in cell size and shape. Exceptions exist, however, and in some instances the ability to predict biological behaviour based on tumour pathology is limited based on currently available methods of analysis.

AETIOLOGY

There are numerous environmental factors responsible for the genesis of skin cancer. Predisposing factors for squamous cell carcinoma include exposure to ultraviolet (UV) radiation, industrial carcinogens (tars and oils), chronic, draining ulcers potentially giving rise to free radical formation, old burn scars, arsenical ingestion, ionizing radiation, and tobacco and betel nut chewing in the case of perioral and intraoral cancers. Sun exposure, particularly UVB, is a primary cause of most basal cell cancers, and both UVB and UVA have been implicated in the cause of melanomas. Products that may activate the carcinogenic process may be generated as a consequence of UV radiation absorbed by epidermal melanin pigment. Because the relevant cellular targets of such carcinogens may reside in the basal cell layer, which is less accessible to their diffusion, the upper layers of the skin may actually be protective (Moan *et al*., 1999). Prolonged exposure to UV radiation in the UVB part of the spectrum, 280–320 m, is a clearly established carcinogenic factor. Eye and hair colour, skin type and constitutive skin pigmentation also have a major role in cutaneous malignancies. Fair-skinned, blue-eyed individuals and lightly pigmented populations show a significantly higher number of cutaneous malignancies than more darkly pigmented populations, underscoring the importance of melanin pigment in protecting against UV radiation induced carcinogens. Individuals with darkly pigmented skin also selectively exhibit an increase in natural killer activity in response to irradiation with low-dose UVB, which could underlie at least partly their resistance to the development of photo dependent skin cancer (Matsuoka *et al*., 1999). It has been postulated, however, that whereas melanin granules discharged by melanocytes into the uppermost epidermal layers may be photoprotective, those within the melanocytes, which normally reside in the basal cell layer, may actually be carcinogenic upon absorption of UV radiation. Evidence to support this includes the observation that certain albino Africans, as compared with normally pigmented Africans, appear to have a relatively smaller risk of developing melanoma as compared with nonmelanoma skin cancer (Moan *et al*., 1999).

A corollary to understanding the roles effects of various environmental carcinogens on skin is the realization that breakdown in specific host defences may facilitate these processes. For example, defective melanin pigmentation or inability to develop photoprotective tanning, resulting in frequent sunburns, is associated with increased risk of developing skin cancer. Genetically determined predisposition to environmental carcinogens includes individuals with inborn errors in DNA repair after UV-induced damage, a condition termed xeroderma pigmentosum. Affected individuals cannot repair UV damage to the skin after exposure to UV radiation, and this leads to early development of cutaneous neoplasms – basal cell carcinoma, squamous cell carcinoma and malignant melanoma. Others suffer from gene mutations, which result in a tendency for dysregulated skin proliferation, resulting in numerous basal cell cancers at an early age (basal cell nevus or Gorlin syndrome). Exposure to ionizing radiation is an environmental factor that predisposes to cutaneous malignancy (X-irradiation, etc.). Immune deficiency may predispose to local defects in the body's ability to fend off deleterious mutations, thus producing an environment permissive to the genesis of skin cancer. Indeed, immune-suppressed populations experience higher rates of cutaneous malignancy than immunologically normal individuals. Moreover, certain viruses may incite altered cell proliferation in the setting of immune deficiency, as is the case of oncogenic human papilloma virus in the setting of renal transplantation and human herpes virus 8 in the setting of Kaposi sarcoma associated with acquired immunodeficiency syndrome (AIDS). In HIV infection, the risk of these cancers increases gradually, while the same risk increases fairly quickly among transplant patients (Mueller, 1999).

SCREENING AND PREVENTION

Environmental factors are potentially conspiring to produce increasing numbers of skin cancers, such as the progressive depletion of the photoprotective ozone layer by anthropogenic pollutants. Ozone depletion increases the amount of biologically harmful solar UV radiation that reaches the surface of the Earth, leading to an increased incidence of cutaneous neoplasms. Indeed, the incidence of malignant melanoma is increasing at an alarming rate, with an estimate of a 1 in 75 lifetime risk of developing this potentially lethal tumour for individuals born in 2000 (Brown and Nelson, 1999). The Caucasian population in Auckland, New Zealand, has the highest incidence of malignant melanoma in the world (Jones *et al*., 1999); Kauai, Hawaii, on the other hand has one of the highest melanoma rates recorded in the United States (Chuang *et al.,* 1999). Accordingly, skin cancer prevention has become increasingly focused on public education regarding avoidance of sun exposure and the use of effective topical creams that significantly block the relevant portion of the UV spectrum. Such measures are particularly important in lightly pigmented individuals, or in those where genetic background or immune status could predispose to cancer formation. Health care professionals and patients alike also have been targeted for educational programmes that enhance early detection of both melanoma and nonmelanoma skin cancers. This approach is particularly important, since even malignant melanoma is 100% curable if it is surgically removed before the development of vertical growth within the dermal layers has occurred. Monitoring of patients with multiple pigmented spots representing potential precursors for or markers of malignant change has also been facilitated by the use of sensitive computerized image analysis approaches for digitized mapping of potentially changing lesions over time.

GROSS AND MICROSCOPIC PATHOLOGY

Precursors of Skin Cancer: Dysplastic Moles and Actinic (Solar) Keratoses

Although some skin cancers appear to develop '*de novo*' in normal skin, many others are preceded by proliferations termed atypical or dysplastic. These lesions have already begun to lose normal responsiveness to control mechanisms that determine the order and uniformity typical of normal skin and benign proliferations. ln terms of malignant melanoma, the association of pigmented 'moles' (naevocellular naevi) with malignant melanoma was made over 160 years ago. However, it was not until 1978 that a genuine precursor of melanoma was described in detail. In that year, Clark and colleagues detailed the characteristics of lesions they termed 'BK' mole (a name derived from the first letters of the last names of the initial two families studied) (Clark *et al*., 1978).

Clinically, BK moles (today more often referred to as dysplastic naevi) are larger than most acquired naevi (often greater than 5 mm across) and may occur as hundreds of lesions on the body surface (**Figure 2a**, inset). They are flat to slightly raised, often with a 'pebbly' surface. Frequently they form target-like lesions, with a darker, raised centre and an irregular, flat periphery. In contrast to most benign moles, dysplastic naevi usually show variability in pigmentation (variegation) and borders that are irregular in contour. Unlike ordinary acquired moles, they have a tendency to occur both on non-sun-exposed and on sun-exposed body surfaces. Dysplastic naevi may occur in multiple members of families prone to the development of malignant melanoma (a condition therefore referred to as the heritable melanoma syndrome) (Reimer *et al*., 1978). Dysplastic naevi also occur as isolated lesions not associated with the heritable melanoma syndrome. In this more common situation, the risk of malignant change appears to be low. Transitions from dysplastic naevi to early melanoma have been documented clinically and histologically within a period as short as several weeks, although the majority of such lesions are relatively stable.

Upon microscopic examination (**Figure 2a–c**), dysplastic naevi consist of compound naevi with both architectural and cytological evidence of abnormal growth. Unlike ordinary moles, where the neoplastic melanocytes tend to be arranged in orderly, small, discrete theques or nests (**Figure 2a**), dysplastic naevus cells form nests that tend to be enlarged and exhibit coalescence (**Figure 2b** and **c**). As part of this process, single naevus cells begin to replace the normal basal cell layer of the epidermis, producing so-called lentiginous hyperplasia. Cellular atypia, consisting of irregular, often angulated, nuclear contours and obliteration of nuclear detail by DNA-rich nuclear contents (hyperchromasia), is frequently observed. Associated alterations in the superficial dermis consist of a sparse inflammatory infiltrate, loss of melanin pigment from presumably destroyed naevus cells, with uptake of this pigment by dermal macrophages (melanin pigment incontinence), and a peculiar linear scarring surrounding the epidermal ridge-like downgrowths that are involved by the naevus (**Figure 2c**). All of these features are of assistance in the histological recognition of a dysplastic naevus.

Several lines of evidence support the belief that some dysplastic naevi are precursors of malignant melanoma. In one study (Greene *et al*., 1985) it was shown that in a large number of families prone to the development of melanoma, over 5% of family members developed melanoma over an 8-year follow-up period. In addition, new melanomas occurred only in individuals with dysplastic naevi. From these and related studies, it has been concluded that the actuarial probability of persons with the dysplastic naevus syndrome developing melanoma is 56% at age 59.

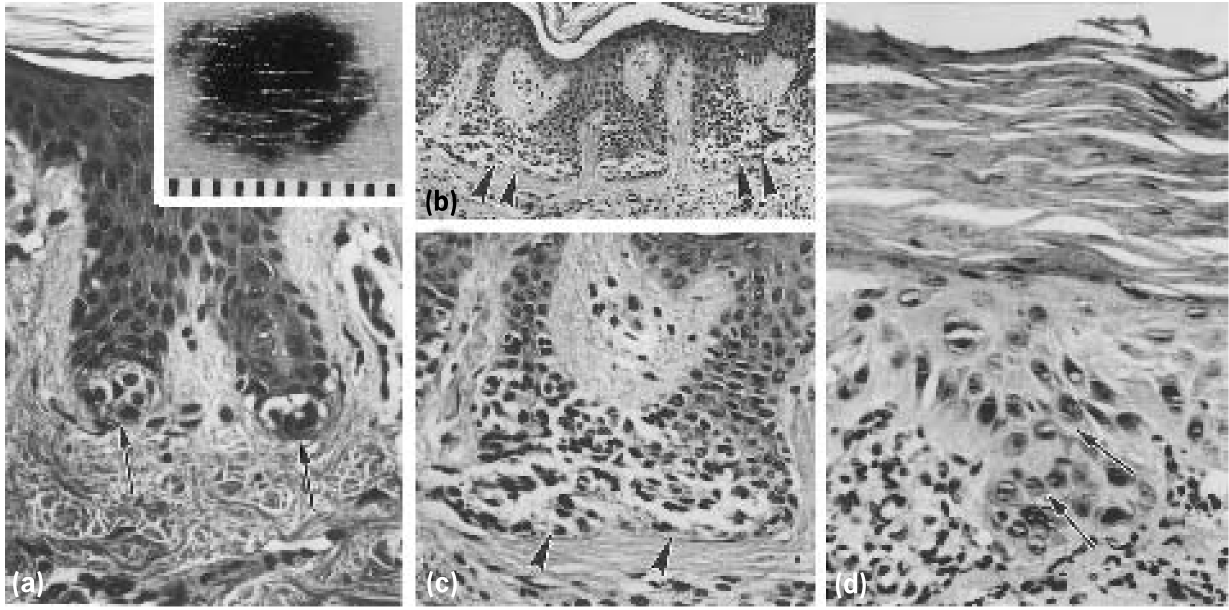


Figure 2 Premalignant dysplasia involving melanocytes and keratinocytes. (a) Unlike normal moles (also called melanocytic naevi) which grow as 'nests' of melanocytes at the tips of epidermal downgrowths called ridges (arrows), dysplastic naevi are characterized by abnormal patterns of growth, including readily recognized coalescence of adjacent melanocytic nests (b and c, arrowheads). The clinical result is a mole with irregular borders and non-uniform colouration (a, inset). Abnormal growth also is witnessed in a potentially premalignant dysplasia of epidermal keratinocytes termed an actinic keratosis (d). On comparison with **Figure 1a**, there is disordered growth of cells containing irregular nuclei within the basal and spinous cell layers (arrows), as well as abnormal formation of a thickened stratum corneum.

Further support of the relationship of melanoma to dysplastic naevi is the observation that the latter demonstrate expression of some abnormal cell surface antigens (Van Duinen *et al*., 1994), chromosomal abnormalities (Caporaso *et al*., 1987) and *in vitro* vulnerability to deleterious and mutagenic effects of UV radiation (Smith *et al*., 1987).

Specific steps have been proposed (Clark *et al*., 1985) whereby benign naevi may undergo aberrant differentiation to become dysplastic and eventually to evolve into melanoma. Parallels may be found in neoplasia involving other organ systems, such as uterine cervix, and thus dysplastic naevi are regarded by some as a paradigm for how certain malignant tumours develop from benign, albeit dysplastic proliferations of cells.

With regard to premalignant dysplasias of non-melanocytic epidermal cells, a series of progressively dysplastic changes typically occur in keratinocytes, a phenomenon analogous to the atypia that precedes carcinoma of the squamous mucosa of the uterine cervix. Excessive chronic exposure to sunlight can induce such premalignant lesions in the epidermis, causing histological changes in the normal keratinocyte maturation pattern and individual cell keratinization. Such a recognized condition is actinic keratosis, also termed solar keratosis. As would be expected, such lesions occur in particularly high incidence in lightly pigmented individuals. Exposure to ionizing radiation, hydrocarbons and arsenicals may induce lesions similar to or identical with actinic keratosis. Clinically, actinic keratoses are characterized by rough, scaling hyperpigmented plaques seen on the sun-exposed sites, especially the face, neck, upper trunk or extremities. Most lesions are usually less than 1 cm in diameter and may be tan–brown, red or skin-coloured. Some lesions may produce so much keratin that a 'cutaneous horn' develops, a phenomenon analogous to the formation of a true animal horn, which also originates from keratin-like protein. Because sites of predilection tend to be those prone to chronic UV damage, the lips may also develop similar lesions, and these are referred to as actinic cheilitis.

Cytologic atypia is seen in actinic keratosis in the lowermost layers of the epidermis and may be associated with increased numbers of basal cells (**Figure 2d**) or, alternatively, with thinning of the epidermis that results in a thin, semitransparent epidermal layer. The atypical basal cells usually contain enlarged, darkly stained nuclei. The superficial dermis contains thickened, blue–grey elastic fibres, a result of abnormal dermal elastic fibre synthesis by sun-damaged fibroblasts (Thielmann *et al*., 1987). The stratum corneum is thickened and, unlike in normal skin, nuclei in the cells in this layer are often retained, a pattern termed 'parakeratosis' (**Figure 2d**).

It appears that not all actinic keratoses (perhaps the majority) do not evolve into skin cancers, and some may actually remain stable or disappear due to immune responses. However, enough do become malignant to warrant conservative local eradication of these potential precursor lesions. This can usually be accomplished by gentle scraping, freezing or topical application of chemotherapeutic agents, which destroy the dysplastic cells. Radical or extensive surgery is not warranted.

Squamous Cell Carcinoma

Squamous cell carcinoma is the most common tumour arising on chronically sun-exposed sites. As such, it is typically detected on facial, extremity and trunk skin of older people. Except for lesions that develop on the lower legs, these tumours have a higher incidence in men than in women. Industrial carcinogens (tars and oils), chronic ulcers and draining osteomyelitis, old burn scars, ingestion of arsenicals, ionizing radiation and in the oral cavity tobacco and betel nut chewing, immunosuppression and genetic factors (xeroderma pigmentosum) are all implicated in the pathogenesis of squamous cell carcinoma, in addition to exposure to sunlight. In the case of tumours induced by sunlight exposure, endogenous melanin pigment is believed to provide an important photoprotective effect. UV radiation is believed to damage DNA by directly forming photoadducts, which contribute to cell mutation and altered proliferation. Normally, many such altered cells may be eliminated by immunosurveillance mechanisms before tumours become clinically apparent. Accordingly, albinos with defective melanin pigment production, patients with xeroderma pigmentosum, an inborn enzyme defect in DNA repair, and those with immunosuppression all tend to have a high incidence of this neoplasm.

Histological variants of SCCs include various well- and poorly differentiated tumours, which may mimic various nonsquamous neoplasms (Murphy and Elder, 1991c). Immunohistochemistry may assist in evaluating antigenic parameters, useful to determine the cell lineage of such tumours. This is important, because poorly differentiated squamous cell carcinomas may occasionally be confused with malignant melanomas, certain lymphomas or even mesenchymal neoplasms. For example, a panel of antibodies to determine the presence of keratin protein (carcinoma), leucocyte common antigen (LCA) (haematopoietic cells), S-100 protein (neural and melanocytic cells) and desmin/vimentin (dermal spindle cell tumours) may be helpful in diagnosing poorly differentiated variants of squamous cell carcinoma.

Squamous cell carcinomas that have not invaded through the basement membrane of the dermoepidermal junction (carcinoma *in situ*) clinically appear as sharply defined, red scaling plaques. More advanced, invasive lesions are nodular, show variable keratin production appreciated clinically as hyperkeratosis and may ulcerate (**Figure 3c**). Well-differentiated lesions may be indistinguishable from keratoacanthoma, a potential variant form with a tendency to regress spontaneously, possibly as a consequence of immune-mediated mechanisms. When the mucosa is involved by squamous cell carcinoma, a zone of white thickening is seen, an appearance caused by a variety of disorders and referred to clinically as leucoplakia.

Unlike the potentially premalignant dysplasia, actinic keratoses (see above), squamous cell carcinoma *in situ* is characterized by cells with atypical (enlarged, angulated, and darkly stained) nuclei completely replacing all layers of the epidermis (**Figure 3a and b**). When over time these cells acquire the ability to break through the basement membrane and enter into the underlying dermis, the process has become invasive. Invasive squamous cell carcinoma (**Figure 3d and e**) exhibits variable differentiation, ranging from tumours formed by polygonal squamous cells resembling those of the stratum spinosum arranged in orderly lobules and exhibiting numerous large zones of keratinization, to neoplasms formed by highly anaplastic, rounded cells with foci of necrosis and only abortive, single-cell keratinization (keratin ‘pearl’ formation).

Unlike some advanced melanomas (see below), invasive squamous cell carcinomas are usually discovered at stages where complete resection results in permanent disease eradication (North *et al.*, 1997; Rowe *et al.*, 1992). Up to 5%, however, may metastasize to regional nodes, particularly deeply invasive tumours that involve skin of the head and neck (Friedman *et al.*, 1985).

As indicated above, the most commonly accepted exogenous cause of squamous cell carcinoma is exposure to UV radiation light with subsequent DNA damage and associated mutagenicity. Individuals who are immunosuppressed as a result of chemotherapy or organ transplantation, or who have xeroderma pigmentosum, are at increased risk of developing malignant skin neoplasms (Penn, 1987). A considerable proportion of these tumours are squamous cell carcinomas, implicating aberrations in local immune networks in the skin in the production of an atmosphere permissive to neoplasia. Interestingly, sunlight, in addition to its effect on DNA, also seems to have a direct and at least a transient immunosuppressive effect on skin by influencing the normal surveillance function of antigen-presenting Langerhans cells in the epidermis (Cooper *et al*., 1985). For example, in experimental animals, it now appears that although Langerhans cells responsible for T lymphocyte activation are injured by UV radiation, similar cells responsible for the selective induction of suppressor lymphocyte pathways are resistant to UV damage (Granstein *et al*., 1987). Moreover, local exposure of the skin to UV radiation may also result in alterations in systemic immunity. Such phenomena could result in both systemic and local imbalances in T cell function that would favour tumour genesis and progression. DNA sequences of certain viruses

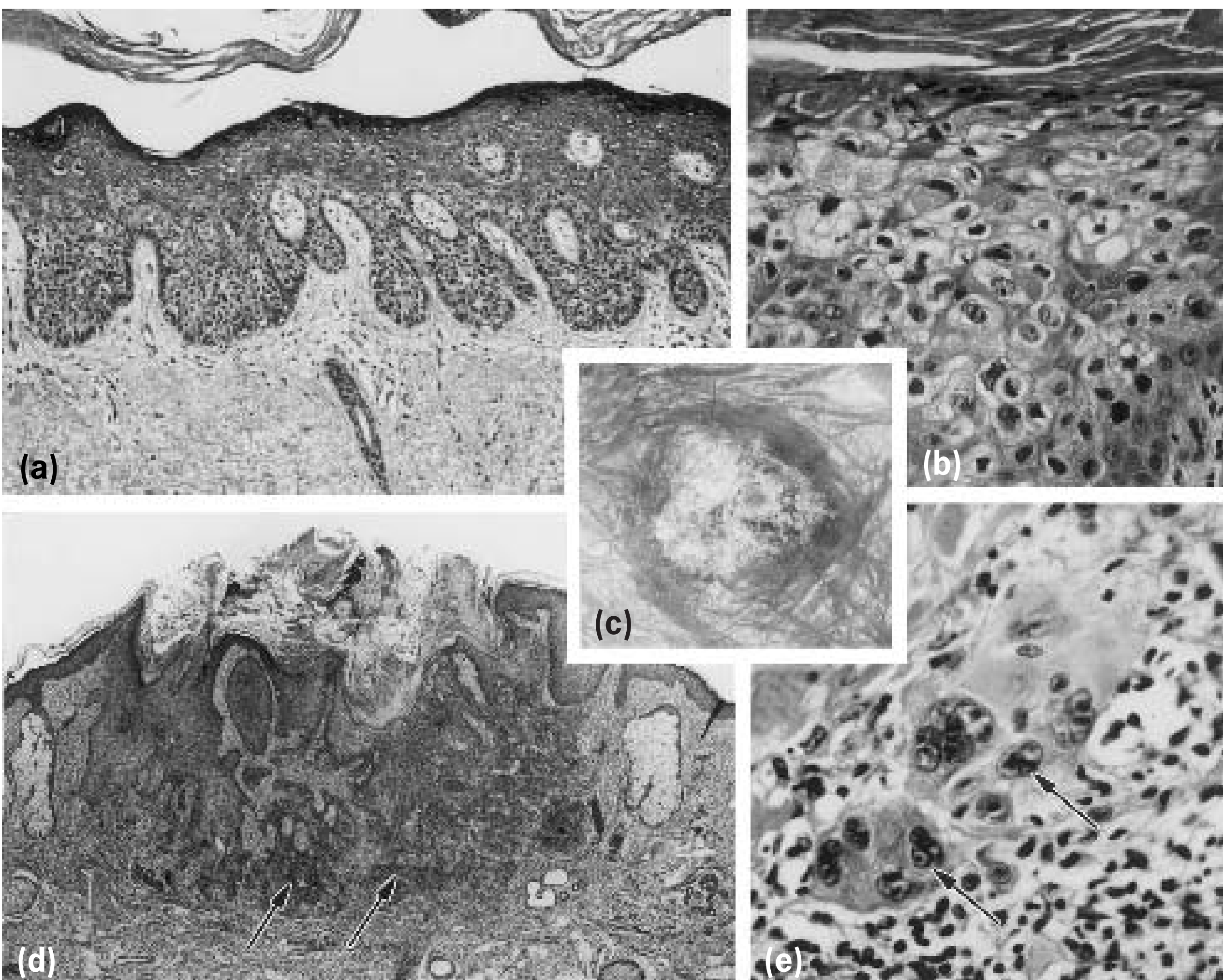


Figure 3 Squamous cell carcinoma. (a) and (b) malignant cells have not as yet invaded into the dermis, but have replaced most of the epidermal layer, a stage referred to as carcinoma *in situ*. (c) The clinical appearance of squamous cell carcinoma is characterized by a scaling, sometimes ulcerated nodule. (d) and (e) invasive squamous cell carcinoma differs from its earlier, *in situ* stage, by showing downward (invasive) growth of malignant cells into the dermis (arrows). These cells (e) have cytological features of malignancy, including enlarged, darkly staining, irregular nuclei (arrows).

(e.g. human papillomavirus HPV36) have been detected in DNA extracted from potential precursors of squamous cell carcinoma (Kawashima *et al*., 1986). This suggests a role for these agents in the causation of these cutaneous epithelial neoplasms. Finally, certain chemical agents appear to have direct mutagenic effects on epidermal keratinocytes by producing DNA adducts with subsequent oncogene activation (Hochwalt *et al*., 1988; Perez *et al*., 1997).

Basal Cell Carcinoma

Basal cell carcinoma by far is the most frequent of all cutaneous cancers. Like squamous cell carcinoma, basal cell carcinomas are slow-growing tumours that rarely metastasize. They also have a tendency to occur at sites subject to chronic sun exposure and in lightly pigmented people. Other predisposing factors include ionizing radiation, chronic scarring and arsenical exposure. As with squamous cell carcinoma, the incidence of basal cell carcinoma increases sharply with immunosuppression and in patients with inherited defects in DNA replication or repair (xeroderm pigmentosum). The rare, dominantly inherited basal cell naevus syndrome (Bale, 1997) is associated with the development of numerous basal cell carcinomas in early life, with abnormalities of bone, nervous system, eyes and reproductive organs, and with a specific gene mutation which now has been characterized.

Clinically, basal cell carcinomas present as smooth-surfaced, pearly papules often containing prominent, dilated subepidermal blood vessels termed telangiectasias

(**Figure 4b**). Some may contain melanin pigment and, as a consequence, may appear clinically similar to pigmented moles or melanomas. Chronic, large lesions may ulcerate (rodent ulcers), and extensive local invasion of bone, facial sinuses and deep subcutaneous nerves may occur after many years of neglect (Dixon *et al.*, 1989).

By light microscopy, tumour cells typically resemble those in the normal basal cell layer of the epidermis. Different histological variants of basal cell carcinoma can be distinguished according to the tumour growth pattern (e.g. nodular, sclerosing, cystic, pigmented). The earliest buds of basal cell carcinoma arise from the epidermis or follicular epithelium and do not characteristically occur on mucosal surfaces. Two patterns of growth include either multifocal origins from the basal cell layer of the epidermis and extending over several square centimeters or more of skin surface (**Figure 4a**), or nodular invasive lesions growing downward deeply into the dermis as cords and islands of darkly stained cells embedded in fibrotic matrix often with a pale blue hue as a consequence of associated mucin deposition (**Figure 4c–e**). Occasional tumours may appear cystic clinically and are found to contain large cystic spaces upon histological examination (**Figure 4d**). The cells at the periphery of the tumour cell cords and islands tend to be arranged radially with their long axes in approximately parallel alignment, a characteristic feature also seen in the bulb of hair follicles and referred to as 'palisading.' The stroma shrinks away from the epithelial tumour nests (**Figure 4d** and **e**), creating clefts or 'separation artifacts' that assist in differentiating basal cell carcinomas from certain follicular appendage tumours also characterized by proliferation of basaloid cells (e.g. trichoepithelioma).

Mutations in the tumour-suppressor gene termed '*PATCHED*' (*PTC*) have been detected in human patients with the basal cell naevus syndrome. In this condition, multiple basal cell carcinomas as well as other anomalies develop according to an inherited pattern. Moreover, mice genetically engineered to overexpress the attachment site or ligand for PTC molecule, which is termed 'sonic hedgehog' or SHH and which mimics the loss of PTC function, develop basal cell carcinomas (Oro *et al.*, 1997). Thus, both SHH and PTC may have important roles in human skin carcinogenesis in the basal cell naevus

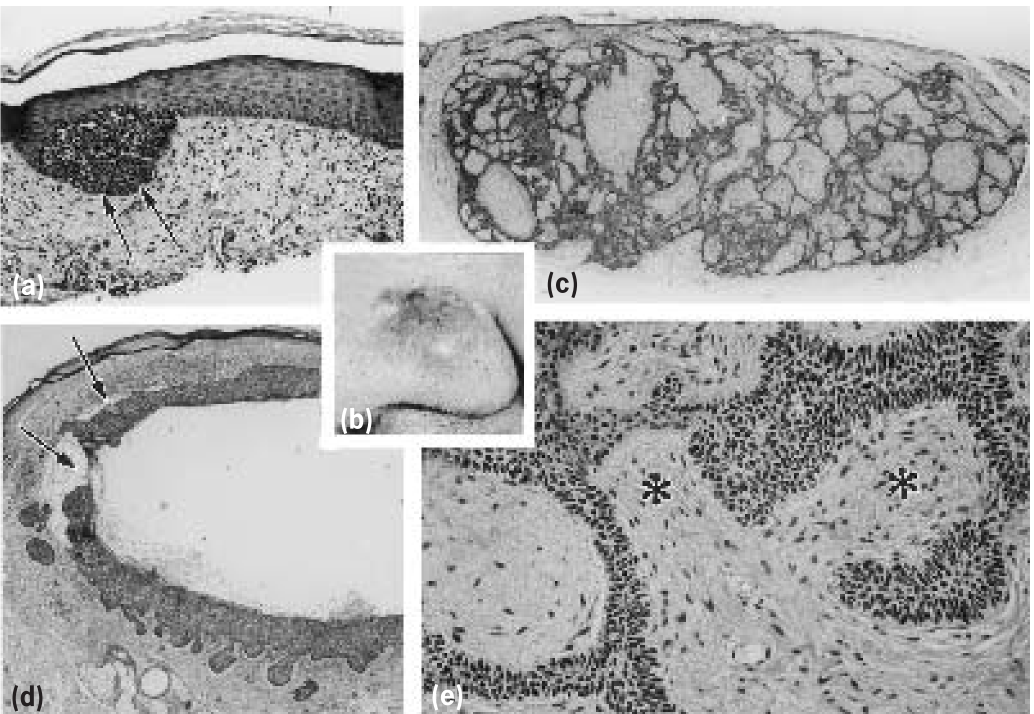


Figure 4 Basal cell carcinoma. (a) Tumours initially are formed by bud-like growth of darkly stained, 'basaloid' cells (arrows), here seemingly arising from the basal cell layer of the epidermis. (b) The clinical appearance of many basal cell carcinomas is that of a smooth-surfaced, 'pearly' nodule containing prominent superficial blood vessels. (c) Some nodular basal cell carcinomas may form an elaborate network of branching and anastomosing basaloid cells, while others (d) may have cystic spaces occupying their centres. (e) The cells forming these common tumours most resemble the epidermal basal cell layer and the basaloid cells that give rise to the hair follicle (matrix cells). The connective tissue that envelops the tumour cells (asterisks) resembles that of the bulb portion of the hair follicle, providing evidence that basal cell carcinoma is a tumour of follicular differentiation. Arrows in d indicate cleft-like spaces of artifactual stromal-epithelial separation.

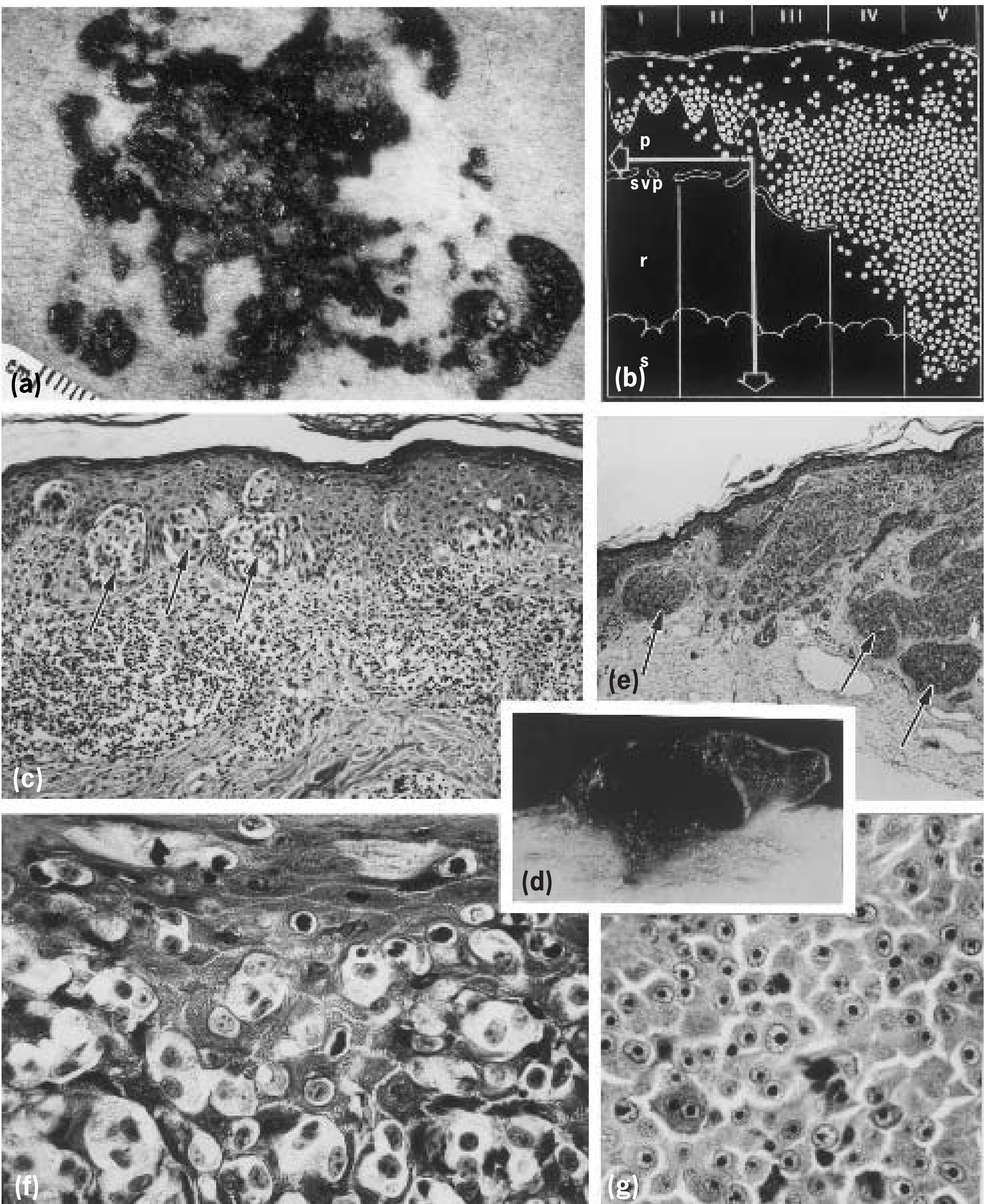


Figure 5 Malignant melanoma. (a) Clinically, melanoma is characterized by a diameter often greater than 1 cm, an irregular border and variation in colour. (b) Melanoma typically grows initially within the epidermal layer and as single cells within the superficial (papillary (p)) dermis, which is directly above an anatomic boundary called the superficial vascular plexus (svp). This early growth is generally curable with surgical excision and is termed the radial phase of growth (indicated by horizontal arrow). If untreated, the melanoma in time grows downward into the deeper (reticular (r)) dermal and subcutaneous (s) layers, an event termed vertical growth (signified by a vertical arrow). This stage carries with it the risk of metastatic spread which correlates with the depth and microanatomical levels of invasion, as

(*Continued*)

syndrome, although their precise roles in the majority of basal cell carcinomas unassociated with genetic patterns of inheritance remain to be elucidated.

Malignant Melanoma

Malignant melanoma is a relatively common neoplasm that as little as several decades ago was considered to be a potentially deadly form of skin cancer. The great preponderance of melanomas arise in the skin, although other sites of origin include the oral and anogenital mucosal surfaces, oesophagus, meninges and eye. All melanomas have the potential for metastasis if left untreated, and therefore the conventional modifier 'malignant' is really unnecessary. As a result of increased public awareness of the earliest signs of cutaneous melanomas, most are cured surgically in the early stages of their biological evolution (Mihm, 1971). Nevertheless, the incidence of malignant melanoma is on the rise, necessitating vigorous surveillance for its development.

As with epithelial malignancies of the skin (see above), sunlight appears to play an important role in the development of cutaneous melanoma. For example, men commonly develop this tumour on the upper back, whereas women tend have a relatively high incidence of melanoma on both the back and the legs. Lightly pigmented individuals are at higher risk for the development of melanoma than darkly pigmented individuals, an observation supported by epidemiological studies examining tumour incidence in equatorial regions versus those characterized by high latitudes. Sunlight, however, does not seem to be the sole predisposing factor for the development of melanoma, and the presence of a pre-existing naevus (e.g. a dysplastic naevus), hereditary factors or even exposure to certain carcinogens (as in the case of experimental melanomas in rodent models) may play a role in lesion development and evolution (Barnhill *et al*., 1993).

Clinically, malignant melanoma of the skin is usually asymptomatic, although itching and focal bleeding may be early manifestations. The most important clinical sign of possible dysplastic or malignant degeneration is change in colour in a pigmented lesion. Unlike benign (nondysplastic) naevi, melanomas usually exhibit striking variations in pigmentation, appearing in shades of black, brown, red, dark blue and grey. Occasionally, regions of white or flesh-coloured hypopigmentation are also present. The borders of melanomas are not smooth, round and uniform, as in naevocellular naevi; rather, they tend to be irregular and often 'notched' (**Figure 5a**). In conclusion, the clinical warning signs of melanoma are (1) enlargement of a pre-existing mole; (2) itching or pain in a pre-existing mole; (3) development of a new pigmented lesion during adult life; (4) irregularity of the borders of a pigmented lesion; and (5) variegation of colour within a pigmented lesion.

Crucial to understanding the complicated histology of malignant melanoma is the concept of radial and vertical growth. Simply described, radial growth indicates the tendency of a melanoma to grow horizontally within the epidermal and superficial dermal layers, often for a prolonged period of time (**Figure 5b**). During this stage of growth, melanoma cells do not have the capacity to metastasize and clinical cure is frequent. Specific types of radial growth phase melanoma are lentigo maligna, superficial spreading and acral/mucosal lentiginous. These can be defined on the basis of architectural and cytological features of growth within the epidermal layer as well as biological behaviour (e.g. lentigo maligna type of radial growth usually occurs on sun-damaged facial skin of the elderly and may continue for as long as several decades before the tumour develops the capacity to metastasize). With time, the pattern of growth assumes a vertical component. The melanoma now extends downward into the deeper dermal layers as an expansile growing mass lacking cellular maturation, without a tendency for the cells to become smaller as they descend into the reticular dermis (**Figures 5d, 5e**). Clinically, this event is heralded by the development of a nodule in the relatively flat radial growth phase, and correlates with the emergence of a clone of cells with true metastatic potential. The probability of metastasis in such as lesion may be predicted by simply measuring in millimeters the depth of invasion of this vertical growth phase nodule below the granular cell layer of the overlying epidermis (Breslow, 1970). Recently, prediction of clinical outcome has been improved further by taking into account factors such as number of mitoses and degree of infiltrative lymphocytic response within the tumour nodule (Clark *et al*., 1989).

Individual melanoma cells generally are considerably larger than naevus cells. They contain large nuclei with irregular contours and have chromatin characteristically clumped at the periphery of the nuclear membrane with prominent red (eosinophilic) nucleoli (**Figures 5f, 5g**).

(**Figure 5** *caption continued*)
indicated by convention as levels I–V. (c) Histological presentation of an early radial growth melanoma, where most of the malignant cells are present as irregular nests within the epidermal layer (arrows). (d) and (e) gross pathological (cross-section) and low-magnification histological representations of a heavily pigmented melanoma that has progressed to the vertical stage of growth. Note in (e) the numerous large nests of darkly stained cells that have found their way into the deeper dermal layers (arrows). Melanoma cells, whether within the epidermis as part of early radial growth (f) or within the dermis as part of vertical growth (g), are distinguished from most naevus cells by their enlarged nuclei which generally contain conspicuous central nucleoli.

These tumour cells proliferate as poorly formed nests or as individual cells at all levels of the epidermis (**Figure 5b**) in the radial phase of growth and, in the dermis, as expansile, balloon-like nodules in the vertical phase of growth (**Figure 5e**). The nature and extent of the vertical growth phase determine the biological behaviour of malignant melanoma, therefore it is important to observe and record vertical growth phase parameters in a pathology report. (See the chapter on *Models for Melanoma and Sarcomas*.)

Mycosis Fungoides (Cutaneous T Cell Lymphoma)

Cutaneous T cell lymphoma (CTCL) represents a spectrum of lymphoproliferative disorders that affect the skin (Murphy and Mihm, 1999). Two types of malignant T cell disorders were originally recognized: mycosis fungoides (MF), which is a chronic proliferative process, and a nodular eruptive variant, mycosis fungoides d'emblée. It is now known that a whole variety of presentations of T cell lymphoma occur, including MF, the eruptive nodular type, and an adult T cell leukaemia or lymphoma type. The latter disorder may present with a rapid progressive downhill course.

Mycosis fungoides is a T cell lymphoproliferative disorder, arises primarily in the skin and may evolve into generalized lymphoma (Murphy, 1988). Most afflicted individuals have disease that remains localized to the skin for many years; a minority has rapid systemic dissemination. This condition may occur at any age, but most it affects persons over 40 years of age.

Clinically, lesions of the MF type of CTCL represent scaly, red–brown patches, raised, scaling plaques that may even be confused with psoriasis, and fungating nodules. Eczema-like lesions describe early stages of disease when obvious visceral or typical nodal spread has not occurred. Raised, indurated and irregularly outlined, erythaematous plaques may then supervene. Systemic spreading correlates with development of multiple, large (up to 10 cm or more in diameter), red–brown nodules. Sometimes plaques and nodules may ulcerate, as depicted in **Figure 6a**. Lesions may affect numerous body surfaces, such as the trunk, extremities, face and scalp. In some individuals, systemic spread and seeding of the blood by malignant T cells is accompanied by diffuse erythema and scaling of the entire body surface (erythroderma), a condition known as Sèzary syndrome.

Histologically, the identification of the Sèzary–Lutzner cells is the hallmark of CTCL of the mycosis fungoides type. These are T helper cells (CD4 antigen-positive), they form characteristically band-like aggregates within the superficial dermis (see **Figure 6b**) and invade the epidermis as single cells and small clusters (Pautrier microabscesses). The tumour cells have markedly infolded nuclear membranes, imparting a 'hyperconvoluted' or 'cerebriform' contour (see **Figures 6c, 6d**). Patches and plaques show pronounced epidermal infiltration by Sèzary–Lutzner cells (epidermotropism), but in more advanced nodular lesions the malignant T cells often lose this epidermotropic tendency, grow deeply into the dermis and eventually seed lymphatics and the peripheral circulation.

The aetiology of CTCL is under active investigation. The discovery that a highly aggressive form of T cell lymphoma or leukaemia in adults is caused by infection of helper T cells by a specific retrovirus (human T cell leukaemia virus, or HTLV-I) (Murphy, 1988) promotes the possibility that conventional CTCL may also have an infectious causation. The proliferating cells in CTCL are clonal populations of lymphocytes of the CD4 subset; they often express aberrant cell surface antigens as well as clonal T cell receptor gene rearrangements, and detection of these features may be of diagnostic assistance in difficult cases (Bakels *et al.*, 1997).

Topical therapy with steroids or UV radiation is often delegated for early lesions of CTCL, while more aggressive systemic chemotherapy is indicated for advanced disease. In patients with circulating malignant cells, exposure of removed from the body blood cells to photosensitizing agents and UVA irradiation, followed by reinfusion into the patient (extracorporeal photopheresis; Rook *et al.*, 1991), has shown promise as a novel therapeutic approach to disseminated CTCL. (See also chapter *Lymph Nodes*.)

Angiosarcoma

Angiosarcoma is a rare malignant tumour showing variable degrees of vascular differentiation. It occurs primarily on the face and scalp, frequently in elderly people. The lesion presents as single or multiple plaques, may progress to become nodules, infiltrate the adjacent soft tissues and ulcerate. Histological examination shows diffuse dermal hypercellularity and irregular vascular spaces formed by malignant endothelial cells. These atypical neoplastic cells have hyperchromatic nuclei with irregular contours. The malignant cells of better differentiated tumours recapitulate vascular spaces, bulging into the vessel lumens and forming intraluminal tufts. Poorly differentiated tumours, on the other hand, often provide sheets composed of solid aggregates of tumour cells (Murphy and Elder, 1991a). Angiosarcoma can develop as a result of previous irradiation and in lymphedematous extremities, such as those produced after surgery for breast cancer or as a result of previous radiation therapy.

An angioproliferative disorder, which may show findings similar to angiosarcoma, is Kaposi's sarcoma. This vascular proliferative disorder occurs as two variant forms: an indolent cutaneous variety, seen in elderly individuals and common in Ashkenazi Jews and Southern Europeans, and a second form, characterized by an aggressive tumour that presents as a component of immunodeficiency states,

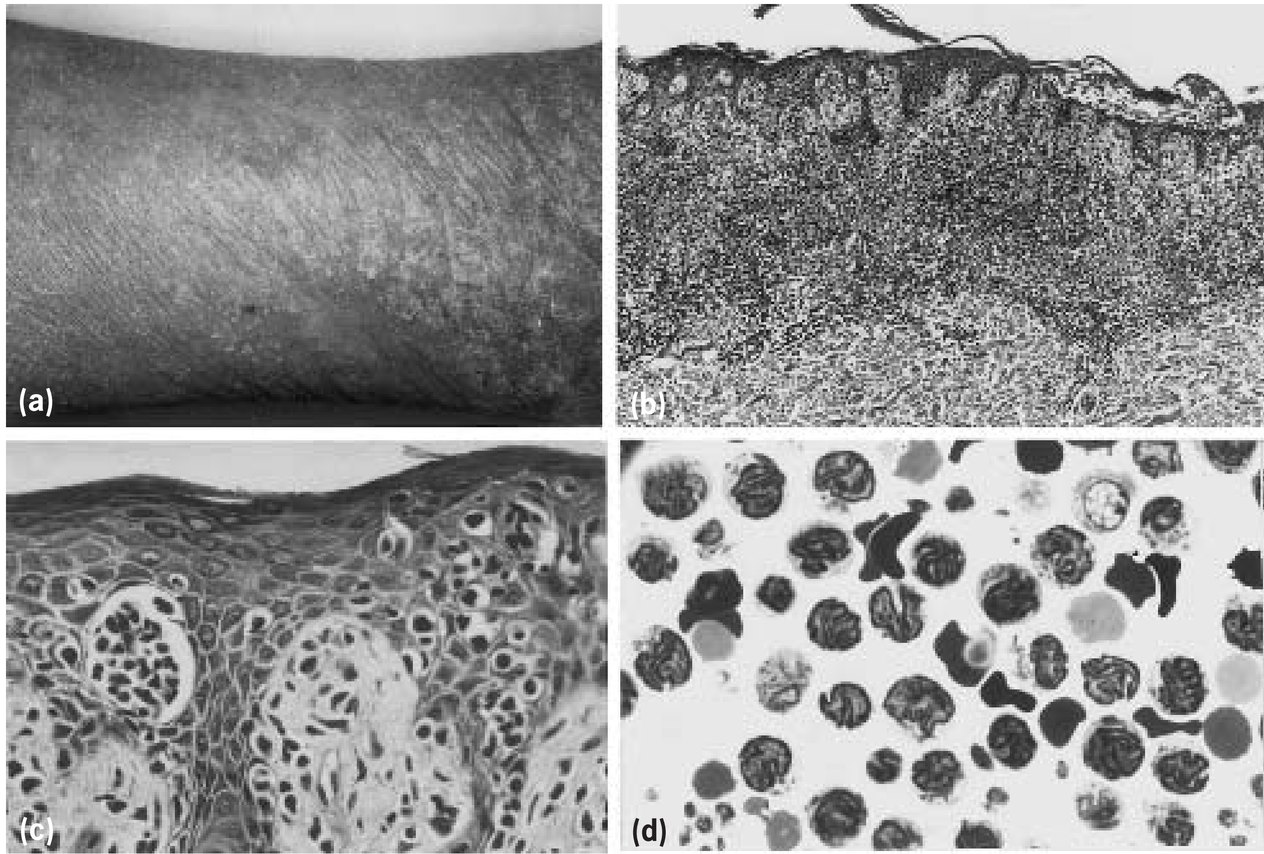


Figure 6 Mycosis fungoides (cutaneous T cell lymphoma). (a) Early lesions may present as pink–red, eczema- or psoriasis-like patches and plaques covered by adherent scales. (b) The biopsy at this stage may show a horizontal band of lymphocytes (many representing malignant T cells) within the upper dermis directly beneath the epidermal layer. (c) Inspection at higher magnification will often reveal large, darkly stained malignant lymphocytes abnormally collecting within the epidermis, where they tend to be found in association with Langerhans cells (see **Figure 1**). (d) Cytologically, the malignant T lymphocytes contain a nucleus with a characteristically infolded contour sometimes likened to the profile of a brain, and therefore referred to as 'cerebriform.'

especially in acquired immunodeficiency syndrome (AIDS). In this setting, its occurrence is associated with a new herpesvirus, human herpesvirus 8 (Murphy and Elder, 1991). Kaposi's sarcoma may begin as small solitary (isolated) purpuric papules, characterized by small zones of intracutaneous haemorrhage and comparable to bruises or petechiae (patch stage). Progression of Kaposi's sarcoma results in raised plaques and nodules, which are often large and destructive clinically. Advanced stage tumours frequently ulcerate. Histological examination during early phases demonstrates a multifocal angioproliferative process characterized by an increase in the number of spindle cells and thin-walled, poorly formed vessels within superficial vascular plexus, lymphoplasmocytic inflammatory infiltrates, and dilated lymphatics. Advanced lesions show expansive nodular growth of atypical spindle cells, containing hyperchromatic nuclei (Murphy and Elder, 1991). Such lesions may be confused with other malignancies of mesenchymal cells.

MOLECULAR GENETIC FINDINGS

In the case of the hereditary form of dysplastic naevi, genetic analyses have demonstrated the trait to be inherited as an autosomal dominant, possibly involving a susceptibility gene located on the short arm of chromosome 1, near the *Rh* locus (1p36) (Greene *et al.*, 1983). Other melanoma-susceptibility genes have been mapped to chromosomes 9p21 and 12q14 (Greene, 1997).

PROGNOSTIC FACTORS

The most common forms of skin cancer, namely squamous cell and basal cell carcinoma, are slowly growing, locally invasive, and tend not to exhibit metastatic behaviour. As mentioned earlier, some squamous cell carcinomas, particularly those that involve facial skin, may metastasize to

local lymph nodes. Basal cell carcinomas rarely metastasize, but when they do, large tumour size tends to be a constant feature. Locally aggressive basal cell carcinomas often will show a sclerosing or diffusely infiltrative pattern of dermal invasion.

Malignant melanoma, in contrast, shows a well-defined ability to metastasize once it has entered the vertical phase of growth. Numerous parameters have been defined which correlate with prognosis for such lesions. Negative attributes are represented by axial (versus acral) site of occurrence, male gender, progressive depth of dermal invasion, high mitotic count, absence of tumour infiltrating lymphocytes at the base of the tumour nodule, presence of regression and documentation of vascular invasion and microscopic satellite formation (Elder and Murphy, 1991). In lesions diagnosed unequivocally as vertical growth phase melanomas, such prognostic variables may be charted on multivariate databases, permitting prediction of 8-year survival with confidence intervals that demonstrate considerable accuracy.

A major drawback in predicting prognosis of melanocytic tumours is represented by a recently recognized, relatively rare but genuine group of neoplasms that possess hybrid or intermediate characteristics between melanocytic naevi and fully evolved malignant melanoma. These tumours have been variously termed borderline melanomas, minimal deviation melanomas, naevoid melanomas, severely atypical melanocytic naevi, severely atypical dermal and epidermal melanocytic proliferations and melanocytic tumours of uncertain malignant potential (MELTUPs) (Elder and Murphy, 1991). Depending on the designation, such problematic, grey-zone tumours may possess some or all of either the architectural or cytological features of melanoma, but do not qualify as fully evolved malignant lesions, leaving the prediction of their biological behaviour an open question. Diagnostic classification of such tumours, even among experts in dermatopathology (Farmer *et al*., 1996), has proven to be extremely problematic. Accordingly, a pathologist has been astute to have identified such a lesion as at least representing significant melanocytic atypia. Often the best that one can do at this juncture is to relate concern regarding the hybrid nature of the histological features and to recommend complete excision, in addition to close clinical follow-up. Hopefully, as more of these problematic lesions are characterized, reproducible diagnostic criteria and prognostic attributes will be identified that will permit more accurate determination of the likelihood of eventual metastasis.

OVERVIEW OF CLINICAL MANAGEMENT

Most squamous and basal cell carcinomas may be adequately treated by complete eradication either by excision or by destructive means, such as curretage, freezing or electrodesiccation. Excision has the advantage of providing for histological analysis to ensure negative margins, although the alternative modalities of local destruction have proven to be highly effective (Fitzpatrick *et al*., 1999), especially when lesions are small and biologically nonaggressive. In situations where tumours are present in close proximity to vital structures, particularly involving facial skin, a modified Mohs technique may be desirable (Leslie and Greenway, 1991). In this setting, multiple frozen sections of marginal tissue are examined at the time of the excision. Lymph node dissection in the case of invasive squamous cell carcinoma may be indicated when there exists suspicious clinical enlargement of draining lymph node chains.

The role of wide excision and lymph node dissection in the setting of vertical growth phase melanoma is potentially problematic. Primary lesions are generally recommended to be excised with a margin that will minimize the likelihood of local recurrence. However, this recommendation is controversial, with some experts arguing that although wider excisions may decrease the incidence of local metastases, they fail to improve survival. This assumption is based on the contention that local metastases of melanoma generally correlate with contemporaneous seeding of distant sites (Ackerman and Scheiner, 1983). In recent years, sentinel lymph node sampling has gained acceptance in the treatment of many vertical growth phase melanomas. In this procedure, a radiographically detectable dye is injected at the site of the primary tumour. This dye is taken up by draining lymphatic vessels and delivered physiologically to regional lymph nodes. The first nodes to be detected by this method are presumed to be candidates for early metastases, and therefore they are surgically removed and studied extensively by histological and sometimes by immunohistochemical analysis for the presence of occult melanoma metastases. Occasionally, entire lymph node chains will be removed as part of a procedure termed prophylactic elective regional lymph node dissection. The clinical rationale for such procedures is to eradicate surgically early metastases while they are theoretically confined to draining lymph nodes. This approach remains to be validated by prospective data. Moreover, it is complicated by the theoretical possibility that latent systemic metastases may occur synchronously with nodal ones, and by the notion that nodal deposits of tumour cells may in some instances eventuate in positive effects in terms of sensitizing the immune system against specific tumour-associated antigens.

In this regard, adjuvant therapy for melanoma has recently focused on immunoenhancing therapies, such as melanoma vaccines and treatment with proinflammatory mediators, such as interferon-α. Although these approaches have provided encouraging preliminary data in the treatment of clinically advanced melanoma metastases (Murphy *et al*., 1993), their role as adjuvants to ameliorate the potential for eventual metastatic spread is not as yet fully defined and awaits further experimental validation.

ACKNOWLEDGEMENT

This work was supported in part by grant CA40358 (G.F.M.) from the National Cancer Institute.

REFERENCES

Ackerman, A. B. and Scheiner, A. M. (1983). How wide and deep is wide and deep enough? A critique of surgical practice in excisions of primary cutaneous malignant melanoma. *Human Pathology*, **14**, 743–744.

Bakels, V., *et al*. (1997). Immunophenotyping and gene rearrangement analysis provide additional criteria to differentiate between cutaneous T-cell lymphomas and pseudo-T-cell lymphomas. *American Journal of Pathology*, **150**, 1941–1949.

Bale, A. E. (1997). The nevoid basal cell carcinoma syndrome: genetics and mechanism of carcinogenesis. *Cancer Investigations*, **15**, 180–186.

Barnhill, R. L., *et al*. (1993). Neoplasms: malignant melanoma. In: Fitzpatrick, T. B., *et al*. (eds), *Fitzpatrick's Dermatology in General Medicine* (Mcgraw-Hill, New York).

Breslow, A. (1970). Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. *Annals of Surgery*, **172**, 902–908.

Brown, T. J. and Nelson, B. R. (1999). Malignant melanoma: a clinical review. *Cutis*, **63**, 275–278, 281–284.

Caporaso, N., *et al*. (1987). Cytogenetics in hereditary malignant melanoma and dysplastic nevus syndrome: is dysplastic nevus syndrome a chromosome instability disorder? *Cancer Genetics and Cytogenetics*, **24**, 299–314.

Chuang, T. Y., *et al*. (1999). Melanoma in Kauai, Hawaii, 1981–1990: the significance of in situ melanoma and the incidence trend. *International Journal of Dermatology*, **38**, 101–107.

Clark, W. H. Jr, *et al*. (1978). Origin of familial malignant melanomas from heritable melanotic lesions: the BK mole syndrome. *Archives of Dermatology,* **114**, 732–738.

Clark, W. H., *et al*. (1985). A study of tumor progression: the precursor lesions of superficial spreading and nodular melanoma. *Human Pathology*, **15**, 1147–1165.

Clark, W. H., Jr, *et al*. (1989). Model predicting survival in stage I melanoma based on tumor progression. *Journal of the National Cancer Institute*, **81**, 1893–1904.

Cooper, K. D., *et al*. (1985). Effects of ultraviolet radiation on human epidermal cell alloantigen presentation: initial depression of Langerhans cell-dependent function is followed by the appearance of T6 Dr^+ cells that enhance epidermal alloantigen presentation. *Journal of Immunology*, **134**, 129–137.

Dixon, A. Y., *et al*. (1989). Factors predictive of recurrence of basal cell carcinoma. *American Journal of Dermatopathology*, **11**, 222–232.

Elder, D. E. and Murphy, G. F. (1991a). Malignant melanoma. Fascicle 2, In: *Atlas of Tumor Pathology*, Fascicle 2, 3rd edn. 154–163 (Armed Forces Institute of Pathology, Washington, DC).

Farmer, E. R., *et al*. (1996). Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. *Human Pathology*, **27**, 528–531.

Fitzpatrick, T. B., *et al*. (1999). Treatment of squamous and basal cell carcinoma, squamous cell carcinoma and basal cell carcinoma. In: Fitzpatrick, T. B., *et al*. (eds), *Fitzpatrick's Dermatology in General Medicine*, (5th edn) 852–862 (McGraw-Hill, New York).

Friedman, H. I., *et al*. (1985). Prognostic and therapeutic use of microstaging of cutaneous squamous cell carcinoma of the trunk and extremities. *Cancer*, **56**, 1099–1105.

Granstein, R. D., *et al*. (1987). Epidermal cells in activation of suppressor lymphocytes: further characterization. *Journal of Immunology*, **138**, 4055–4062.

Greene, M. H. (1997). Genetics of cutaneous melanoma and nevi [review]. *Mayo Clinic Proceedings*, **72**, 467–474.

Greene, M. H., *et al*. (1983). Familial cutaneous malignant melanoma: autosomal dominant trait possibly linked to the Rh flocus. *Proceedings of the National Academy of Sciences of the USA*, **80**, 6071–6075.

Greene, M. H., *et al*. (1985). The high risk of melanoma in melanoma prone families with dysplastic nevi. *Annals of the Internal Medicine*, **102**, 458–465.

Hochwalt, A. E., *et al*. (1988). Mechanism of H-ras oncogene activation in mouse squamous carcinoma induced by an alkylating agent. *Cancer Research*, **48**, 556–558.

Jones, W. O., *et al*. (1999). Incidence of malignant melanoma in Auckland, New Zeland: highest rates in the world. *World Journal of Surgery*, **23**, 732–735.

Kawashima, M., *et al*. (1986). Characterization of a new type of human papillomavirus (HPV) related to HPV5 from a case of actinic keratosis. *Virology,* **154**, 389–394.

Leslie, D. F. and Greenway, H. T. (1991). Mohs micrographic surgery for skin cancer. *Australasian Journal of Dermatology*, **32**, 159–164.

Matsuoka, L. Y., *et al*. (1999). Immunological responses to ultraviolet B radiation in black individuals. *Life Sciences*, **64**, 1563–1569.

Mihm, M. C. (1971). The clinical diagnosis, classification and histogenetic concepts of the early stages of cutaneous malignant melanomas. *New England Journal of Medicine*, **284**, 1078–1082.

Moan, J., *et al*. (1999). Epidemiological support for a hypothesis for melanoma induction indicating a role for UVA radiation. *Photochemistry and Photobiology*, **70**, 243–247.

Mueller, N. (1999). Overview of the epidemiology of malignancy in immune deficiency. *Journal of Acquired Immune Deficiency Syndrome*, **21**, 5–10.

Murphy, G. F. (1988). Cutaneous T cell lymphoma. In: Fenoglio-Preiser, C. M. (ed.), *Advances in Pathology*. 131–148 (Year Book Medical Publishers, Chicago).

Murphy, G. F. and Elder, D. E. (1991a). Angiosarcoma. In: *Atlas of Tumor Pathology*, (Fascicle 1, 3rd edn), 211–214 (Armed Forces Institute of Pathology, Washington, DC).

Murphy, G. F. and Elder, D. E. (1991b). Kaposi's sarcoma. In: *Atlas of Tumor Pathology*, Fascicle 1, 3rd edn, 214–219 (Armed Forces Institute of Pathology, Washington, DC).

Murphy, G. F. and Elder, D. (1991c). Non-melanocytic tumors of the skin. In: *Atlas of Tumor Pathology*, Fascicle 1, 3rd edn, 61–154 (Armed Forces Institute of Pathology, Washington, DC).

Murphy, G. F. and Mihm, M. C., Jr (1999). Mycosis fungoides (cutaneous T-cell lymphoma), the skin. In: Cotnar, R. S., Kumar, V. and Robbins, S. L. (eds), *Robbins' Pathologic Basis of Diseases*. 1190 (Saunders, Philadelphia).

Murphy, G. F., *et al*. (1993). Autologous melanoma vaccine induces inflammatory responses in melanoma metastases: relevance to immune regression and immunotherapy. *Journal of Investigative Dermatology*, **100**, 335–341.

North, J. H., Jr, *et al*. (1997). Advanced cutaneous squamous cell carcinoma of the trunk and extremity: analysis of prognostic factors. *Journal of Surgical Oncology*, **64**, 212–217.

Oro, A. E., *et al*. (1997). Basal cell carcinoma in mice over-expressing sonic hedgehog. *Science* **276**, 817–821.

Penn, I. (1987). Neoplastic consequences of transplantation and chemotherapy. *Cancer Detection and Prevention*, **1**, 149–157.

Perez, *et al*. (1997). P53 oncoprotein expression and gene mutations in some keratoacanthomas. *Archives of Dermatology*, **133**, 189.

Reimer, R. R., *et al*. (1978). Precursor lesions in familial melanoma: a new genetic preneoplastic syndrome. *Journal of the American Medical Association*, **239**, 744–746.

Rook, A. H., *et al*. (1991). Combined therapy of the Sezary syndrome with extra-corporeal photochemotherapy and low dose interferon alpha: clinical, molecular, and immunologic observations. *Archives of Dermatology*, **127**, 1535–1540.

Rowe, D. E., *et al*. (1992). Prognostic factors for local recurrence, metastasis, and survival rates in squamous cell carcinoma of the skin, ear, and lip. Implications for treatment modality selection. *Journal of the American Academy of Dermatology,* **26**, 976–990.

Smith, P. J., *et al*. (1987). Abnormal sensitivity to UV-radiation in cultured skin fibroblasts from patients with hereditary cutaneous malignant melanoma and dysplastic nevus syndrome. *International Journal of Cancer,* **30**, 39–45.

Starink, T. M., *et al*. (1985). The cutaneous pathology of Cowden's disease: new findings. *Journal of Cutaneous Pathology* **12**, 83–93.

Thielmann, H. W., *et al*. (1987). DNA repair synthesis in fibroblast strains from patients with actinic keratosis, squamous cell carcinoma, basal cell carcinoma, or malignant melanoma after treatment with ultraviolet light, *N*-acetoxy-2-acetyl-aminofluorene methyl methanesulfonate, and *N*-methyl-*N*-nitrosourea. *Journal of Cancer Research and Clinical Oncology*, **113**, 171–186.

Van Duinen, C. M., *et al*. (1994). The distribution of cellular adhesion molecules in pigmented skin lesions. *Cancer*, **73**, 2131–2139.

FURTHER READING

Bruce, A. J. and Brodland, D. G. (2000) Overview of skin cancer detection and prevention for the primary care physician. *Mayo Clinic Proceedings*, **75**, 491–500.

de Villiers, E. M., *et al.* (1999). Human papillomaviruses in non-melanoma skin cancer. *Seminars in Cancer Biology*, **9**, 413–422.

Hadshiew, I. M., *et al*. (2000). Skin aging and photoaging: the role of DNA damage and repair. *American Journal of Contact Dermatitis*, **11**, 19–25.

Hussain, S. P. and Harris, C. C. (2000). Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutation Research*, **462**, 311–322.

Leffell, D. J. (2000). The scientific basis of skin cancer. *Journal of the American Academy of Dermatology*, **42**, 18–22.

Murphy, G. F. and Mihm, M. C. (1999) Recognition and evaluation of cytologic dysplasia in acquired melanocytic nevi. *Human Pathology*, **30**, 506–512.

Tenkate, T. D. (1999) Occupational exposure to ultraviolet radiation: a health risk assessment. *Reviews on Environmental Health*, **14**, 187–209.

van Kranen, H. J. and de Gruijl, F. R. (1999). Mutations in cancer genes of UV-induced skin tumors of hairless mice. *Journal of Epidemiology*, **9**, S58–S65.

Woodhead, A. D., *et al.* (1999). Environmental factors in non-melanoma and melanoma skin cancer. *Journal of Epidemiology*, **9**, S102–S114.

Chapter 35

Oral Cavity and Major and Minor Salivary Glands

Paul L. Auclair
Maine Medical Center, Portland, ME, USA

Karen Rasmussen
Maine Center for Cancer Medicine, Scarborough, ME, USA

CONTENTS

- Normal Development and Structure
- Tumour Pathology: Squamous Cell Carcinoma
- Screening and Prevention: Gross/Histopathology/Preinvasive
- Tumour Pathology: Salivary Gland Carcinomas

NORMAL DEVELOPMENT AND STRUCTURE

During the third week of development of the human embryo, a deep groove develops below the forebrain that represents the future oral cavity. This groove is known as the primary oral fossa, or stomadeum, and it is lined by ectoderm that forms an additional ectodermal pouch, Rathke's pouch, that gives rise to the anterior lobe of the pituitary gland and to the dental organs that will form the enamel of the deciduous and permanent teeth. The stomadeum is separated from the endodermally lined foregut by the buccopharyngeal membrane that ruptures in the fourth week of development. In an adult the oral mucosa is covered by stratified squamous epithelium that is keratinized in areas exposed to masticatory forces, and is supported by fibrous tissue containing salivary glands, adipose tissue, skeletal muscle and bone. The red portion of the lips (vermilion) represents a transition zone from skin to mucosa but does not contain eccrine or sebaceous glands or hair follicles.

The salivary gland primordia develop as buds of primitive stomadeal epithelium that proliferate as strands into the underlying oral ectomesenchyme. This results in the formation of three pairs of major salivary glands the parotid, submandibular and sublingual glands and between 500 and 1000 lobules of minor salivary glands. Although the parotid is a single contiguous structure, the facial nerve courses through its centre, essentially dividing the gland into superficial (lateral) and deep (medial) lobes. The parotid gland is rich in lymphoid tissue, and normally contains from 3–24 lymph nodes, most in the superficial lobe (Ellis and Auclair, 1996). The minor glands are usually numerous in all oral mucosal sites except for the anterior hard palate and gingiva. The glands consist of secretory acini composed of mucous or serous acinar cells. Saliva then flows to the intercalated portion of the branching ductal system, lined by low cuboidal epithelium, then to the larger striated duct portion formed by columnar cells that have large numbers of mitochondria, and finally to the excretory ducts that are lined by stratified squamous epithelium where they merge with the oral epithelium. The acini are surrounded by myoepithelial cells that have ultrastructural features of both epithelium and smooth muscle.

TUMOUR PATHOLOGY: SQUAMOUS CELL CARCINOMA

Epidemiology

More than 90% of all cancers of the oral mucosa and lip vermilion are squamous cell carcinomas (epidermoid carcinoma). Although other forms of cancer, including sarcomas, melanomas and lymphomas, also affect the oral mucosa and jaws, only squamous cell carcinoma will be discussed in this section; adenocarcinomas that arise from the major and minor salivary glands will be discussed in the following section.

In the United States, squamous cell carcinomas of the oral mucosa and lip vermilion account for about 3% of all cancers of all sites, representing about 22 000 new cases each year. As shown in **Figure 1**, the incidence rates are strongly age-related. The death rate per 100 000 population

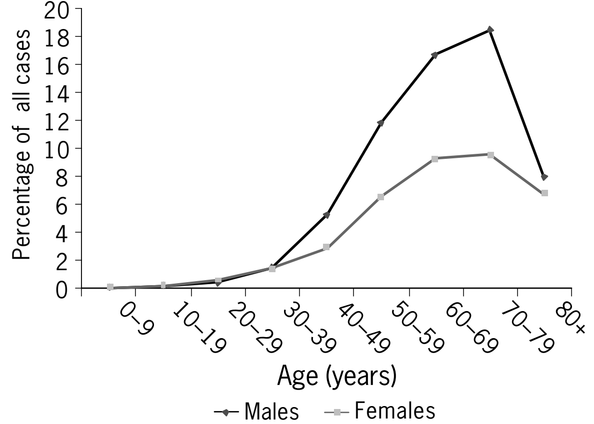


Figure 1 Age- and gender-specific incidence rates of oral and pharyngeal cancer. While the incidence rate for oral squamous cell carcinoma is greater in men than women, the rates in both increase dramatically with advancing age.

Table 1 Oral cancer death rates: selected countries[a]

Country	Male	Rank	Country	Female	Rank
Hong Kong	14.8	1	Singapore	4.8	1
France	14.3	2	Hong Kong	4.8	2
Singapore	12.8	3	Kuwait	2.4	3
Hungary	12.5	4	Cuba	1.8	4
Puerto Rico	9.0	5	Malta	1.7	5
Italy	6.4	11	Puerto Rico	1.4	11
Germany	6.0	13	United States	1.4	13
Spain	5.7	16	Canada	1.3	16
Canada	4.5	22	Luxembourg	1.2	22
Denmark	3.8	28	Belgium	1.0	28
United States	3.8	30	Germany	1.0	30
England and Wales	2.8	37	Netherlands	0.8	37
Venezuela	2.2	42	Finland	0.7	42
Mexico	1.7	45	Chile	0.7	45
Greece	1.6	46	Japan	0.6	46
Ecuador	0.9	50	Korean Republic	0.3	50

[a]*Source*: World Health Organisation data as adapted by the American Cancer Society, 1992.

for oral cancer (see **Table 1**) ranges in men from 0.9 in Ecuador to 14.8 in Hong Kong and in women from 0.3 in the Korean Republic to 4.8 in Singapore. In the United States the mortality rate ranges from 1.4 in Utah to 7.0 in the District of Columbia. In most populations males have an incidence rate about three times that of females and the incidence and mortality rates in the United States are greater among blacks than whites.

The tongue and floor of the mouth are the most commonly affected intraoral sites. There are exceptions, such as occurs in some Indian and South American cultures where carcinoma of the hard palate is endemic because of the custom of 'reverse' smoking (chutta). Women in particular hold the lit end of a slowly burning, hand-made cigarette inside the mouth in close proximity to the palatal mucosa, a site that rarely gives rise to carcinoma in most parts of the world.

Aetiology

It is estimated that about 90% and 60% of all deaths in males and females, respectively, caused by oral cancer can be attributed to cigarette smoking. Like cancers at other sites caused by smoking, the risk of developing cancer is dose related and is related to carcinogenic aromatic hydrocarbons. The relative risks are estimated to be at least five and as high as 10–24 for patients who smoke 40 cigarettes per day and this risk increases the longer one smokes. Pipe and cigar smoking appear to result in at least as great a risk. The association of the use of smokeless ('spit') tobacco with oral carcinoma is less clear. The prevalence of oral carcinoma is low in some countries where the use of smokeless tobacco far exceeds that of cigarette smoking (La Vecchia *et al.*, 1992; Vigneswaran *et al.*, 1995). On the other hand, some studies have shown that smokeless tobacco users are at increased risk and that about one-half of such cancers occur at the site where the tobacco is placed (Winn *et al.*, 1981). Another variation of topical use is seen in India and Southeast Asia where areca nut, tobacco and slaked lime are wrapped with betel leaf and chewed for long periods each day. This results in a precancerous scarring condition known as oral submucous fibrosis.

The fact that only a fraction of individuals with heavy exposure to tobacco and alcohol develop cancers of the upper aerodigestive tract suggests that there may be genetic differences between individuals that influence their susceptibility to these environmental agents. Heritable differences in head and neck cancer susceptibility have been found for nearly every step of tumorigenesis including carcinogen metabolism, DNA repair and progression as influenced by oncogenes and tumour-suppressor genes (Spitz, 1994; Khuri *et al.*, 1997; de Andrade *et al.*, 1998; Jahnke *et al.*, 1999). There is also increasing evidence to suggest that some of the population variance in response to therapy is due to interindividual genetic differences (Khuri *et al.*, 1997).

Because many heavy smokers are also heavy drinkers, it has been difficult to link alcohol use directly with oral cancer, but some investigators have estimated the relative risk at 2–6-fold. Alcohol acts synergistically with tobacco and together they increase a person's risk to 40-fold. Alcoholic beverages contain nitrosamines and hydrocarbons and, additionally, it has been proposed that contaminants or metabolites may promote malignant transformation (Blot, 1992). Malnutrition and vitamin deficiency may have a contributory role in oral cancers.

Vitamins A, B and C, are independently related to a reduced incidence of oral carcinoma and the risk of oral cancer in vitamin E users is half that of others according to one study (Gridley *et al*., 1992).

Cancer of the lower lip is strongly related to excessive exposure to ultraviolet light with a wavelength range of 2900–3200 Å, especially in fair-skinned individuals. Actinic cheilitis, similar in name and biology to actinic keratosis of the skin, represents the premalignant clinical condition. Just as in the skin, extensive damage is done to the collagen in the lamina propria (solar elastosis), but whether or not this event has any influence on epithelial transformation is unknown. Exposure to therapeutic X-irradiation is associated with an increased risk for the development of both carcinomas and sarcomas. Evidence of infection with human papillomavirus (HPV) has been found in clinically normal oral mucosa, benign and malignant neoplasms that arise from it, and some of the metastatic tumours. About 35% of oral cavity tumours have been found to contain HPV, usually the 'high-risk' types of the virus. HPV positivity correlates with age (< 60 years) and gender (male), but not with tobacco or alcohol use (McKaig *et al*., 1998). Nevertheless, its exact role remains elusive.

In summary, it appears that the pathogenesis of oral carcinomas is multifactorial with suppressor genes acting in association with growth factors, viruses, chemical carcinogens and oncogenes (Scully, 1993).

SCREENING AND PREVENTION: GROSS/HISTOPATHOLOGY/PREINVASIVE

Lesions/Ultrastructure/Immunohistochemistry

Squamous cell carcinoma (SCC) of the oral mucosa has numerous clinical appearances and is often preceded by premalignant lesions showing epithelial dysplasia or carcinoma *in situ*. The most common early form of premalignant or malignant disease, known as erythroplakia, is an asymptomatic, well-defined, erythematous macule or plaque that may have a finely granular surface texture. About 90% of erythroplakias show either dysplasia or carcinoma. SCC may also appear as a white patch but only between 5 and 25% of clinical white patches will show epithelial dysplasia or carcinoma, the remaining representing hyperkeratosis or other benign conditions. However, in sites designated 'high-risk,' including the ventral surface and lateral border of the tongue (**Figure 2; see colour plate section**), anterior floor of the mouth and soft palatal complex, the risk is much greater. Importantly, it is usually impossible to distinguish clinically between benign, premalignant and malignant disease. Often mucosal lesions show both red and white components (**Figure 3; see colour plate section**); the presence of an erythematous area in a lesion for which an obvious source of irritation is not evident should be biopsied and reviewed microscopically.

Advanced tumours may present as exophytic masses that often have a papillary surface (**Figure 4; see colour plate section**) and as nonhealing ulcers that reveal an endophytic, indurated nodule on palpation. Carcinomas that involve the gingiva, alveolar mucosa and tooth extraction sockets may first be discovered as irregular radiolucencies; those of the vermilion of the lower lip typically present as crusted ulcers (**Figure 5; see colour plate section**).

The most important morphological alterations that indicate a premalignant condition of the oral mucosa involve the squamous cell nuclei. They are enlarged, more darkly stained then normal, have irregular shapes and large, dark nucleoli and reveal increased and abnormal mitotic figures (**Figure 6; see colour plate section**). Cells showing these features extend above the parabasal cell layer in a haphazard arrangement, are crowded and often form bulbous rete ridges. The term carcinoma *in situ* signifies that the entire thickness of the spinous cell layer demonstrate these changes, but invasion has not yet occurred.

Grossly, squamous cell carcinoma is firm, has a glistening, heterogeneous grey–white cut surface and is often poorly delineated. Microscopically, individual epithelial cells or cords or islands of cells showing dysplastic changes infiltrate the underlying connective tissue (**Figure 7; see colour plate section**). When the tumour cells show a striking resemblance to normal squamous cells and show prominent keratinization, the tumour is considered 'well differentiated.' 'Moderately differentiated' carcinomas show a greater degree of cellular variability, less resemblance to normal squamous cells and much less keratinization. 'Poorly differentiated' tumours show very limited, focal resemblance to their cell of origin. Two benign oral lesions that often show an epithelial proliferation known as pseudocarcinomatous hyperplasia that histologically mimics well differentiated squamous cell carcinoma are granular cell tumour and necrotizing sialometaplasia.

The tendency of a patient to develop more than one oral mucosal primary carcinoma is know as 'field cancerization.' In a study of over 21 000 patients the rate of development of second tumours was 3.7% per year and the risk of a second was 2.8 times greater than expected (Day and Blot, 1992). The risks are highest in patients who continue to smoke and persist for more than 5 years after diagnosis of the initial cancer. Frequent periodic follow-up oral examinations for this possibility are essential.

Field cancerization is believed to be the result of multiple related lesions that arise from a field defect. This situation is similar to that seen in bladder cancer (Sidransky, 1997) and diffuse gastric cancer, where a single transformed cell clonally propagates and then spreads throughout a region. Subsequently, independent additional mutations occur in the

dispersed cells, giving rise to multiple malignancies over time, all of which share the original transforming mutations in common (Sidransky *et al.*, 1992; Sidransky, 1997). In head and neck squamous cell carcinoma (HNSCC), shared specific mutations in chromosomes 9p and 3p, as well as shared patterns of X-chromosome inactivation have been identified in synchronous and metachronous lesions along with independent mutations in each lesion (Nees *et al.*, 1993; Worsham *et al.*, 1995; Gauri *et al.*, 1996; Partridge *et al.*, 1997; Lydiatt *et al.*, 1998; Califano *et al.*, 1999). These findings indicate that multiple head and neck neoplasms may arise from a single clone which spreads, producing separate lesions that acquire additional mutations and progress individually.

There are several forms of squamous cell carcinoma that have important clinicopathological features different from the 'conventional' form previously described. Verrucous carcinoma is a low grade form that represents less than 10% of all oral squamous cell carcinomas. It most often occurs in patients who have used chewing tobacco for many years and appears as a white, papillary lesion that spreads slowly but progressively along the mucosal surface over one to several years (**Figure 8; see colour plate section**). Unlike conventional SCCs that may have a papillary configuration, the nuclear morphology is bland; characteristic architectural features, including broad, elongated, rounded rete ridges and papillary surface projections filled with parakeratin, facilitate the diagnosis. Invasive carcinomas develop in about 20% of cases mandating extensive tissue sampling. Proliferative verrucous leukoplakia is a term used to describe the white oral plaques that occur in some patients that, as they spread, change from having a smooth to fissured to granular to papillary surface texture (Hansen *et al.*, 1985). Some develop into verrucous carcinoma (**Figure 9; see colour plate section**) or conventional invasive SCC.

An aggressive variant is basaloid squamous carcinoma that most often occurs in the hypopharynx and base of the tongue. Histologically, the tumour is characterized by having well delineated nests of basaloid cells, some with hyaline cores and pseudoglandular structures, often resembling the solid variant of adenoid cystic carcinoma or small cell undifferentiated carcinoma. Weak immunoreactivity for neuron-specific enolase is seen in 75% of cases (Banks *et al.*, 1992) and ultrastructurally the basaloid cells reveal rare tonofilaments and varying numbers of desmosomes (Wain *et al.*, 1986). There is a poor prognosis; 40% of patients die within 17 months. Another variant is spindle cell carcinoma, a name chosen for the tumour's characteristic histopathological feature. The prominent spindle element resembles sarcomas but the focal presence of typical squamous cell carcinoma, dysplastic surface epithelium and immunoreactivity of the spindle cells for cytokeratin are all helpful in the distinction. These tumours are often clinically pedunculated, grow rapidly and about one-third are associated with prior therapeutic radiation.

Molecular Genetic Findings

Oral cavity and salivary gland neoplasms, like other solid tumours, arise from normal tissue that has undergone a series of genetic alterations. These changes include activating mutations of proto-oncogenes in addition to inactivation of tumour-suppressor genes, apoptotic genes and sometimes cell adhesion genes. For at least some cancers, tumour development also involves abnormalities in the cell's DNA repair system such that somatic mutations accumulate leading to malignancy (Sherbet and Lakshmi, 1997).

Investigations into the genetic-basis HNSCCs have resulted in a general molecular progression model for these neoplasms (**Figure 10**) (Califano *et al.*, 1996; Sidransky, 1997). This model resembles those developed by similar means for bladder cancer and non-small cell lung cancer (Sidransky *et al.*, 1992; Minna *et al.*, 1997) more than it corresponds to the colorectal cancer mutation paradigm (Fearon and Vogelstein, 1990). Whether this general progression model will hold true for all squamous cell

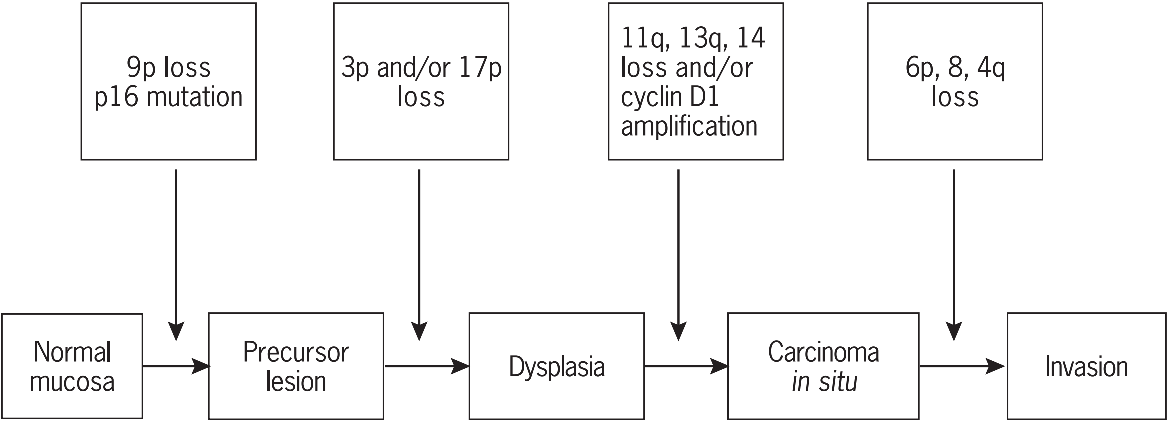


Figure 10 Preliminary molecular progression model for head and neck squamous cell carcinoma.

carcinomas of the head and neck remains to be seen. Some recent genetic evidence suggests that primary head and neck cancers arising in different locations are, in fact, separate clinical entities with independent genetic aetiologies and should be treated as such (Takes *et al*., 1998; Gleich *et al*., 1999).

Part of the rationale for grouping all head and neck cancers together, or at least all HNSCCs, has to do with shared aspects of disease management, their natural history (e.g. field cancerization and tendency for synchronous and metachronous lesions) and the common role of carcinogens in their origin, particularly the role of tobacco and alcohol exposure (Sidransky, 1997; Spitz, 1994). In molecular genetic studies, this tendency to group various head and neck cancers together makes it difficult to interpret the literature with regard to any particular subset, such as oral cavity lesions. Some research aimed at identifying specific genes involved in tumorigenesis and progression for head and neck cancers in general did not include oral cavity tumours in the study sample. A relatively small number of investigations focused on oral cavity lesions specifically. Therefore, it is unclear whether a conserved set of gene abnormalities is responsible for all head and neck cancers (or at least all HNSCCs) with minor variations on the theme, or whether each neoplasm has a distinct aetiology as has been suggested recently (Takes *et al*., 1998; Gleich *et al*., 1999)

Chromosomal changes detectable by cytogenetic methods provided the first clue to the identification of specific genes involved in tumorigenesis and progression. The chromosomal regions involved in rearrangement, deletion, amplification, etc., in tumour cells indicate sites where oncogenes, tumour-suppressor genes or other genes contributing to tumour development reside. The chromosomal changes most frequently reported for HNSCC include deletions of 3p, 5q, 8p, 9p, 18q and 21q (Sidransky, 1997), amplifications of 11q (Callander *et al*., 1994; Sidransky, 1997) and breakpoints at 1p22, 3p21, 8p11 and 14q (Sidransky, 1997). These common abnormalities identify potential tumour-suppressor genes and oncogenes contributing to HNSCC progression. Their correlation with certain stages of tumour progression allowed for their relative placement, chronologically, in a preliminary genetic progression model (**Figure 10**) (Jares *et al*., 1994; Sidransky, 1997). Multiple regions of deletion on chromosome arm 13q indicate the presence of a tumour-suppressor gene implicated specifically in oral cavity and supraglottal SCC (Gupta *et al*., 1999).

Further refinement of the location of putative tumour-suppressor genes contributing to head and neck cancers is provided by loss of heterozygosity (LOH) analysis to define the minimal regions of loss within the larger common areas of chromosomal deletion. Correlation of these regions defined by molecular markers, with the known map position of various tumour-suppressor genes, reveals likely candidates for gene mutation (followed by loss of the corresponding normal allele on the deleted chromosome). Evidence from LOH analysis and direct mutational screening indicates that a number of tumour-suppressor genes play a role in HNSCCs, including *TP53*, *VHL*, *CDKN2/INK4/p16*, *p21/WAF/CIP* and *TGFβR* (Brachman, 1994; Sidransky, 1997). Molecular genetic studies on oral cavity lesions specifically have implicated *TP53*, cell cycle inhibitors *p16* and *p27* and a tumour-suppressor gene other than *RB* or *BRCA1* located on the long arm of chromosome 13 (Gupta *et al*., 1999; Kanekawa *et al*., 1999; Riese *et al*., 1999; Venkatesan *et al*., 1999).

Despite the abundance of characterized oncogenes believed to play a role in tumorigenesis in humans, relatively few have been convincingly linked to the progression of any primary tumour. In head and neck cancer, there are a few cases of direct alteration of an oncogene or its overexpression and stage-specific progression. Well documented examples of oncogene activation in HNSCCs and corresponding tumour-progression include *cyclin D1* and *EGFR* (Brachman, 1994; Sidransky, 1997). There is strong evidence that overexpression of *cyclin D1* and the ligand–receptor pair HGF–c-Met play a role in progression of oral cavity neoplasms (Marshall and Kornberg, 1998; Matthias *et al*., 1998).

A variety of other genes have been identified that, when mutant, are involved in uncontrolled cell proliferation or the mechanisms by which benign lesions becomes invasive and metastatic. Some examples of additional genes that appear to promote development of various head and neck cancers including oral cavity tumours are *bcl-2*, *E-cadherin*, *EMS-1*, *telomerase*, *Ki-67*, *retinoic acid receptor*, *e1F4E*, *K19*, *GADD 153*, and *Cat D*, *B* and *L*. These genes encode apoptosis factors, cell adhesion molecules, translation initiation factors, cysteine proteases and other proteins for which the function of the altered gene product is unclear (Brachman, 1994; Kos *et al*., 1996; Drachenberg, *et al*., 1997; Friedman *et al*., 1997; Takes *et al*., 1998; Vo and Crowe, 1998; Crowe *et al*., 1999; Los *et al*., 1999; Nathan *et al*., 1999).

Numerous investigators have attempted to determine the clinical relevance of specific gene mutations in terms of their value as diagnostic markers or their reliability in predicting outcome and likelihood of recurrence. Some genes believed to play a role in the development of oral cavity tumours show potential as prognostic indicators or molecular markers of response to therapy, most notably low levels of p27 protein and poor response to treatment (Venkatesan *et al*., 1999), *GADD153* (growth arrest and DNA damage gene) mRNA levels as a predictor of response to cisplatin (Los *et al*., 1999), overexpression of *bcl-2* and poor prognosis, and *cyclin D1* polymorphism as a predictor of clinical outcome (Friedman *et al*., 1997). In addition, flow cytometric DNA content measurements (indicative of aneuploidy) can provide significant prognostic information, such as serving as an independent predictor of metastatic potential and clinical outcome

(Ensley and Maciorowski, 1994; Hemmer *et al*., 1999). Abnormal chromosome copy numbers measured by FISH in exfoliated epithelial cells of tumours and clinically normal margins can detect subclinical tumorigenesis (i.e. chromosome imbalances indicative of preneoplastic cells in the 'field') (Barrera *et al*., 1998).

Prognostic Factors

It has been shown that 14% of epithelial dysplastic lesions that were not excised progressed to invasive carcinoma within a follow-up period of 20 years, but 50% remained unchanged and 15% regressed (Lumerman *et al*., 1995). Nonetheless, the potential progression of any dysplastic lesion to carcinoma must be presumed and, when possible, should be excised with microscopically uninvolved surgical margins. Only 6.2% of 65 patients who had excision developed SCC compared with 15.4% of 91 patients who received no treatment (Lumerman *et al*., 1995).

Prognosis of oral squamous cell carcinoma is most directly related to the clinical stage of the tumour. The TNM staging system and Stage Grouping are shown in **Table 2**. Prior to 1992, 'fixation' of a node was considered in determining the N category, but now only the size of the node is evaluated, and bilateral nodal involvement is N2 rather than N3.

Most staging protocols do not include microscopic features. Nonetheless, some investigators have suggested that tumour thickness or depth of invasion independently correlates with patient outcome (Spiro *et al*., 1986). One study, for instance, found that tumours that invaded less than 4 mm, 4–8 mm and more than 8 mm showed metastatic rates of 8.3, 35 and 83%, respectively (Shingaki *et al*., 1988). Others have shown that tumour thickness correlates best with lesions of the vermilion border.

Overview of Present Clinical Management

Treatment failure most often is associated with uncontrolled regional disease. For most tumours, surgery with or without adjunctive postoperative radiotherapy is considered standard therapy, but patients with detectable distant metastases are usually treated palliatively. Radiation includes both external beam and interstitial brachytherapy. Local recurrence is more likely to follow excision of tumours with positive surgical margins so the goal is to establish tumour-free margins of 1–2 cm. The risks of occult cervical metastases in clinically negative necks is significant so elective neck dissections are often performed. Identification of positive lymph nodes in elective neck dissections has been reported to be 19, 40–50 and 25–54%, respectively, for tumours of the gingiva, floor of the mouth and oral tongue (de Braud *et al*., 1989). Elective radiation to the regional lymph nodes of the neck is often performed instead.

Systemic and regional chemotherapy is used mainly for palliation of patients with tumours that can not be treated effectively with surgery or radiation, and to help improve regional control of aggressive tumours. The drugs used most often singly or in combination with others are methotrexate, bleomycin, *cis*-platinum and 5-fluorouracil.

Survival benefits with chemotherapy have not been demonstrated. Research and early clinical trials using monoclonal antibodies, such as antiepidermal growth factor antibody C225, are under way and, it is hoped, will offer effective additional treatment options in the future (Huang *et al*., 1999).

It has only recently become possible to integrate information about an individual's genetic susceptibility and/or knowledge of specific genetic abnormalities in an individual's tumour with overall risk and response to particular therapies. Retinoid chemoprevention in head and neck cancers represents one of the most promising attempts at this sort of 'rational' intervention, in which levels of *RAR-B* expression in oral lesions predict response to intervention with retinoids (Khuri *et al*., 1997). The mechanism of action of retinoids and role of *RAR-B* in tumorigenesis are becoming clear (Grandis *et al*., 1996; Lotan, 1996; Cheng and Lotan, 1998; Vo and Crowe, 1998).

Gene therapy is also under exploration as a possible intervention for oral cavity and salivary gland cancers. This strategy provides a less invasive means of therapy compared with traditional management, and may prove particularly effective in combating aspects of recurrence associated with field cancerization. Adenovirus-mediated *p53* gene therapy has reached phase II clinical trials (Clayman *et al*., 1999), and combination chemo- and gene therapy is under initial investigation in mice as a means of treating salivary gland tumours (O'Malley and Li, 1998).

TUMOUR PATHOLOGY: SALIVARY GLAND CARCINOMAS

Epidemiology

The annual incidence of both benign and malignant salivary gland tumours reported in most studies varies from 0.4 to 6.5 cases per 100 000 population. They account for about 2% of all neoplasms of the head and neck. Primary epithelial tumours comprise approximately 90% of the tumours seen, with most of the remaining cases being soft tissue tumours, lymphomas and metastatic tumours. Adenocarcinomas, on which this discussion is focused, represent between 21 and 45% of the primary tumours. Current classifications of salivary gland tumours include 23 different types of carcinomas, and all but one, primary squamous cell carcinoma, represent adenocarcinomas.

The average age of patients with salivary carcinomas is 47 years, but two of the most common types, mucoepidermoid carcinoma and acinic cell adenocarcinoma, have

Table 2 TNM staging system for oral squamous cell carcinoma[a]

T	*Primary tumour*
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
TIS	Carcinoma *in situ*
T1	2 cm or less in greatest dimension
T2	More than 2 cm, but not more than 4 cm, in greatest dimension
T3	More than 4 cm in greatest dimension
T4 (lip)	Tumour invades adjacent structures such as cortical bone, inferior alveolar nerve, floor of mouth, skin of face
T4 (oral cavity)	Tumour invades adjacent structures such as cortical bone, muscle of tongue, maxillary sinus, skin. Superficial erosion alone of bone/tooth socket by gingival primary is not sufficient to classify as T4
N	*Regional lymph nodes*
NX	Regional lymph node cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single ipsilateral cervical lymph node(s) less than 3 cm
N2	Metastasis in a single ipsilateral node more than 3 cm but not more than 6 cm in greatest dimension; or in multiple ipsilateral nodes, none more than 6 cm; or in bilateral or contralateral lymph nodes, none more than 6 cm
N2a	Metastasis in single ipsilateral lymph node 3–6 cm
N2b	Metastasis in multiple ipsilateral nodes, none more than 6 cm
N2C	Metastasis in bilateral or contralateral lymph nodes, none more than 6 cm
N3	Metastasis in a lymph node more than 6 cm in greatest dimension
N3a	Clinically palpable ipsilateral node(s), at least one more than 6 cm
N3b	Bilateral clinically palpable bilateral nodes
N3c	Clinically palpable contralateral nodes
M	*Distant metastasis*
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Clinical stage – grouping of carcinoma of the oral cavity

Stage	**T**	**N**	**M**
Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IVA	T4	N0	M0
	T4	N1	M0
	Any T	N2	M0
Stage IVB	Any T	N3	M0
Stage IVC	Any T	Any N	M1

[a] Adapted from American Joint Commission on Cancer, 1997, *AJCC Staging Manual*, 5th edn. 24–27 (Lippincott-Raven, Philadelphia).

a peak incidence in the third and fourth decades, unlike most of the other types. Females are more often affected than males but there is variation according to the tumour type. About 55, 10, 0.3 and 34% of the carcinomas occur in the parotid, submandibular, sublingual and minor glands, respectively. However, the proportion of all salivary gland tumours at a particular site that are malignant varies significantly. For instance, while less than 25% of parotid tumours are malignant, 45% of submandibular tumours, 50% of minor gland tumours and nearly 90% of sublingual gland tumours are carcinomas. Furthermore, in some minor gland sites, including the tongue, floor of the mouth

and retromolar area, 80–90% are malignant. Most parotid gland tumours arise in the lateral (superficial) lobe and present as preauricular swellings. Those that develop in the deep lobe often expand into the parapharyngeal space and manifest as pharyngeal swellings.

Aetiology

Little is known about the cause of most salivary gland carcinomas, but there is a strong relationship with exposure to ionizing radiation. Studies of the survivors of the atomic bombings of Hiroshima and Nagasaki demonstrated an 11-fold increased risk for development of salivary carcinoma. The risk was greatest in patients closest to the hypocentre and during a period 12–16 years after exposure. Patients exposed to therapeutic radiation also appear to be at increased risk.

Although there is a strong association of cigarette smoking with the development of Warthin tumour, a benign neoplasm, no such association has been shown for salivary carcinomas. Occupations associated with increased risk of development of salivary gland carcinomas in a small number of cases include asbestos mining, industries with significant use of rubber products and plumbing (Auclair *et al*., 1991). A strong association of lymphoepithelial carcinoma, a rare salivary gland malignancy, with Epstein–Barr virus has been shown. Hamilton-Dutoit *et al*. (1991) found EBV genomes in all 18 such tumours removed from Greenlandic and Alaskan Eskimos that they studied.

Gross/Histopathology/Preinvasive Lesions/Ultrastructure/ Immunohistochemistry

Most patients with a salivary gland tumour, whether benign or malignant, present with an asymptomatic mass and both pain and rapid growth occur in both benign and malignant tumours (**Figures 11 and 12; see colour plate section**). However, facial paralysis occurs in about 12% of carcinomas but rarely in benign tumours and, when present, tumour fixation or ulceration suggest malignancy.

Most salivary gland carcinomas arise *de novo* rather than from morphologically recognizable premalignant disease. The most notable exception is carcinoma that arises from mixed tumours (carcinoma ex mixed tumour). The incidence of malignant transformation in mixed tumours, which is the most common of all benign salivary gland tumours, increases with duration. The incidence in tumours present for 5 years or less is 1.6% compared with 9.6% for those present more than 15 years (Eneroth and Zetterberg, 1974). Microspectrophotometric analysis has shown a similar tetraploid fraction in benign mixed tumours of long duration and in carcinomas.

Salivary gland carcinomas demonstrate extremely diverse cellular and architectural features (**Figure 13; see colour plate section**). Unlike carcinomas at many sites, most do not demonstrate a significant degree of nuclear pleomorphism and, therefore, their recognition as malignant is based on their characteristic growth patterns. Histological grading is based on one of four methods. For most salivary gland carcinomas there is a single grade determined by classification. For instance, acinic cell adenocarcinoma or basal cell adenocarcinoma are low grade whereas salivary duct carcinoma or undifferentiated carcinoma are high grade. The other three methods are uniquely applied to individual tumours: adenocarcinoma, not otherwise specified, is graded on its cytomorphological features, adenoid cystic carcinoma on the predominant growth either as cribriform-tubular (intermediate grade) or solid (high grade), and mucoepidermoid carcinoma on specific criteria that include the presence or absence of growth characteristics and cytomorphological features (Auclair *et al*., 1992).

Immunohistochemical staining shows that in normal gland, intercalated, striated and excretory ductal epithelium react strongly with antibodies for keratin intermediate filaments, but acinar and myoepithelial cells react weakly or not at all. Myoepithelial cells are reactive with smooth muscle actin antibodies and variably with antiglial fibrillary acidic protein antibodies. Immunoreactivity for S-100 protein is seen in intercalated duct and myoepithelial cells in glandular tissue that is next to tumour or severe inflammation but otherwise variable. The usefulness of immunohistochemical studies in diagnostic surgical pathology of salivary gland tumours is very limited. It is helpful in demonstrating neuroendocrine differentiation in small cell undifferentiated carcinoma. It is also invaluable in the identification of benign and malignant mesenchymal tumours and metastatic tumours such as malignant melanoma that involve the parotid glands.

Molecular Genetic Findings

Far fewer cytogenetic and molecular genetic studies have been carried out on salivary gland tumours than on SCCs of the oral cavity and, of those, most have investigated genetic alterations in pleomorphic adenoma and adenocarcinoma. For various salivary gland tumours, the most common chromosomal abnormalities reported are 11q;19p translocations (Martins *et al*., 1997), rearrangement at 6p23, amplifications and other abnormalities of 8q12 (Voz *et al*., 1998) and LOH at 3p, 6q, 8p, 8q and 12q (Gillenwater *et al*., 1997). The genetic lesions at 8q12 are associated with overexpression of *PLAG1*, the pleomorphic adenoma proto-oncogene which appears to be an early event in tumorigenesis (Voz *et al*., 1998). Tumour-suppressor genes involved in salivary gland tumorigenesis include *TP53*, *CDKN2*/*INK4*/*p16*, and an unidentified locus on 6q (Pignataro *et al*., 1998; Quemao *et al*., 1998; Suzuki *et al*.,

1998). Other genes that appear to play a role in development of salivary gland neoplasms include *TNF-α*, *bcl-2*, *MDM2* and the oncogenes *MYC* and *CDK4* which reside at 8q and 12 q (Rao *et al*., 1998; Soini *et al*., 1998). At present there is no genetic progression model for any salivary gland neoplasm or for salivary gland tumours in general.

Prognostic Factors

Clinical stage of disease correlates well with patient outcome for major gland disease. Staging largely depends on the size of the primary tumour and presence or absence of local neoplastic extension (**Table 3**). Local extension is the

Table 3 Staging system for major salivary glands[a]

Primary Tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1*	Tumour 2 cm or less in greatest diameter without extraparenchymal extension
T2*	Tumour more than 2 cm but not more than 4 cm in greatest dimension without extraparenchymal extension
T3*	Tumour having extraparenchymal extension without seventh nerve involvement and/or more than 4 cm but not more than 6 cm in greatest dimension
T4	Tumour invades base of skull, seventh nerve, and/or exceeds 6 cm in greatest dimension
Regional lymph nodes (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension
N2	Metastasis in a single ipsilateral lymph node, more than 3 cm but not more than 6 cm in greatest dimension, or in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension, or in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension
N2a	Metastasis in a single ipsilateral lymph node more than 3 cm but not more than 6 cm in greatest dimension
N2b	Metastasis in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension
N2c	Metastasis in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension
N3	Metastasis in a lymph node more than 6 cm in greatest dimension
Distant metastasis (M)	
MX	Presence of distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Clinical stage – grouping of carcinoma of the major salivary glands

Stage	**T**	**N**	**M**
Stage I	T1	N0	M0
	T2	N0	M0
Stage II	T3	N0	M0
Stage III	T1	N1	M0
	T2	N1	M0
Stage IV	T4	N0	M0
	T3	N1	M0
	T4	N1	M0
	Any T	N2	M0
	Any T	N3	M0
	Any T	Any N	M1

[a] Adapted from American Joint Commission on Cancer, 1997, *AJCC Staging Manual*, 5th edn. 53–55 (Lippincott-Raven, Philadelphia).

clinical or macroscopic evidence of invasion of skin, soft tissues, bone or nerve. Regional nodes are those within or immediately adjacent to the salivary gland and the deep cervical lymph nodes. There are no comparable staging criteria for intraoral minor gland tumours, but it has been shown that the criteria for intraoral squamous cell carcinoma can be applied successfully (Spiro *et al*., 1991).

The grade of the tumour also influences the prognosis. For instance, in two recent studies, 3.3% of patients with low-grade mucoepidermoid carcinomas died of disease compared with 46.3% of those with high-grade tumours (Auclair *et al*., 1991; Goode *et al*., 1998). Tumour site also appears to affect biological behaviour with tumours of the submandibular gland of the same type and grade having the worst prognosis.

Overview of Present Clinical Management

The principal therapy for salivary gland carcinomas is surgical and the best results correlate with complete initial tumour resection. Partial parotidectomy, usually removal of the lateral lobe, with preservation of the facial nerve, if possible, is indicated for low-grade carcinomas whereas total parotidectomy is performed for high-grade carcinomas and recurrent low-grade tumours. Total removal of the submandibular or sublingual gland is indicated for carcinomas in those sites. Low-grade carcinomas of the minor glands are removed by wide local excision whereas high-grade tumours are more radically excised. For tumours of the hard palate, bone is not removed for low-grade carcinomas except for those exceptional cases that erode or infiltrate the palatal bone. For tumours of all sites, prophylactic neck dissection is indicated for high-grade carcinomas and in any patient in whom clinically suspicious lymph nodes are discovered. Radiotherapy is used postoperatively for some high-grade tumours and for treating residual tumour when positive surgical margins are found. Neutron beam therapy has shown promise in controlling locoregional disease but needs further study (Spiro, 1998). Chemotherapy is often used for palliation of patients with recurrent, unresectable disease.

REFERENCES

Auclair, P. L., *et al*. (1991). Salivary gland neoplasms: general considerations. In: Auclair, P. L., *et al*. (eds), *Surgical Pathology of the Salivary Glands*. 135–164 (W. B. Saunders, Philadelphia).

Auclair, P. L., *et al*. (1992). Mucoepidermoid carcinoma of intraoral salivary glands. Evaluation and application of grading criteria in 143 cases. *Cancer*, **69**, 2021–2030.

Banks, E. R., *et al*. (1992). Basaloid squamous cell carcinoma of the head and neck. A clinicopathologic and immunohistochemical study of 40 cases. *American Journal of Surgical Pathology*, **16**, 939–946.

Barrera, J. E., *et al*. (1998). Malignancy detection by molecular cytogenetics in clinically normal mucosa adjacent to head and neck tumours. *Archives of Otolaryngology and Head and Neck Surgery*, **124**, 847–851.

Blot, W. J. (1992). Alcohol and cancer. *Cancer Research*, **52**, Suppl., 2119s–2123s.

Brachman, D. G. (1994). Molecular biology of head and neck cancers. *Seminars in Oncology*, **21**, 320–329.

Califano, J., *et al*. (1996). Genetic progression model for head and neck cancer. *Cancer Research*, **56**, 2488–2492.

Califano, J., *et al*. (1999). Second esophageal tumours in patients with HNSCC: an assessment of clonal relationships. *Clinical Cancer Research*, **5**, 1862–1867.

Callander, T., *et al*. (1994). PRAD-1(CCND1)/Cyclin D1 oncogene amplification in primary head and neck squamous cell carcinoma. *Cancer*, **74**, 152–158.

Cheng, Y. and Lotan R. (1998). Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G-protein coupled receptor. *Journal of Biological Chemistry*, **273**, 35008–35015.

Clayman, G. L., *et al*. (1999). Adenovirus-mediated wild-type p53 gene transfer as a surgical adjuvant in advanced head and neck cancers. *Clinical Cancer Research*, **5**, 1715–1722.

Crowe, D. L., *et al*. (1999). Keratin 19 downregulation by oral squamous cell carcinoma lines increases increases invasive potential. *Journal of Dental Research*, **78**, 1256–1263.

Day, G. L. and Blot, W. J. (1992). Second primary tumours in patients with oral cancer. *Cancer*, **70**, 14–19.

De Andrade, M., Amos, C. I. and Foulkes, W. D. (1998). Segregation analysis of squamous cell carcinoma of the head and neck: evidence for a major gene determining risk. *Annals of Human Genetics*, **62**, 505–510.

De Braud, F., *et al*. (1989). Metastatic squamous cell carcinoma of an unknown primary localized to the neck: advantages of an aggressive treatment. *Cancer*, **64**, 510–515.

Drachenberg, C. B., *et al*. (1997). Comparative study of invasive squamous cell carcinoma and verrucous carcinoma of the oral cavity: expression of bcl-2, p53 and Her-2/neu, and indexes of cell tunover. *Cancer Detection and Prevention*, **21**, 483–489.

Ellis, G. L. and Auclair, P. L. (1996). *Tumors of the Salivary Glands. Fascicle 17, Atlas of Tumor Pathology, 3rd Series. 1–8* (Armed Forces Institute of Pathology, Washington, DC).

Eneroth, C. M. and Zetterberg, A. (1974). Malignancy in pleomorphic adenoma. A clinical and microspectrophotometric study. *Acta Otolaryngologia (Stockholm)*, **77**, 426–432.

Ensley, J. F. and Maciorowski, Z. (1994). Clinical applications of DNA content parameters in patients with squamous cell carcinomas of the head and neck. *Seminars in Oncology*, **21**, 330–339.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model of colorectal tumorigenesis. *Cell*, **61**, 759–767.

Friedman, M., *et al*. (1997). Prognostic significance of Bcl-2 expression in localized squamous cell carcinoma of the head

and neck. *Annals of Otolaryngology, Rhinolaryngology and Laryngology*, **106**, 445–450.

Gauri, B., *et al*. (1996). Multiple head and neck tumours: evidence for a common clonal origin. *Cancer Research*, **56**, 2484–2487.

Gillenwater, A., *et al*. (1997). Microsatellite alterations at chromosome 8q loci in pleomorphic adenoma. *Archives of Otolaryngology and Head and Neck Surgery*, **117**, 448–452.

Gleich, L. L., *et al*. (1999). Variable genetic alterations and survival in head and neck cancer. *Archives of Otolaryngology and Head and Neck Surgery*, **125**, 949–952.

Goode R. K., *et al*. (1998). Mucoepidermoid carcinoma of the major salivary glands: clinical and histopathologic analysis of 234 cases with evaluation of grading criteria. *Cancer*, **82**, 1217–1224.

Grandis, J. R., *et al*. (1996). Retinoic acid normalizes increased gene transcription rate of TGF-alpha and EGFR in head and neck cell lines. *Nature Medicine*, **2**, 237–240.

Gridley G., *et al*. (1992). Vitamin supplementation and the reduced risk of oral and pharyngeal cancer. *American Journal of Epidemiology*, **135**, 1083–1092.

Gupta, V. K., *et al*. (1999). Multiple regions of deletion on chromosome arm 13q in head and neck squamous cell carcinoma. *International Journal of Cancer*, **84**, 453–457.

Hamilton-Dutoit, S. J., *et al*. (1991). Undifferentiated carcinoma of the salivary gland in Greenlandic Eskimos: demonstration of Epstein–Barr virus DNA by *in situ* nucleic acid hybridization. *Human Pathology*, **22**, 811–815.

Hansen J. L., *et al*. (1985). Proliferative verrucous leukoplakia: a long-term study. *Oral Surgery, Oral Medicine and Oral Pathology*, **60**, 285–290.

Hemmer, J., *et al*. (1999). DNA aneuploidy by flow cytometry is an independent prognostic factor in squamous cell carcinoma of the oral cavity. *Anticancer Research*, **19**, 1419–1422.

Huang, S. M., *et al*. (1999). Epidermal growth factor receptor blockade with C225 modulates proliferation, apoptosis, and radiosensitivity in squamous cell carcinomas of the head and neck. *Cancer Research*, **59**, 1935–1940.

Jahnke, V., *et al*. (1999). Genetic predisposition for the development of head and neck carcinomas. *Laryngorhinootologie*, **78**, 24–27.

Jares, P., *et al*. (1994). PRAD-1/Cyclin D1 gene amplification correlates with messenger RNA overexpression and tumour progression in human laryngeal carcinomas. *Cancer Research*, **54**, 4813–4817.

Kanekawa, A., *et al*. (1999). Chromosome 17 abnormalities in squamous cell carcinoma of the oral cavity, and its relationship with p53 and Bcl-2 expression. *Anticancer Research*, **19**, 81–86.

Khuri, F., *et al*. (1997). Molecular epidemiology and retinoid chemoprevention of head and neck cancer. *Journal of the National Cancer Institute*, **89**, 199–213.

Kos, J., *et al*. (1996). Prognostic significance of cathepsins D, B, H, and L and their protein inhibitors in breast, head and neck, and melanoma cancer (Meeting Abstract). *Proceedings of the Annual Meeting of the American Association of Cancer Research*, **37**, A622.

La Vecchia, C., *et al*. (1992). Trends of cancer mortality in Europe, 1955–1989: I. Digestive sites. *European Journal of Cancer*, **28**, 132–235.

Los, G., *et al*. (1999). Quantitation of the change in GADD153 mRNA levels as a molecular marker of tumour response in head and neck cancer. *Clinical Cancer Research*, **5**, 1610–1618.

Lotan, R. (1996). Retinoids and their receptors in modulation of differentiation, development, and prevention of oral premalignant lesions. *Anticancer Research*, **16**, 2415–2419.

Lumerman H., *et al*. (1995). Oral epithelial dysplasia and the development of invasive squamous cell carcinoma. *Oral Surgery Oral Medicine Oral Pathology, Oral Radiology and Endodontics*, **79**, 321–329.

Lydiatt, W. M., *et al*. (1998). Molecular support for field cancerization in the head and neck. *Cancer*, **82**, 1376–1380.

Marshall, D. D. and Kornberg, L. J. (1998). Overexpression of scatter factor and its receptor (c-met) in oral squamous cell carcinoma. *Laryngoscope*, **108**, 1413–1417.

Martins, C., *et al*. (1997). Cytogenetic characterisation of Warthin's tumour. *Oral Oncology*, **33**, 344–347.

Matthias, C., *et al*. (1998). Polymorphism within the cyclin D1 gene is associated with prognosis in patients with squamous cell cancer of the head and neck. *Clinical Cancer Research*, **4**, 2411–2418.

McKaig, R. G., *et al*. (1998). Human papilloma virus and head and neck cancer: epidemiology and molecular biology. *Head and Neck*, **203**, 250–265.

Minna, J. D., *et al*. (1997). Cancer of the lung. In: DeVita, V., *et al*. (eds), *Cancer: Principles and Practice of Oncology*. 849–857 (Lippincott-Raven, New York)

Nathan, C. A., *et al*. (1999). Expression of eIF4E during head and neck tumorigenesis: possible role in angogenesis. *Laryngoscope*, **109**, 1253–1258.

Nees, M., *et al*. (1993). Expression of mutated p53 occurs in tumour-distant epithelia of the head and neck cancer patients: a possible molecular basis for the development of multiple tumours. *Cancer Research* **53**, 4189–4196.

O'Malley, B. W. Jr and Li, D. (1998). Combination gene therapy for salivary gland cancer. *Annals of the New York Academy of Sciences*, **842**, 163–170.

Partridge, M., *et al*. (1997). Field cancerization of the oral cavity: comparison of the spectrum of molecular alterations in cases presenting with both dysplastic and malignant lesions. *Oral Oncology*, **33**, 332–337.

Pignataro, L., *et al*. (1998). p53 and cyclin D1 protein expression in carcinomas of the parotid gland. *Anticancer Research*, **18**, 1287–1290.

Quemao, L., *et al*. (1998). A refined localization of deleted regions in chromosome 6q associated with salivary gland carcinomas. *Oncogene*, **16**, 83–88.

Rao, P. H., *et al*. (1998). Nonsyntenic amplification of MYC with CDK4 and MDM2 in mixed tumour of the salivary gland. *Cancer Genetics and Cytogenetics*, **105**, 160–163.

Riese, U., *et al*. (1999). Tumor suppressor gene p16(CDKN2) mutation status and promoter inactivation in head and neck cancer. *International Journal of Molecular Medicine*, **4**, 61–5.

Scully, C. (1993). Oncogenes, tumour suppressors and viruses in oral squamous cell carcinoma. *Journal of Oral Pathology*, **22**, 337–347.

Sherbet, G. V. and Lakshmi, M. S. (1997). Clonal evolution of the metastatic phenotype. In: Sherbet, G. V. and Lakshmi, M. S. (eds), *The Genetics of Cancer – Genes Associated with Cancer Invasion, Metastasis and Cell Proliferation*. 4–20 (Academic Press, London).

Shingaki, S., *et al*. (1988). Evaluation of histologic parameters in predicting cervical lymph node metastasis of oral and oropharyngeal carcinoma. *Oral Surgery, Oral Medicine and Oral Pathology*, **66**, 683–688.

Sidransky, D. (1997). Cancer of the head and neck. In: DeVita, *et al*. (eds), *Cancer: Principles and Practice of Oncology*. 735–740 (Lippincott-Raven, New York).

Sidransky, D., *et al*. (1992). Clonal origin of bladder cancer. *New England Journal of Medicine*, **326**, 737–740.

Soini, Y., *et al*. (1998). Apoptosis is inversely related to bcl-2 but not bax expression in salivary gland tumours. *Histopathology*, **32**, 28–34.

Spiro, R. H. (1998). Management of malignant tumours of the salivary glands. *Oncology*, **12**, 671–680.

Spiro, R. H., *et al*. (1986). Predictive value of tumour thickness in squamous carcinoma confined to the tongue and floor of mouth. *American Journal of Surgery*, **152**, 345–350.

Spiro, R. H., *et al*. (1991). The importance of clinical staging of minor salivary gland carcinomas. *American Journal of Surgery*, **162**, 330–336.

Spitz, M. (1994). Epidemiology and risk factors for head and neck cancer. *Seminars in Oncology*, **21**, 281–288.

Suzuki, H. and Fujioka, Y. (1998). Deletion of the p16 gene and microsatellite instability in carcinoma arising in pleiomorphic adenoma of the parotid gland. *Diagnoses in Molecular Pathology*, **7**, 224–231.

Takes, R. P., *et al*. (1998). Differences in expression of oncogenes and tumour suppressor genes in different sites of head and neck squamous cell carcinoma. *Anticancer Research*, **18**, 4793–4800.

Venkatesan, *et al*. (1999). Prognostic significance of p27 expression in carcinoma of the oral cavity and oropharynx. *Laryngoscope*, **109**, 1329–1333.

Vigneswaran N., *et al*. (1995). Tobacco use and cancer. A reappraisal. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics*, **80**, 178–182.

Vo, H. P. and Crowe, D. L. (1998). Transcriptional control of retinoic acid responsive genes by cellular retinoic acid binding protein II modulates RA mediated tumour cell proliferation and invasion. *Anticancer Research*, **18**, 217–224.

Voz, M. L., *et al*. (1998). The recurrent translocation t(5;8)(p13;q12) in pleomorphic adenomas results in upregulation of the PLAG1 gene. *Oncogene*, **16**, 1409–1416.

Wain, S. L., *et al*. (1986). Basaloid-squamous carcinoma of the tongue, hypopharynx, and larynx: report of 10 cases. *Human Pathology*, **17**, 1155–1166.

Winn, D. M., *et al*. (1981). Snuff dipping and oral cancer among women in the southern United States. *New England Journal of Medicine*, **304**, 745–749.

Worsham, M., *et al*. (1995). Common clonal origin of synchronous primary head and neck squamous cell carcinomas. *Human Pathology*, **26**, 251–261.

FURTHER READING

McKinnell, R. G., *et al*. (eds) (1998). *The Biological Basis of Cancer*. (Cambridge University Press, New York).

Schantz, S. P., *et al*. (1997). *Tumors of the nasal cavity and paranasal sinuses, nasopharynx, oral cavity, and oropharynx*. In: DeVita, V. T., *et al*. (eds), *Cancer: Principles and Practice of Oncology*, 741–801 (Lippincott-Raven, Philadelphia).

Sessions, R. B., *et al*. (1997). Tumors of the salivary glands and paragangliomas. In: DeVita, V. T., *et al*. (eds), *Cancer: Principles and Practice of Oncology*. 830–847 (Lippincott-Raven, Philadelphia).

Silverman, S., Jr (1998). *Oral Cancer*, 4th edn (B.C. Decker, London).

Web Sites

http://rex.nci.nih.gov/NCI_Pub_Interface/raterisk
http://www.cancer.org
http://www.spohnc.org
http://cancernet.nci.nih.gov/canlit/canlit.htm
http://www3.ncbi.nlm.nih.gov/Omim/searchomim.html

Chapter 36

Respiratory System

Phillip S. Hasleton
Wythenshawe Hospital, Manchester, UK

CONTENTS

- Normal Upper Respiratory Tract
- Upper Respiratory Tract Tumours
- Lung
- Lung Tumours

The upper and lower respiratory tracts are the site of many common primary and secondary tumours. This area is exposed to many noxious influences, some environmental and others occupational.

NORMAL UPPER RESPIRATORY TRACT

The nose warms, humidifies and filters air via hairs and conchae (turbinates) (**Figure 1**). These cause alterations in airflow, trapping particles larger than 6 μm in diameter in nasal mucus.

The nasal sinuses are collections of air cells called the frontal, sphenoidal, maxillary and ethmoid sinuses. The inner nasal cavity and sinuses, with their ostia in the lateral nasal walls, are lined by ciliated, pseudostratified columnar epithelium. A type of mucus cell called the goblet cell appears in the sinuses. Beneath the epithelium are seromucinous glands, which produce additional mucus, IgA and other immunoglobulins as defence mechanisms.

The epiglottis prevents aspiration of food and other materials into the respiratory tract. Because of the antigen load, in the nasopharynx there are large masses of lymphoid tissue – the adenoids, the palatine, tubal and lingual tonsils and aggregates of lymphoid tissue, which circle the pharyngeal wall (Waldeyer's ring). This forms part of a mucosa-associated lymphoid tissue (MALT), part of the immunological defence of the lung and gastrointestinal tract (**Figure 1**). With antigenic stimulation, especially in childhood, these areas enlarge. Any marked enlargement of the nasopharyngeal tonsil causes mouth breathing. This affects the efficiency of nasal function, impairing pulmonary function.

The larynx is divided into the supraglottis, glottis and subglottis (**Figure 2**). The larynx acts as a vibrator, via the vocal cords for speech. During normal breathing the cords are held wide open to allow air passage. With speech the folds close, so air causes vibrations. The intrinsic laryngeal muscles are innervated by the recurrent laryngeal branch of the vagus nerve. Lung cancer metastases may entrap the vagus, causing hoarseness.

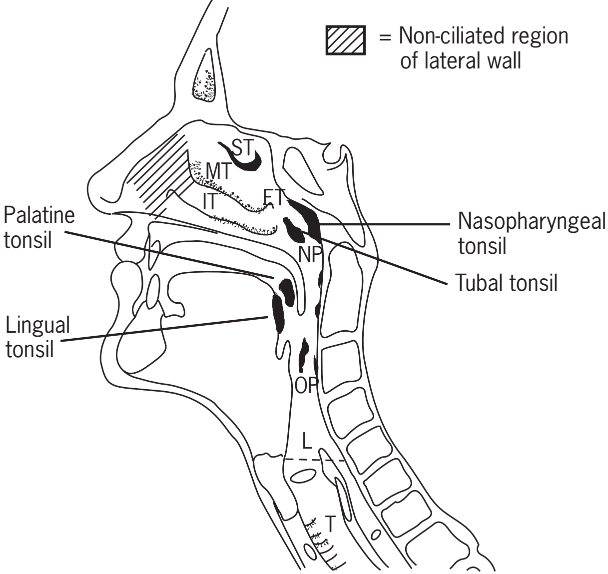


Figure 1 Diagram of nose, turbinates and larynx, showing relationships of the nasal cavity, nasopharynx (NP), oropharynx (OP), larynx (L) and distribution of lymphoid tissue (Waldeyer's ring). ST, MT and IT = superior, middle and inferior turbinates. ET = eustachian tube opening. Dotted line marks lower limit of the upper respiratory tract; below are the larynx (L), trachea (T) anteriorly and oesophagus (O) posteriorly. (Adapted from Jones, 1994.)

The epiglottis is lined by stratified squamous epithelium but in the lower half it gives way to a ciliated pseudostratified columnar type, characteristic of most of the larynx. The false cords enclose fibroadipose tissue, admixed with striated muscle and many seromucinous glands. The true cords are lined by stratified squamous epithelium and may contain some melanocytes along with elastic

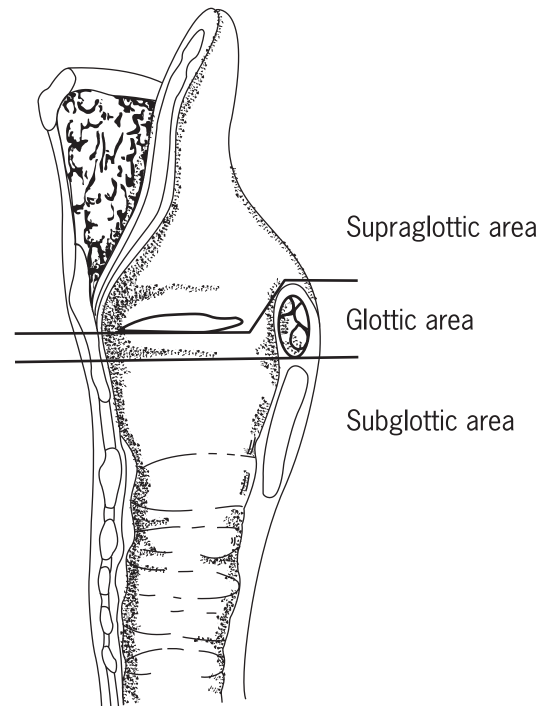


Figure 2 Division of the larynx into supraglottic, glottic and subglottic areas. [From Gregor, 1998, Jones, Phillips and Hilgers (eds.) *Diseases of the Head and Neck, Nose and Throat* (Arnold, London).]

tissue. The larynx is supported by a cartilaginous framework, connected by ligaments.

The trachea has a series of C-shaped cartilages extending into the bronchi, joined by fibroelastic membranes forming a hollow tube. Posteriorly lies the trachealis muscle. The midline cervical trachea lies anterior to the oesophagus. Subglottic tracheal lesions may cause oesophageal problems and vice versa. The isthmus of the thyroid is anterior to the second to fourth rings.

The trachea divides into the main bronchi. The right follows the general direction of the trachea, the left diverges at a greater angle. Thus aspiration is commoner in the right lung. The trachea and main bronchi are conducting airways. Bronchial cartilage is progressively decreased with distance from the trachea. The airways continue dividing into respiratory bronchioles, alveolar ducts and finally alveoli (**Figure 3; see colour plate section**). The terminal bronchiole is the smallest airway lined by bronchial epithelial cells.

The trachea and main bronchi have an inner lining of immunoglobulin and mucus for protection. Beneath is a lining ciliated, pseudostratified epithelium with a variable number of goblet cells attached to a basement membrane.

The cilia maintain the mucociliary escalator, causing upward passage of mucus and entrapped organisms or particulate matter to be expectorated. The ciliary shaft or cilium is a cytoplasmic extension from the surface of the cell. Just above the basement membrane are Feyrter cells [neuroendocrine (NE) cells], whose role in adults is unknown. They have clear cytoplasm and may occur as clusters, termed neuroepithelial bodies. NE cells contain dense core neurosecretory granules and secrete hormones. In the first 3 months of life, where there is relative hypoxia, they may act as chemoreceptors. These cells are the precursors of neuroendocrine tumours.

There is a surface, non-ciliated, bronchiolar secretory cell termed a Clara cell. These contain electron-dense membrane-bound inclusions and a few osmiophilic myelin bodies and produce surfactant apoprotein A.

In the subepithelial tissue there is collagen, elastin, nerves, lymphatics and small blood vessels. In addition there are serous and mucous glands (**Figure 4; see colour plate section**). The former produce lysozyme and if there is chronic cough they are converted to the mucous glands. The epithelium is regenerated by small pyramidal basal cells attached to the basement membrane.

Bronchus-associated lymphoid tissue (BALT) consists of subepithelial, mucosal, lymphoid follicles, containing both B and T lymphocytes. It develops at points of particle deposition, suggesting that it is stimulated as a response to inhaled particles. It is absent at birth and commoner in smokers.

The respiratory zone of the lung begins at the respiratory bronchiole and continues into alveolar ducts and alveoli. The alveolus is cup-shaped and thin-walled. Its cells can only be identified by electron microscopy. Up to 96% of the alveolar wall is covered by type I pneumocytes (**Figure 5; see colour plate section**). These are specialized cells, which cannot regenerate if damaged. The cytoplasm is thin to facilitate gas transfer between the alveolus and the pulmonary capillary. The edges of adjacent cells are bound by tight junctions, which restrict the movement of ions and water. These save the alveoli from flooding with water, as in pulmonary oedema.

Some 7% of the alveolar surface is covered with type II pneumocytes, which lie in the corners of alveolar walls. These cells form surfactant from their intracytoplasmic osmiophilic lamellar bodies. Surfactant is a phospholipid [dipalmitoylphosphatidylcholine (DPPC)]. This lowers the surface tension in the alveoli at the air/water interface. It acts similarly to a patch of oil on a road surface, repelling water. Any oedema can therefore be converted into droplets and they are removed by the pulmonary lymphatics.

The interstitial space is the part of the septal wall, which lies between the alveolar epithelial and capillary endothelial basement membrane. Normally it is inconspicuous but is distended in any form of alveolar damage. It contains macrophages, myofibroblasts, mast cells and occasional collagen and elastic fibres. Any thickening of this space causes alveolar diffusion problems. The interstitial connective tissue forms a continuous sheet with that surrounding blood vessels and bronchioles. This is efficient for removing fluid from alveoli into pulmonary lymphatics.

Pulmonary lymphatics aid the spread of infection and tumours. These are present around pulmonary blood

vessels at an alveolar level and in the pleura. These drain directly into the mediastinal nodes, in the upper lobes especially. The lymphatics can be traced to the respiratory bronchioles and continue around small bronchi and bronchioles forming a plexus outside muscle. If the lymphatics are distended, Kerley B lines are formed on chest radiographs.

UPPER RESPIRATORY TRACT TUMOURS

These tumours are well documented in the book *Tumours of the Upper Aerodigestive Tract* (Mills *et al*., 2000).

Tumours of the Nasopharynx

Clinical Features

Most nasal tumours cause nasal obstruction, occasionally epistaxis and facial pain, irrespective of their histology.

Squamous Papillomas

These are benign exophytic tumours with no association with human papillomavirus (HPV) and are seen in the oropharynx, larynx and trachea. They are covered by a bland, stratified, keratinizing squamous epithelium lining a fibrovascular core.

Schneiderian Papillomas

The ciliated columnar epithelium lining most of the nasal and paranasal cavity is called the Schneiderian membrane. The commonest tumour arising from this epithelium is the inverted Schneiderian papilloma, seen most commonly on the lateral nasal wall in the paranasal sinuses. They present as nasal polyps growing through the lateral wall within the maxillary or ethmoid sinuses. The nests of epithelium grow down into the underlying stroma with dilated ductal structures lined by multiple layers of epithelium. The epithelium may be squamous, ciliated, columnar or transitional and is often thick, being sometimes over 20 layers or more (**Figure 6; see colour plate section**). Mitoses are usually confined to the lower epithelial levels but nuclear pleomorphism may be identified. They may co-exist with a squamous cell carcinoma and any papillomas should be thoroughly sampled for evidence of carcinoma. These tumours have a recurrence rate of up to 75% when treated by local excision.

Fungiform Schneiderian papillomas are not associated with malignancy and arise predominantly from the nasal septum as polypoid tumours. The epithelium is similar to that of the inverted tumour but there are admixed mucous cells and little nuclear pleomorphism. Unlike papillary squamous cell carcinoma, there is no mitotic activity or dyskeratosis. They have a low recurrence rate.

Squamous Cell Carcinoma

These are relatively rare in the nasopharyngeal region. The distribution is maxillary antrum 58%, nasal cavity 30%, ethmoid sinuses 10% and frontal and ethmoid sinuses 1% each (Lewis and Castro, 1972). They are commonest on the lateral nasal wall. Tumours in the maxillary antrum may be misdiagnosed as chronic sinusitis, delaying treatment. Sinonasal squamous carcinomas occur predominantly in males, often in their 60s.

There is an association with cigarette smoking, nickel mining and refining and chromium exposure and, more recently, formaldehyde has been suggested as a carcinogen. HPV types 16 and 18 DNA have been found in 14% of sinonasal squamous cell carcinomas but the high sensitivity of polymerase chain reaction (PCR) suggests that contamination cannot be totally excluded (Furuta *et al*., 1992).

The tumours are typically papillary, polypoid and are usually moderately or well differentiated squamous cell carcinomas. Nasal lesions spread to the submental and submandibular nodes and they enter the facial and superficial parotid deep cervical nodes. Spread of paranasal sinus carcinoma to regional nodes is uncommon when the tumour is confined to the sinus cavity.

Sinonasal Carcinoma

This is strongly associated with hardwood dust exposure but additional 'risk' occupations are logging, milling, exposure to leather dust and softwood dusts (Cecchi *et al*., 1980). Most of these tumours occur in men, with origins in descending order of frequency in the ethmoid sinuses, nasal cavity, maxillary antrum and indeterminate. They present with the usual symptoms, as well as rhinorrhea or a mass in the cheek. The tumours mimic ulcerated or haemorrhagic inflammatory nasal polyps.

There is malignant, small and large intestinal type mucosa along with Paneth, goblet and argentaffin cells. These tumours do not usually express carcinoembryonic antigen (CEA), unlike metastatic small or large bowel carcinomas. Sinonasal adenocarcinomas are usually chromogranin-positive, which is less common in colonic carcinomas.

The optimum treatment is surgical resection with radiation therapy to the region of the tumour. Grade and subtype correlate with survival. The better prognosis is with the papillary subtype as opposed to sessile or alveolar mucoid variants (Batsakis *et al*., 1963). Woodworkers also appear to have a better survival than patients who present without this risk factor.

Nasopharyngeal Carcinoma

Nasopharyngeal carcinoma arises near the fossa of Rosenmuller (the internal opening of the middle ear canal) causes middle ear obstruction and involves the cervical nodes. These tumours are classified into keratinizing squamous cell

carcinoma, nonkeratinizing carcinoma and undifferentiated carcinoma. These tumours affect a broad age range, including children. There is a distinct bimodal age distribution with peaks in the second and sixth decades of life.

Cigarette smoking may play a part in lymphoepithelial carcinoma but this cell type is especially predominant in Chinese patients, even when they emigrate to other countries. It is associated with Epstein–Barr virus (EBV). In Hong Kong it accounts for 18% of all malignancies compared with 2% in the United States (Digby *et al.*, 1941). Chinese patients in Singapore with lymphoepithelial carcinoma have an increased frequency of histocompatibility antigens HLA-A2 and HLA-BW46 (Simons *et al.*, 1976).

Clinically the nasopharynx may be normal, full or show surface granularity. An obvious carcinoma is uncommon. It is usually a poorly differentiated squamous cell carcinoma often with a marked inflammatory component, including lymphocytes, plasma cells and eosinophils (lymphoepithelioma) (**Figure 7; see colour plate section**). They may be confused with a lymphoma because of the prominent inflammation. Mitoses are typically numerous, ranging from 5 to 10 per 10 high power fields. The tumour cells have a syncytial growth pattern or form cohesive nests and cords. They tend to metastasize to regional nodes, where they may be first detected. Distant metastases to lung, brain, liver and bone are common. Lymphoepitheliomas are usually treated by radiation.

Other variants of squamous carcinoma are adenoid, angiosarcoma-like, papillary, spindle cell, basaloid and adenosquamous.

Squamous carcinoma also occurs in the oropharynx and hypopharynx. In both of these sites surgery is the mainstay of treatment followed by radiation.

Neural Tumours

Nasal glioma is a heterotopia due to failure of the developing frontal lobe to retract completely via the foramen caecum. They are therefore encephaloceles, not true tumours. Ectopic pituitary tissue may be seen in the same area and present as a pituitary adenoma.

Nasopharyngeal neurofibromas, neuromas and neurilemmomas are similar to those seen in other areas of the body.

Melanotic neuroectodermal tumour of infancy (MNTI) occurs in infants under the age of 1 year; 70% are seen in the anterior maxillary area. These are lobular, well-circumscribed, unencapsulated tumours, measuring up to 13 cm in diameter. They are grey/white to dark brown with nests of small neuroblastic cells, often surrounded by larger melanin-containing cells with an alveolar or tubular pattern. Nuclear pleomorphism, mitoses and necrosis are uncommon.

Paraganglioma (carotid body tumour) can be seen in the nasal cavity, paranasal sinuses, nasopharynx and larynx. They are no different from extra-adrenal paraganglial tumours of the autonomic nervous system in other parts of the body.

Olfactory Neuroblastoma

There are no known risk factors for this tumour, which has a bimodal distribution, at the ages of 15 and 55 years. In addition to the usual symptoms, headaches, visual disturbances and anosmia may also occur. The ethmoid sinus is often involved. Radiologically the tumour has a 'dumbbell' appearance extending across the cribriform plate (**Figure 8; see colour plate section**).

They are polypoid, vascular, red/tan masses set high in the nasal cavity. The tumour may have circumscribed nests of cells, which are small with round nuclei and little cytoplasm or grow as diffuse sheets of neoplastic cells set in a background of capillaries and little intervening stroma. There is a fibrillary cytoplasmic background in haematoxylin and eosin-stained sections. Necrosis is a poor prognostic indicator.

Most of these tumours stain positively for synaptophysin. Treatment is by surgical excision with adjuvant radiotherapy or chemotherapy.

Undifferentiated Sinonasal Carcinoma

This is rare but is seen in both young adults and the elderly. Typically it presents as a large fungating mass obstructing the nasal cavity and invading surrounding structures. There are nests, sheets, ribbons and trabeculae of polygonal cells, often with an organoid appearance. Mitoses are prominent and there is extensive vascular permeation. This tumour must be differentiated from a lymphoepithelioma.

Primary Small Cell Carcinoma

This is similar to that seen in the lung and may arise rarely in the sinonasal region. Malignant melanoma is rarely seen as a primary tumour developing in the nasal cavity or the paranasal sinuses.

Lymphoid Tumours

Non-Hodgkin Lymphoma (NHL)

Up to 10% of NHL may involve Waldeyer's ring (adenoids, oropharynx, tonsilar and lymphoid tissue at the base of the tongue). After gastrointestinal lymphomas, it is the commonest extranodal lymphoid tumour. Most involve the tonsil but 35% involve the nasopharynx. It is often difficult to determine if a lymphoma is strictly confined to the nose or paranasal sinuses. There is often involvement of multiple sinuses. They are usually localized, affecting a wide age span from the first to the tenth decades.

The tumours are greyish/white, firm and up to 10 cm in diameter. Under the REAL classification 85% are B cell lymphomas, half of which are diffuse large cell

lymphomas (**Figure 9; see colour plate section**). There is a diffuse proliferation of large cells with immunoblastic features or diffuse mixed and large cells. Mantle cell and low-grade lymphomas also occur in Waldeyer's ring. A subset of nasopharyngeal lymphomas are of T/NK phenotype, especially in the Far East. These are probably human T cell leukaemia/lymphoma-virus induced, while others are more closely related to sinonasal T/NK cell lymphoma (Tomita *et al.*, 1996).

Radiation is the treatment of choice and chemotherapy is added for locally advanced or disseminated tumours. (See also chapter *RNA Viruses*.)

Sinonasal T/NK Cell Lymphoma

Virtually no B cells are seen in normal nasal epithelium but there are T cells. T helper cells outnumber the T cytotoxic/suppressor cells. T/NK cell lymphoma used to be termed midline malignant reticulosis, lethal midline granuloma and midfacial destructive lesion, amongst others. These terms are now no longer used. T/NK lymphoma is an angiocentric, immunoproliferative lesion. It is commoner in Asians. The disease can occur at any age with a median of 47 years with a male predominance (Aozasa *et al*., 1995).

The patients present with mid-facial destruction, involving the nasal cavity, paranasal sinuses and in some cases the palate and other mid-facial structures. It may extend to involve the larynx. There is sometimes collapse of the nasal bridge, ulceration, necrosis, septal perforation, fever and weight loss. The tumour often spreads to other extranodal sites, such as skin, subcutaneous tissue, gastrointestinal tract and testis.

Histologically there is extensive necrosis, which may involve the cartilage and bone. Small lymphocytes, plasma cells, immunoblasts and some cytologically atypical lymphocytes, as well as polymorphs, eosinophils and histiocytes are seen. There is much angioinvasion, causing thrombosis. Mitoses are frequent. T/NK lymphoma is typically reactive for CD2, a T cell marker, and CD 56, a marker of NK cells. However, clonal T cell receptor gene rearrangements are absent in this lymphoma.

Other Lymphomas of the Sinonasal Tract

These include extramedullary plasmacytoma, post-transplant lymphoproliferative disorder (PTLD) and rare cases of Hodgkin's disease. The histology of these tumours does not differ from those elsewhere in the body.

Vascular Tumours

Haemangioma is a hamartoma. They are common in the nose.

Nasopharyngeal Angiofibroma

This usually arises in the posterior nasal cavity and nasopharyngeal wall. It is seen almost exclusively in male adolescents. The tumours have testosterone receptors (Gown *et al*., 1993). Their vascularity can be reduced by oestrogen therapy. The mean age of the patients is 15 years and the presenting symptoms are those of any other nasal lesion. They are firm, sessile or polypoid tumours with a fleshy appearance. There is a fibrous stroma containing many vessels of varying sizes. Some are thick-walled with an obvious media (**Figure 10; see colour plate section**). The stroma consists of characteristic spindled and stellate shaped mesenchymal cells and haphazardly arranged collagen. Mitoses are rare. The deeper portions of an angiofibroma away from the central feeding vessels may be hypovascular. Angiofibromas lack the lobular arrangement of a capillary haemangioma. Treatment is by surgical excision but recurrence is common.

Glomus tumours, although rare, may occur in the nose.

Other Tumours

These include fibroma, solitary fibrous tumour, aggresive fibromatosis, osteoma, osteosarcoma, often involving the maxilla, teratoma involving the nasopharynx and metastatic tumours especially from breast, kidney and lung.

Other soft tissue tumours to involve the nasopharyngeal area are leiomyoma, chondroma, chondrosarcoma, rhabdomyosarcoma (nose) and leiomyosarcoma (nose).

Larynx

Squamous papillomas may arise from the true cords, at other sites in the larynx, the oropharynx and the trachea. They may be part of juvenile laryngeal papillomatosis, which may be single or multiple. 'Juvenile' is arbitrarily classified as presenting before the age of 20 years. Multiple is usually defined as more than three lesions.

Most laryngeal papillomas are probably caused by HPV, types 6 and 11 (Travis *et al.*, 1999). These tumours are glistening, nodular, exophytic masses and, depending on their site and size, cause hoarseness, stridor or respiratory distress. A rare complication is squamous cell carcinoma, which may develop in the larynx or lung.

Histologically there are multiple layers of orderly squamous epithelium with no atypia, covering a fibrovascular core. A few cases may show varying degrees of dysplasia. Prominent surface keratin or intraepithelial dyskeratosis suggests verucca vulgaris or veruccous carcinoma. Squamous carcinomas usually arise in the juvenile lesions. Treatment is by surgery but recurrence is frequent. Solitary adult and juvenile papillomas are often cured by surgery.

Keratosis and Dysplasia

Keratosis is an epithelial area with a marked degree of orthokeratosis or parakeratosis unassociated with underlying epithelial proliferation (AFIP). Typically keratosis is due to long-term tobacco abuse and affects the vocal cords. The mucosa is thickened and white. 'Leucoplakia' is a

clinical term meaning 'white plaque' and not a pathological description. Keratotic epithelium usually separates from the vocal cords. This is helpful diagnostically since an intact basal layer helps exclude invasive carcinoma.

Non-dysplastic keratosis shows a normal or increased thickness of squamous epithelium with a prominent granular layer and overlying layers of orthokeratin admixed with parakeratosis. Maturation is orderly and mitoses are basal. If there is no dysplasia there appears to be only a minimal risk of developing a subsequent carcinoma.

Dysplasia is graded into mild, moderate and severe and has similarities to the system used in the cervix except that keratinization is more common in the larynx. Mild dysplasia involves the basal layer, moderate extends upwards to involve usually two-thirds but does not involve the superficial epithelium. Severe dysplasia shows increased mitoses and involves the full thickness of the epithelium. The more severe the dysplasia, the greater is the chance of an invasive carcinoma co-existing or developing in the larynx. Involvement of the underlying seromucinous glands by dysplastic epithelium is not considered to be evidence of invasion.

Invasive Squamous Cell Carcinoma

Squamous carcinoma of the larynx is associated with cigarette smoking and excess alcohol intake. 'Social' use of alcohol does not appear to cause an increased risk. There is no association between asbestos exposure and carcinoma of this site. Carcinoma of the larynx is divided into supraglottic, glottic and subglottic types. This division has relevance for surgical management. Transglottic carcinoma is a term applied to carcinomas bridging the laryngeal ventricle and involving both the vocal cord and supraglottic regions. Most are glottic carcinomas extending into the supraglottis.

Most carcinomas of the larynx are glottic, arising anteriorly on the mobile part of the cord. They cause hoarseness. The majority of the remaining laryngeal carcinomas are supraglottic and arise from the epiglottis, ventricles, false cords and aryepiglottic folds. These give rise to changes in voice quality and may cause difficulty in swallowing or the feeling of a mass in the throat. Subglottic carcinomas account for 5% or less of all laryngeal carcinomas.

The tumours are usually ulcerated and vary in size from small lesions to large masses, which may almost obstruct the laryngeal lumen (**Figure 11; see colour plate section**). There are varieties of squamous cell carcinoma, including spindle cell. Since the true cords have a limited lymphatic supply, tumours at this site have a good prognosis. Thus they are often cured by radiation or limited surgical resection. With an increase in size the cords become fixed and the tumour may extend outside the larynx, involving nodes. Subglottic carcinomas are often circumferential and extend beyond the larynx, penetrating the cricothyroid membrane. These cases are likely to involve cervical nodes.

Early carcinomas are usually cured with limited surgery or radiation but laryngectomy is one of the treatments for larger tumours. Recently chemotherapy has been shown to play a role. The size of the tumour and degree of differentiation are important factors in determining prognosis. Patients with laryngeal carcinoma have an increased risk of developing another tumour, especially in the lung or elsewhere in the head or neck.

Genetic Changes in Laryngeal Carcinoma

P53 status over-expression does not result in cell cycle arrest in some studies, although a correlation with this oncogene and early stage glottic cancer as well as those tumours recurring locally has been described (Narayana *et al*., 1998). It may also relate to decreased survival (Bradford *et al*., 1997). C-erbB-2 plays no part in prognosis in this tumour (Krecicki *et al*., 1999). Retinoblastoma protein, however, does relate to survival (Dokiya *et al*., 1998). Cyclin D1 overexpression identified patients with poor grade laryngeal carcinoma, tumour extension, lymph node involvement and poor histological differentiation (Bellacosa *et al*., 1996).

Adenosquamous Carcinoma

This may affect the larynx. These are usually high-grade tumours, similar to adenosquamous carcinoma in the lung. Adenocarcinoma is very rare in the larynx and the possibility of a secondary tumour should be considered.

Neural Laryngeal Tumours

Neurilemmomas, neurofibromas, granular cell tumour, paraganglioma, malignant melanoma and carcinoid tumour [this includes all neuroendocrine carcinomas (see the section on the lung) can all be identified in the larynx. These tumours are histologically identical with the pulmonary lesions described below.

Lymphoma

Localized laryngeal non-Hodgkin lymphoma is rare. Extramedullary plasmacytoma has been well described in the larynx, where it causes a subepithelial tumour.

Haemangiomas may arise in the larynx. In infants they present with respiratory distress and affect the subglottic region. Adult haemangiomas are commoner and usually affect the supraglottic and glottic areas. They may cause cough and haemoptysis, as well as hoarseness. Inflammatory myofibroblastic tumour may involve the larynx and in some cases it is caused by *Mycobacterium avium intracellulari*, often in HIV-positive or immunosuppressed patients. The commonest site is on the vocal cords, where they cause hoarseness and stridor.

Carcinomas may metastasize to the larynx, the commonest tumour is melanoma followed by breast, kidney, lung, prostate and gastrointestinal tract carcinomas. Soft tissue laryngeal tumours include liposarcoma, chondrosarcoma and rarely rhabdomyosarcoma, leiomyosarcoma and malignant fibrous histiocytoma.

Trachea

The trachea is rarely involved by tumours, the commonest being polyposis and adenoid cystic carcinoma but any of the neoplasms mentioned below may affect this area.

LUNG

Cigarette Smoking

This is one of the major causes of mortality and morbidity in the Western world. As some of the tobacco sales decline in this area, the producers are ensuring that Third-world countries are targetted.

Cigarette smoke yields more than 4000 constituents. These include carbon monoxide, hydrogen cyanide, aldehydes, cadmium, ammonia, nicotine and benz[*a*]anthracene, a potent carcinogen. In addition, there are aromatic hydrocarbons and other toxic substances, which may be tumour initiators, such as benzanthracenes and benzopyrenes. There are suspended water droplets, which vary in size and have central resinous cores. These smoke droplets are absorbed on bronchial walls and propelled on the mucociliary escalator back to the mouth; 98% of smoke particles landing in the bronchi are removed by cilia within 24 h. Smaller particles enter alveoli and, if undissolved, are ingested by macrophages and removed to lymphatics.

Nicotine makes cigarettes addictive. The faster a cigarette is smoked, the more nicotine is present in the mainstream smoke; 85% is absorbed in the lungs and it causes an increase in heart rate, blood pressure and cardiac output.

Cigarette smoking may be active or passive. The latter occurs in smokers and nonsmokers, in the same environment as a smoker. Passive smoking increases the risk of lung cancer by 26% and of ischaemic heart disease by 23%. In addition, passive smoking is linked with an increased incidence of asthma and chest infections in children.

This habit affects virtually every organ system, ranging from adenolymphoma of the parotid to peripheral vascular disease. (See the chapter *Tobacco Use and Cancer*.)

LUNG TUMOURS

These may be primary or secondary, benign or malignant. Because the lung receives the entire cardiac output, tumour metastases are common.

Classification of Lung Tumours

There has been an increase in new pulmonary lesions described in the last 18 years (Travis *et al*., 1999). In addition to bronchial squamous dysplasia, the concept of atypical alveolar hyperplasia, as a precursor for adenocarcinoma of lung, is becoming well established. It is impossible in a chapter such as this to give a comprehensive description of all lung tumours and the reader is referred to specialist texts (Hasleton, 1996).

Benign Lung Tumours

These may be central or peripheral. Central indicates they involve main bronchi and peripheral, the lung parenchyma. The site of the neoplasm determines signs and symptoms. Central tumours cause collapse of a lobe or lung, with recurrent infections leading to bronchiectasis. All central lung tumours present with similar symptoms, i.e. cough, recurrent chest infections and haemoptysis, due to ulceration of the surface of the tumour.

Peripheral lung tumours are often detected as a chance radiological finding, known as 'a solitary pulmonary nodule.' As a rough guide, 40% of solitary nodules are malignant and 60% are benign. The benign lesions may well be inflammatory or non-neoplastic rather than benign tumours. For any individual patient these figures are meaningless and do not help with diagnosis. By convention a nodule is 4–6 cm in diameter; a larger lesion is termed a mass, often suggesting malignancy. The treatment options for a solitary nodule are outside the scope of this book but with computed tomography (CT) scans and fine-needle biopsy, some of the diagnostic responsibility rests with the histopathologist.

Radiologically benign lesions tend to have a smooth circumscribed periphery; malignant ones are larger with irregular margins. A calcified nodule can be followed up and is considered benign. If a nodule has been radiologically stable in size for 2 years and the patient is below 35 years of age, it is malignant in only 1–5% of cases. The widespread policy in remaining solitary nodules of unknown aetiology is resection.

It is unnecessary for anyone to memorize every benign lung tumour. If one remembers the normal bronchial wall components, i.e. epithelium, neuroendocrine cells, connective tissue, muscle, cartilage, fat and nerves, and the fact that the mucous and serous glands of the bronchial wall act as a minor salivary gland, the nature of the majority of benign tumours can be predicted. Only the important or commoner ones will be considered here.

Hamartoma

These tumours develop in adults and have an abnormal karyotype. The most characteristic aberration is a 6p21 rearrangement but a 12q14–15 rearrangement has also been identified. These molecular abnormalities indicate these 'hamartomas' are true neoplasms and a better term is 'mesenchymoma.' This tumour is twice as common in smokers and is associated with bronchial carcinoma. They are often peripheral and range from 1 to 4 cm in diameter. The cut surface is grey (**Figure 12; see colour**

plate section) but if fat is prominent it is yellow. If the tumour is central, there is distal bronchiectasis. They consist of cartilage, bone, fat, loose myxoid tissue and islands of ciliated or columnar epithelium.

Papilloma of the Bronchus

Papillomas are classified into solitary or multiple. The solitary lesions are rare, presenting in middle-aged smokers as a central tumour. It grows as a wart-like lesion into the bronchial lumen and consists of nonkeratinizing squamous epithelium. Papillomatosis is seen in children. HPV (human papillomavirus) types 6 and 11 can be identified in both solitary and multiple tumours. One-third show carcinoma *in situ* or invasive carcinoma and thus papillomas require at a minimum close follow-up.

Glandular and mixed squamous and glandular papillomas are solitary, benign central tumours.

Neuroendocrine Tumours

There are no real premalignant neuroendocrine lesions. DIPNECH (diffuse idiopathic pulmonary neuroendocrine hyperplasia) has been placed in the preinvasive section of the new WHO classification. The author has seen very rare cases, associated with tumourlets and typical carcinoids. However, DIPNECH is a nonspecific reaction to airway inflammation and/or fibrosis. There is neuroendocrine (NE) cell hyperplasia in association with peripheral carcinoid tumours. However, adjacent to central typical carcinoids there is a decrease in NE cells. No *in situ* lesion has been found in association with small cell carcinoma, although dysplastic squamous epithelium may lie over the tumour.

A tumourlet is defined by the WHO as 'a microscopic, peribronchiolar, nodular aggregate of uniform, round to oval or spindle-shaped cells with moderate amounts of cytoplasm and morphology similar to the cells of carcinoid tumours.' An arbitrary size of 0.5 cm or less is given for these lesions; larger tumours are termed bronchial carcinoids. Tumourlets usually occur with focal pulmonary fibrosis and bronchiectasis but not diffuse interstitial pulmonary fibrosis. They are found incidentally. They may be mistaken for small-cell carcinoma but lack this tumour's cytological features and have no mitoses.

Bronchopulmonary Carcinoid Tumours

As shown in **Figure 13** there is a spectrum of biological behaviour in these tumours, ranging from the 'benign' bronchopulmonary carcinoid to small-cell lung carcinoma (SCLC). The latter neoplasm accounts for 15–25% of all lung malignancies. This figure does not imply there is any transition from one tumour type to another. Typical carcinoid (TC), atypical carcinoid (AC) tumours and large-cell neuroendocrine carcinomas (LCNEC) are rare, together accounting for 2–3% of lung tumours. Nearly all patients with AC, SCLC and LCNEC are cigarette smokers. While TC and AC are morpologically similar to neuroendocrine tumours elsewhere in the body they do not always behave in a similar fashion. For example, the carcinoid syndrome is relatively rare in bronchial carcinoids but commoner with hepatic metastases from ileal tumours. Bone production is rare in gastrointestinal carcinoids but commoner in bronchial tumours.

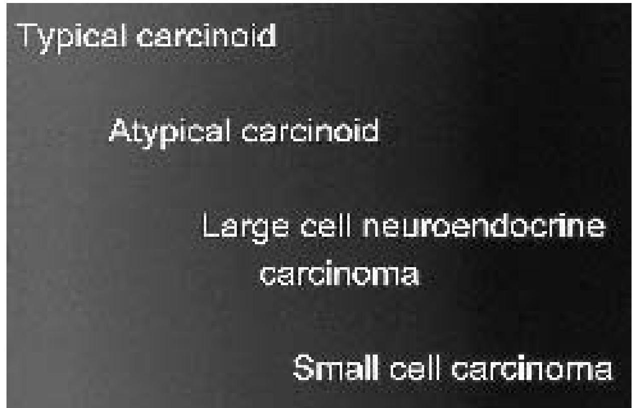


Figure 13 Diagram of neuroendocrine tumours of the lung, excluding non-small cell carcinomas with neuroendocrine features. This diagram does not imply that there is any biological transition from typical carcinoid into atypical carcinoid, etc. It merely depicts the increasing malignancy of these tumours with large-cell neuroendocrine tumours and small cell neuroendocrine carcinoma being the most malignant of the neoplasms.

TC may be central or peripheral. Central tumours give the signs and symptoms mentioned above but nearly 20% of cases may be asymptomatic. Endocrine manifestations are rare but hormones and peptides, including calcitonin, bombesin and cytokines are easily detectable immunohistochemically. This paradox may be due to the small tumour volume and the efficiency of endothelial cells in detoxicating such products. The tumours cause haemorrhage, due to the presence of many small blood vessels. These vessels are caused by TGF (transforming growth factor)-α formation by the tumour. Macroscopically they are yellow/white (**Figure 14; see colour plate section**) with foci of haemorrhage. Up to 30% of tumours contain bone, due to elaboration of osteogenic cytokines, including TGF-β and IGF-1.

TC has trabecular, insular or acinar growth patterns (**Figure 15; see colour plate section**). Other histological patterns have been documented. There should be fewer than two mitoses/2 mm^2/10 high-power fields (hpf) and no necrosis. These tumours may show cytological atypia, especially on frozen section (Sheppard, 1997), increased cellularity and lymphatic invasion. If there are tumour deposits in lymph nodes, the lesion may still be

classed as typical if the histology is as described above. Special stains are not always necessary to confirm the diagnosis, since the histological pattern may be distinctive. Ultrastructurally they contain dense core neurosecretory granules, typical of NE cells.

AC may be central or peripheral, have the NE architecture as described above, and a mitotic rate of 2–10/ 2 mm^2/10 hpf and/or punctate necrosis.

LCNEC are usually peripheral, tan tumours with focal necrosis and haemorrhage. Histologically there is a neuroendocrine appearance, with large cells, irregular pleomorphic nuclei and low nuclear/cytoplasmic ratios. The cytoplasm is abundant, granular and eosinophilic. Necrosis is usually geographic and mitoses are frequent, with a mean of 75 per 2 mm^2 per 10 hpf. Prognosis is poor.

SCLC ('oat cell carcinoma') is described with malignant tumours.

Bax, *bcl-2* and *p53* in Neuroendocrine Tumours

The *bax* gene is one of the main effectors of apoptosis and can be considered a tumour suppressor gene. *Bcl-2* is an oncogene and can block both p53-mediated and p53-independent apoptosis. There is an inverse correlation between the scores of *Bax* and *bcl-2* expression in NE tumours. A predominant *Bax* expression is seen in low-grade NE tumours (TC or AC) and mainly *bcl-2* expression in small-cell and large-cell lung cancers. The p16-retinoblastoma pathway is normal in typical carcinoids but abnormal in the higher-grade NE tumours. (Dosaka-Akita *et al*., 2000) P53 mutation or stabilization is absent in TC. AC may show focal (less than 10%) or patchy P53 positivity, are more aggressive and have significantly shorter survival times than those without P53 staining (Brambilla and Brambilla, 1999).

Carcinoma of Lung

Epidemiology

The lung is the commonest site of cancer worldwide and is in first place in all areas of Europe and North America. Lung cancer appears to be rising in incidence by 70% in women and 30% in men (Travis *et al*., 1995). In the USA the largest percentage increases in age-adjusted rates are for small-cell carcinoma and adenocarcinoma (60% each) with a smaller change for the latter tumour (14%). In the EEC, lung cancer accounts for 21% of all cancer deaths in men and the corresponding figure for women is 4%. This latter figure is increasing and worldwide lung cancer is the fifth most frequent cancer in women. This is due to the social acceptability of smoking, the targetting of women in tobacco advertisements and appetite suppression caused by cigarettes. A decline in incidence in all cell types is predicted soon in the USA. The rate in the West will probably continue to decrease balanced by an increase the Third-world countries, now targetted by tobacco companies. This proposed decrease does not take into account any drug habits of teenagers.

Tobacco is associated with most of the major histological types of lung cancer. The increase in lung cancer risk according to the number of cigarettes smoked appears stronger for squamous and small cell carcinomas than for adenocarcinoma. The increase in this latter cell type has yet to be explained. In the USA there has been a decline in tar and nicotine levels due to filter-tipped cigarettes. There has been a progressive introduction of specially processed tobacco and perforated cigarette paper. These changes may have altered the types of carcinogens in the cigarette smoke, affecting the histological type.

Aetiology

Cigarette Smoking

Cigarette smoking (see earlier) is the most important cause of lung cancer. With an increase in sales of cigarettes at the beginning of the 1900s, changes were made in the types of tobacco used. The smoke was milder and thus easier to inhale. pH was adjusted so that absorption did not occur across the oral mucosa. Subjects had to inhale smoke into the lung to absorb substantial amounts of nicotine. This deep inhalation and absorption of the tobacco's toxic and carcinogenic substances enhanced the lung cancer risk. The evidence linking cigarette smoking with lung cancer started to accumulate in the 1930s. Four retrospective studies showing the relationship between lung cancer and smoking were published in the 1950s, each showing a consistent statistically significant association. Relative risk increases in a stepwise fashion with the increased number of cigarettes smoked. The age at which the subject begins to smoke is also critical.

Other types of tobacco inhalation, ranging from pipes and cigars in the West to bidis in Asia, also correlate with a significant risk for lung cancer.

Involuntary (Passive) Smoking

This is a combination of sidestream smoke, emitted into the air from a burning cigarette between puffs and some of the mainstream smoke, exhaled by the smoker. The potential carcinogenicity of passive smoking was highlighted by the increased risk of lung cancer in nonsmoking women married to smokers. It is now estimated that in the USA at about 2000 deaths per year are due to passive smoking.

Genetic Predisposition

Some 85–90% of smokers who are consuming 20 or more cigarettes a day will not develop lung cancer. Therefore, host factors must be important in altering the risk/predisposition to the development of this disease. There is

mounting evidence that some of the genetic changes predisposing to lung cancer are inherited in a mendelian character. First-degree relatives of lung cancer relatives have a 2.4-fold increased risk of lung cancer or other nonsmoking-related cancers. Patients treated for lung/laryngeal cancer have an increased risk of developing second lung tumours. More tangible evidence of linkage between hereditary and lung cancer has been shown in relatives of patients with retinoblastoma, with a 15-fold risk of lung cancer developing in carriers. Lung cancer is seen in some families with Li–Fraumeni syndrome. This is an inherited cancer-prone condition, due to mutation of the *p53* tumour-suppressor gene. Both the *RB* and the *p53* gene are mutated or inactivated in most small-cell and non-small-cell lung cancers.

Occupation

This is a complex subject since employees may be exposed to more than one potentially carcinogenic substance. The proportion of lung cancer attributable to a given occupational exposure has been estimated at 10–15%. Tobacco smoke acts as a strong confounder in the association.

A prime example of industrially induced cancer is asbestos. Cigarette smoke and asbestos have a multiplicative effect in increasing the incidence of lung cancer.

Metals and Gases

Arsenic and its compounds, chromates, nickel, beryllium and cadmium are associated with an excess of lung cancer deaths.

Hydrocarbons, derived from coal or petroleum and polycyclic aromatic hydrocarbons, such as dibenzanthracene and benzo[*a*]pyrene, are known carcinogens. Exposure to hydrocarbons with increased lung cancer risk is seen in coke oven workers, gas-house workers and aluminum workers, exposed to pitch volatiles (tar). Car and bus exhausts also contain hydrocarbons and truck and bus drivers have an increased risk of lung carcinoma.

Radiation

The increased risk of lung cancer in radiation was first shown in Schneeberg due to radon gas in the mines. Radon is a decay product of naturally occurring uranium. An increased risk of lung cancer is seen in uranium miners in Colorado and fluorspar miners in Newfoundland.

Pulmonary Fibrosis

An increased incidence of lung cancer complicates interstitial pulmonary fibrosis or other significant causes of pulmonary fibrosis and COPD. However, the earlier concept of pulmonary scar cancer is no longer regarded as valid since it has been shown that the fibrosis is a reaction to the tumour and not the cause. The increase in type III collagen in lung cancer is a host response (Madri and Carter, 1984).

Viruses and Cancer

The role of Herpes virus in the causation of papillomatosis has already been mentioned. Human immunodeficiency virus (HIV) infection is only seen as a cause of lung cancer in young individuals with a history of heavy tobacco smoking as well as a moderately advanced immunodeficiency status (Tirelli *et al*., 2000). These patients were considerably younger than most patients with carcinoma at this location. Lymphoepithelioma-like carcinoma of lung is EBV-related in Asian patients, but not usually in Caucasians. This tumour is not usually related to cigarette smoking.

Chromosomal Changes in Lung Cancer

SCLC shows deletions on the short arm of chromosome 3, but this loss has also been observed in non-small-cell lung cancer.

Oncogenes regulate normal growth and development, but when activated promote tumour formation. Dominant oncogene mutations are seen in the *myc* family in SCLC and in the *ras* family in NSCLC (non-small-cell lung cancer). The *myc* family encode nuclear phosphoproteins, which bind to DNA and probably have transcriptional regulatory functions. Amplification of c-, N- and L-*myc* has been detected in SCLC.

The *ras* family of proto-oncogenes consists of H-*ras*, K-*ras* and N-*ras*. These encode related 21-kDa membrane-associated proteins, which probably have a role in transduction of growth signals. *Ras* gene mutations are found in 20–40% of NSCLC, especially adenocarcinomas. They are not seen in SCLC. K-*ras* mutations in NSCLC tumours is associated with shortened survival. K-*ras* mutations correlate with smoking in lung adenocarcinoma. This suggests that exposure to the carcinogens in tobacco smoke activate K-*ras* mutations.

The C-*ErbB-2* proto-oncogene encodes a transmembrane tyrosine-specific protein kinase, p185 neu, which acts as a putative growth factor receptor. Overexpression of p185 neu is common in NSCLC and adenocarcinoma and is associated with shortened survival.

p53 and the retinoblastoma (*RB*) gene are tumour suppressor genes. *RB* gene is always mutated in SCLC but in only 20% of NSCLC. p53 acts as a suppressor of cell division. p53 mutations are seen in 75% of SCLC and 50% of NSCLC. The commonest mutation is a guanine to thymine transversion.

Premalignant Lesions

Three types of the preinvasive epithelial lesions are documented in the new WHO classification. These are

1. squamous dysplasia and carcinoma *in situ*;
2. atypical adenomatous hyperplasia (AAH);
3. diffuse idiopathic pulmonary neuroendocrine hyperplasia (DIPNECH).

Preinvasive Squamous Lesions of the Bronchi

There is a great impetus in the United States and Japan, and beginning in Europe, to detect early bronchial carcinomas, so they may be resected early, hopefully producing a cure. LIFE (laser imaging fluorescence endoscopy) uses the different characteristics of malignant and premalignant tissues to enhance the bronchoscopist's ability to detect small neoplastic bronchial lesions. In high-risk groups, such as smokers, it is six times more sensitive than white light bronchoscopy in detecting preneoplasia (Lam *et al*., 1998). One study spawned from this technique showed deletions of 8p21–23 commenced early during the multistage development of lung cancer, at the hyperplasia/metaplasia stage in smokers without cancer. Allelic deletions persisted for up to 48 years after smoking cessation (Witsuba *et al*., 1999).

Auerbach *et al*. (1957) demonstrated a range of changes in smokers and patients with carcinoma of the lung, explaining the above molecular studies. These changes ranged from loss of normal ciliated lining cells with basal cell hyperplasia, low columnar nonciliated epithelium or squamous metaplasia and increasing degrees of dysplasia (**Figure 16; see colour plate section**) through to carcinoma *in situ* and invasive carcinoma. They meticulously blocked out entire bronchial trees and showed all these changes co-existed in any one cancerous lung.

Basal (reserve cell) hyperplasia, immature squamous hyperplasia and squamous metaplasia may be misdiagnosed as dysplasia by the unwary. As with any other dysplastic lesion elsewhere in the body, the diagnosis of preneoplasia should be made with great caution in the presence of active chronic inflammation, as well as adult respiratory distress syndrome (ARDS) (Hasleton and Roberts, 1999). Previous radiotherapy and/or chemotherapy may cause misdiagnosis. These two latter treatment modalities should be suspected if there are large, bizarre nuclei and plentiful cytoplasm.

Atypical Adenomatous Hyperplasia (AAH)

AAH is defined by the WHO as a focal lesion, often 5 mm or less in diameter, in which the involved alveoli and respiratory bronchioles are lined by monotonous, slightly atypical cuboidal to low columnar epithelial cells with dense nuclear chromatin, inconspicuous nucleoli and scant cytoplasm (**Figure 17; see colour plate section**). The size of the lesion, which is arbitrary, distinguishes it from bronchioloalveolar carcinoma (BAC). Other features delineating AAH from BAC are given below in the differential diagnosis.

Geographical Differences in Incidence of AAH and BAC

There appears to be a geographical difference in the incidence of AAH and BAC between Japan and the UK. Many papers relating to these two lesions originate from the East, suggesting there are biological variations in the incidence of lung cancer between different continents. However, AAH has been described in the UK (Kerr *et al*., 1994). There appears to be an increase in adenocarcinomas, both in the East and the USA (Travis *et al*., 1995). This is not mirrored in the UK, where squamous cell carcinoma remains the predominant cell type in lung cancer.

Clinical and Pathological Considerations

AAH is asymptomatic and is an incidental pathological finding at lobectomy or pneumonectomy (Miller, 1990). It is found incidentally in up to 10% of surgically resected lungs for carcinoma (Kitamura *et al*., 1996). AAH may co-exist with squamous carcinoma of the lung and metastatic colonic and renal cell carcinomas, but is seen most commonly with primary pulmonary adenocarcinoma.

The lesions are minute (smaller than 5 mm) white lesions, where the airways may still be identified, after magnification. Histologically there is a uniform proliferation of atypical cuboidal or low columnar epithelial cells with no mitoses growing along the alveolar septa. There are no admixed ciliated or mucus-secreting cells. There is variation in the cell density, with larger nuclei having increased variation in size, shape and nuclear hyperchromatism. Eosinophilic intranuclear inclusions may be seen. These are not specific for neoplastic cells, being present in reactive alveolar epithelium at times. The cells are ultrastructurally Type II pneumocytes or Clara cells. The alveolar septa may be thickened and infiltrated by lymphocytes.

Immunohistochemistry

Urine protein 1, which is identical with Clara specific 10-kDa protein, may be expressed in 70% of overt BAC but not in AAH lesions (Kitamura *et al*., 1999). Ki-67, a marker of cell proliferation, has demonstrated AAH has a proliferative activity intermediate between normal cells and adenocarcinoma (Kitaguchi *et al*., 1998). A useful marker may be cyclin D1. The frequency of lesions with this cell cycle control protein showing overexpression is high in AAH cycle (47–89%) but decreased in 'early' adenocarcinoma (28%) and overt adenocarcinoma (35%) (Kurasono *et al*., 1998).

Molecular Studies

Both 3p and 9p deletions may be found in certain populations of AAH cells. In another study, 3p, 9p and 17p showed loss of heterozygosity (LOH) in 18, 13 and 66% of AAH cases, respectively. The corresponding carcinomatous lesions showed LOH in 67, 50 and 17%, respectively. This suggests that AAH lesions with moderate or severe atypia are a preneoplastic stage of lung adenocarcinoma.

There is allelic loss of tumour suppressor gene loci in the tuberous sclerosis complex-1 (TSC-1)-associated regions on the long arm of chromosome 9 (9q) and on the short arm of chromosome 16 (16p) in human lung carcinoma; 24% of adenocarcinomas show LOH on 9q and partial LOH on 9q.

The incidence of associated AAH is significantly higher in adenocarcinomas harbouring a partial LOH in the TSC-1-associated regions. These results suggest that TSC-1-associated regions are new candidate loci for tumour-suppressor genes in lung adenocarcinoma, especially when accompanied by multiple AAH lesions (Suzuki *et al*., 1998).

Differential Diagnosis

In a small biopsy AAH or BAC may be part of an invasive adenocarcinoma or an intrapulmonary metastasis. The distinction of AAH from BAC has been considered above. AAH and multiple synchronous lung cancers may be part of the Li–Fraumeni syndrome, due to a constitutional mutation of the *p53* gene (Nadav *et al*., 1998).

Alveolar or papillary adenomas may be misdiagnosed as AAH on small biopsy specimens. These two entities are solitary tumours.

Honeycomb lung (end-stage interstitial fibrosis) may show bronchiolar epithelial proliferation but the distribution of the epithelium is irregular. The epithelium does not grow along slightly thickened uniform alveoli with an intact architecture, as in BAC, but is part of an extensive fibroblastic process.

Organizing pneumonia is associated with Type II cell hyperplasia but there is intra-alveolar fibrosis.

Atypical type II proliferations suggest radiotherapy or chemotherapy-induced change, as with Busulphan or adult respiratory distress syndrome. In cytology specimens ARDS shows bizarre type II or bronchiolar cells. Typically the sheets or balls of cells in BAC are composed of monotonous cells, uniform in size and shape. They have round to oval nuclei and a finely granular chromatin. In ARDS there is group to group variability of cells, often with marked variation in the nuclear/cytoplasmic ratio.

Alveolar hyperplasia can be seen in lymphangioleiomyomatosis (Lantuejoul *et al*., 1997).

Prognosis

There was no difference in the age, follow-up interval or survival rate in one study of 13 patients with no detectable nodules preoperatively compared with cases showing nodules on CT scan.

Some cases of BAC have a good prognosis (Logan *et al*., 1996). The 5-year survival rate in stage 1 disease was 72.9% and stage 2 60.6%, falling to 27.1% in stage 3a and 0% in stages 3b and 4. The 5-year survival rate was higher (64.6%) in AAH associated with well-differentiated adenocarcinoma, as opposed to the presence of this lesion in adenocarcinoma with pulmonary metastases.

Tumourlets

Tumourlets are micronodular, neuroendocrine cell proliferations extending beyond the bronchial/bronchiolar walls, forming aggregates with organoid and nested patterns and measuring less than 0.5 cm in diameter. This measurement, like that of AAH, is purely arbitrary. Any lesion larger than 0.5 cm is diagnosed as a pulmonary carcinoid. Tumourlets are often associated with fibrosis in the walls of bronchiectatic cavities but are uncommon in interstitial pulmonary fibrosis, for some unknown reason. A possible explanation is the prominent inflammation seen in bronchiectasis, whereas in interstitial pulmonary fibrosis inflammation is an early event and is usually less prominent than the fibrosis at the time of biopsy. The differential diagnosis is considered under DIPNECH below but the unwary on a biopsy may suggest small cell lung carcinoma. It does not have the mitotic activity or the lack of cytoplasm of this tumour.

Diffuse Idiopathic Neuroendocrine Hyperplasia (DIPNECH)

DIPNECH is a proliferation of neuroendocrine cells limited to the bronchiolar epithelium (**Figure 18; see colour plate section**). There are increased numbers of scattered single cells, small nodules (neuroendocrine bodies) or linear proliferations of neuroendocrine cells within the bronchiolar epithelium. It is typically associated with obliterative bronchiolar fibrosis and co-existing interstitial or airway fibrosis or inflammation should be absent. This is because, especially with inflammation, there is a proliferation of NE cells.

DIPNECH is a rare condition and may be part of a diffuse neuroendocrine proliferation or identified adjacent to a peripheral carcinoid (Miller and Muller, 1995). DIPNECH may present as an interstitial disease (Armas *et al*., 1995). Rare cases may show the full gamut with DIPNECH, tumourlets and peripheral typical carcinoid tumours (Miller *et al*., 1978).

Differential Diagnosis

Minute meningiothelioid nodules are perivenular, interstitial aggregates of small regular cells with no airway contact. The cells often have a lobular, whorled arrangement, reminiscent of chemodectomas. Meningiothelioid nodules co-express vimentin and epithelial membrane antigen (EMA). DIPNECH is positive with neuroendocrine markers, such as NCAM (neural cell adhesion molecule), chromogranin and synaptophysin.

Carcinoma of Lung

All lung carcinomas show histological heterogeneity. If one accepts that there is a common stem cell, consequently there are mixtures of squamous and adeno or small cell carcinoma. However, major heterogeneity is found in only 5% of cases. The heterogeneity explains why after chemotherapy for SCLC a squamous or adenocarcinoma may develop at the site of the original tumour.

Synchronous tumours arise at the same time as a primary, while a metachronous one consists of a second

tumour occurring after resection of the initial neoplasm. The incidence of synchronous and metachronous tumours is difficult to elucidate from the literature but approximately 1% of common lung neoplasms are multiple.

Classification of Lung Tumours

The WHO has now produced a new lung and pleural tumour classification (1999). The main tumour variants are given in **Table 1**.

Clinical Presentation Due to Local Disease

Central tumours cause obstructive symptoms, including cough, haemoptysis, wheezing and stridor, similar features to benign neoplasms. The haemoptysis is caused by bronchial ulceration but in a few cases there may be infiltration into a large pulmonary artery, causing life-threatening haemorrhage. In addition, there may be bone pain, dyspnea, clubbing and dysphagia. More than 90% of patients with lung cancer are symptomatic at presentation, reflecting advanced disease. This is especially the case with SCLC, which is assumed to be a systemic disease at presentation by most oncologists. However, some patients have stage I disease and in some centres have been offered surgery and chemotherapy, although the role of surgery has yet to be established in patients with more extensive nodal disease (Lucchi *et al*., 1997). Patients with general systemic symptoms or metastases in both small cell and non-small-cell lung cancer have a poor prognosis.

Pancoast tumours (superior sulcus tumours) are localized, arising posteriorly at the apex of the upper lobe near the brachial plexus. They infiltrate C8, T1 and T2 nerve roots, causing pain, temperature changes and muscle atrophy in the shoulder and arm innervated by these nerve roots. Horner's syndrome is caused by involvement of the sympathetic chain and stellate ganglion, giving unilateral enophthalmos, ptosis and miosis. Superior vena caval obstruction presents as oedema and plethora of the face, as well as dilated neck and upper torso veins. Hoarseness is due to recurrent laryngeal nerve entrapment, seen more frequently in left upper lobe tumours, as the left recurrent laryngeal nerve loops around the aortic arch. Tumour can involve the phrenic nerve, paralysing a hemi-diaphragm. The oesophagus may be infiltrated causing dysphagia and, if the pleura is involved, an effusion occurs.

Metastases are common in SCLC, with 20% of cases metastatic at presentation. Squamous cell carcinoma tends to remain intrathoracic whereas adeno- and large-cell carcinoma show metastases to regional nodes, liver, gut, adrenals, central nervous system and bone.

Table 1 Classification of lung tumours

Epithelial tumours
- Benign
  - Papilloma
  - Adenoma
- – alveolar and papillary
    - – salivary gland type
    - – mucous gland
    - – pleomorphic
    - – mucinous
- Preinvasive lesions
  - Squamous
  - Atypical adenomatous hyperplasia
  - Diffuse idiopathic pulmonary neuroendocrine cell hyperplasia
- Malignant
  - Squamous cell carcinoma (and variants, including basaloid))
  - Small-cell carcinoma (and variants)
  - Adenocarcinoma (and variants)
  - Large-cell carcinoma (and variants, including large-cell neuroendocrine carcinoma and lymphoepithelial carcinoma)
  - Adenosquamous carcinoma
  - Carcinomas with pleomorphic, sarcomatoid or sarcomatous elements (including carcinosarcoma and pulmonary blastoma)
  - Carcinoid tumour – typical
    – atypical
  - Carcinomas of salivary gland type (including mucoepidermoid and adenoid cystic carcinomas)
  - Unclassified
  - Soft tissue tumours

Mesothelial tumours
Miscellaneous tumours (including hamartoma and sclerosing haemangioma)
Lymphoproliferative disorders
Secondary tumours

Paraneoplastic Syndromes

This term identifies symptoms and signs secondary to cancer, occurring at a site distant from the tumour or its metastases. They are caused by the production of products, such as polypeptide hormones, hormone-like peptides, antibodies, immune complexes, etc., by the tumour. Nonmetastatic hypercalcaemia is commonest in squamous cell carcinoma. The squamous carcinoma cells secrete a parathormone-related protein (PTH-rP), which shows a limited sequence homology with parathyroid hormone. Cushing's syndrome, the commonest, is due to ectopic ACTH production, usually seen in SCLC. The syndrome of inappropriate antidiuretic hormone secretion (SIDDH) is seen mainly with SCLC. In half the cases there is ectopic vasopressin secretion from the tumour. In the remainder there is abnormal release of this peptide from the posterior pituitary because of altered or defective chemoreceptor control. Gynaecomastia develops because of increased levels of β-hCG (human chorionic gonadotrophin) produced by the tumour. This hormone is most commonly seen in germ cell tumours of the mediastinum and gonads and very rarely as a primary pulmonary tumour.

Neurological Syndromes

These are most often associated with an autoimmune reaction, in which the tumour shares antigens with normal nervous tissue. In recent years, antibodies reactive with nuclear and cytoplasmic antigens and neurons throughout the CNS and peripheral ganglia have been identified in patients with these syndromes. The Lambert–Eaton myasthenic syndrome is uncommon and is associated with SCLC. There is muscle weakness and fatigue and it is most pronounced in the pelvic girdle and thighs. There may be dysarthria, dysphagia and blurred vision. Hypertrophic osteoarthropathy is characterized by finger clubbing, periosteal bone formation (see below) and arthritis. Other neurological syndromes include autonomic neuropathy and subacute sensory peripheral neuropathy.

The enlargement of the distal phalanx of the digits due to an increase in the connective tissue in the nail bed is termed clubbing and is almost always a feature of hypertrophic osteoarthropathy. It is seen especially in squamous and adenocarcinomas but has also been documented in other diseases, including congenital cyanotic heart disease and cystic fibrosis. It may be neurogenic (vagally mediated), hormonal owing to a high oestrogen or growth hormone or vascular owing to arterio-venous shunts.

For treatment purposes lung carcinomas are divided into SCLC and NSCLC, the former only rarely having surgery. The commonest non-small-cell carcinomas are squamous, adeno- and large cell.

Squamous Cell Carcinoma

These are often central, in main or segmental bronchi, or peripheral. They may show an endobronchial growth pattern or infiltrate between the cartilaginous rings initially, in time destroying them to invade surrounding tissue. The tumour is solid, grey/white but may show cavitation (**Figure 19; see colour plate section**). Such cavity formation may cause clinical misdiagnosis as apical cavitating tuberculosis. Very rarely both diseases may co-exist. There is often related bronchiectasis and obstructive pneumonitis, due to bronchial obstruction. The pneumonitis causes both radiological and macroscopic problems in definition of the true extent of the tumour. This problem may be accentuated if there is a prominent inflammatory component to the tumour.

Adjacent to the tumour there may be carcinoma *in situ* as well as chronic bronchitic changes. In addition, there is squamous metaplasia and reserve cell hyperplasia. The tumour shows varying degrees of squamous differentiation, lying in a fibrous stroma with a varying amount of acute and chronic inflammation. Foci of recent and old haemorrhage are seen. Well-differentiated tumours show keratin pearls and keratinization (**Figure 20; see colour plate section**). Intercellular bridges should be sought and paradoxically can be best identified with a mucin stain. Rarely mucin can be seen in the malignant cells but more than 10% of the tumour should show a distinctive glandular component before an adenosquamous carcinoma is diagnosed. Tumour giant cells imply a poor prognosis. Clear cell change, due to glycogen accumulation, is also a feature of some tumours but has no prognostic connotation. The nuclei are hyperchromatic and may show prominent nucleoli. The tumour may encircle the bronchus and a fibre-optic biopsy may only reveal fibrosis.

Squamous cell carcinomas manifest earlier than other types of pulmonary malignancy because of obstructive symptoms. They may involve lymph nodes by direct spread. There may be lymphatic and vascular invasion, but the prognostic significance of vascular invasion in non-small-cell carcinoma is uncertain. Some studies suggest that vascular invasion indicates a poor prognosis, others show it has no effect on survival, since the tumour-associated desmoplasia causes occlusive intimal fibrosis in thin walled pulmonary arteries and veins.

A small subgroup of squamous carcinomas is termed basaloid carcinoma. This tumour grows exophytically in proximal bronchi. Mediastinal pleura or adipose tissue are often invaded. There is a solid lobular or anastamotic trabecular pattern with small, moderately pleomorphic, cuboidal or fusiform cells. There is peripheral pallisading and a high mitotic rate. Other subtypes of squamous carcinomas are clear cell, not to be confused with adenocarcinoma, papillary, pleomorphic and spindle cell and small cell variant of squamous.

Differential Diagnosis

Conditions to be considered in the differential diagnosis are florid squamous metaplasia in the bronchus or states associated with an inflammatory process, dysplasia, which if high grade can be very difficult to distinguish on a small biopsy, any benign lung tumour with squamous metaplasia on the surface, especially granular cell myoblastoma, benign squamous cell papilloma, tumours with a squamous component, such as carcinosarcoma, spindle cell sarcomas, metastatic tumours with a squamous appearance, including sarcomas, which can be primary or secondary, and mesothelioma. Special stains, including mucins and cytokeratins, are often helpful in determining the cell of origin and in some cases the site.

It may be difficult with poorly differentiated tumours to differentiate squamous from adenocarcinoma. In such cases if mucin stains are diffusely negative the term 'non-small cell carcinoma' is used. This enables medical oncologists to give appropriate treatment, as they need to distinguish this tumour from small-cell lung carcinoma.

Treatment and Prognosis

Treatment depends on the stage of the disease. Cases with stage IIIA disease or less are treated by surgery though the role of surgery has yet to be established in N2 and N3 disease. Other options for more advanced disease are radiotherapy, either direct beam or intraluminal, brachytherapy and chemotherapy.

Prognosis depends on the stage. Five-year survival rates are stage I, 50%, stage II, 30% and stage IIIA, approximately 10% (Mountain, 1988).

Adenocarcinoma

It is often impossible to distinguish a bronchial primary from secondary pulmonary tumour. If an adenocarcinoma is thought to be secondary, a primary should be sought in the stomach, colon, pancreas, breast, ovary, prostate or kidney. Special stains, such as cytokeratins 7 and 20, TTF-1 (thyroid transcription factor 1, but commonly expressed in lung cancer), PSA (prostate-specific antigen), ER (oestrogen receptor) and PR (progesterone receptor) may help in identifying the site of the primary tumour.

Adenocarcinomas are usually peripheral, well-circumscribed masses (**Figure 21; see colour plate section**). If the pleura is involved there is fibrosis and puckering. These tumours may be central, arising from bronchial mucous glands. The peripheral tumours usually occur in females while the bronchial gland type tends to have a male predominance. There is no significant survival difference between the two variants. Adenocarcinomas vary in size and may occupy an entire lobe. They sometimes contain carbon pigment and may show marked scarring. The term 'scar carcinoma' is no longer used, since the stroma is a desmoplastic response to tumour rather than arising in a pulmonary scar (pulmonary fibrosis). Adenocarcinomas may be single or multiple and this may create confusion with metastases. Pleural seeding is common and may mimic a mesothelioma.

Histologically there are different growth patterns with tubular, papillary, acinar, signet ring and solid variants (**Figure 22; see colour plate section**). It is common to have a mixture of the above patterns in any one tumour. The cells are large, polygonal and tend to be discohesive with a high nuclear/cytoplasmic ratio. Some cases have mucin vacuoles. However, 30% of adenocarcinomas show no mucin on special stains. Spindle cell and giant cell foci may be identified. The tumour spreads aerogenously and may show a peripheral bronchioloalveolar pattern. This should not be diagnosed as bronchioloalveolar carcinoma. It rapidly invades lymphatics, blood vessels and the pleura and spreads to distal sites.

Differential Diagnosis

This includes any secondary adenocarcinoma, as detailed above, adenosquamous carcinoma, muoepidermoid tumour of the bronchus, germ cell tumours, carcinoid tumours with glandular foci, pulmonary blastoma and in the case of a pleural biopsy, mesothelioma. The stains mentioned above will help to differentiate *some* primary pulmonary tumours from secondary ones.

Trainee pathologists may misdiagnose benign lesions, such as radio or chemotherapy-induced change, type II cell hyperplasia associated with interstitial pulmonary fibrosis and resolving ARDS, alveolar adenoma, sclerosing haemangioma and atypical adenomatous hyperplasia.

Treatment and Prognosis

Surgery is the most appropriate treatment option. Survival is related to stage, stage I has a 50% 5-year survival rate, stage II approximately 25%, and stages IIIA and B, less than 10% (Mountain, 1988). These figures are pre-CT scans and the Japanese literature suggests they are detecting earlier tumours with this modality.

Bronchioloalveolar Carcinoma (BAC) (Alveolar Carcinoma)

This is a subtype of adenocarcinoma. It is the commonest tumour in North American women. There are greyish/white nodules with a central scar in some cases. If close to the pleura there is puckering and fibrosis, as in adenocarcinoma, but the tumour does not usually infiltrate this layer. Mucinous tumours have a glistening appearance. The tumour may involve an entire lobe (**Figure 23; see colour plate section**) or lung, often giving bronchorrhoea.

There are two main subtypes, mucinous and nonmucinous. Nonmucinous consist of Clara or type II cells and mucinous have goblet or mucin-producing cells. Both grow along intact alveolar walls as a single layer or occasionally forming papillae. If stromal invasion is identified, the tumour is classified as adenocarcinoma. In nonmucinous BAC, the cells are cuboidal with an eosinophilic, ciliated cytoplasm and prominent nuclei. In mucinous BAC there are tall columnar, mucinous cells that are well differentiated and lack cilia. The nuclei are uniform and show varying degrees of hyperchromasia. The cells produce much mucin. Discrete satellite nodules are seen in either type.

The differential diagnosis is similar to adenocarcinoma as detailed above. It is impossible on a small biopsy to differentiate invasive adenocarcinoma from BAC or AAH.

Treatment and Prognosis

The optimum treatment is resection and the 5-year survival rate is 42.1%, with a higher survival rate for stage I disease.

Small-cell Lung Carcinoma (SCLC)

This tumour grows rapidly and may occasionally present as metastases without any visible primary tumour. A primary tumour in the upper respiratory tract, oesophagus or cervix may metastasize to the lung. This tumour is responsive to chemotherapy. It presents as hilar masses with extension into lymph nodes (**Figure 24; see colour plate section**). The tumour is soft, white and shows extensive necrosis. In advanced cases the bronchial lumen is obstructed by extrinsic compression. There are two main histological variants.

The first is 'classical' small cell carcinoma (**Figure 25; see colour plate section**) and the second a combined

small cell carcinoma. In classical SCLC there are sheets of small, hyperchromatic nuclei with nuclear moulding and little cytoplasm. There is a high mitotic rate. Because of the high cell turnover and necrosis, DNA leaches out and is taken up by blood vessels ('Azzopardi effect'). A combined SCLC has the above pattern and adeno-, large-cell or squamous carcinomatous components.

This tumour stains positively with NCAM (neural cell adhesion molecule), synaptophysin and chromogranin. Lymphovascular invasion occurs early and distant metastases are common. These are seen in bone marrow, liver, kidney, adrenals, cerebrum, cerebellum, meninges, regional and cervical lymph nodes. After chemotherapy, recurrent tumour may be predominantly squamous or adenocarcinoma.

Differential Diagnosis

Benign lesions may be confused with SCLC, especially if there is crush artifact. In the presence of this, a diagnosis of SCLC should never be made, since any chronic inflammatory condition, such as tuberculosis, may have crush artifact. Tumours that may be confused with SCLC are tumourlets, other neuroendocrine tumours, on a small biopsy, squamous, large-cell and adenocarcinomas, which may have an SCLC component, metastatic carcinomas and lymphomas.

Treatment and Prognosis

This is neoplasm is usually treated with chemotherapy, since cases with stage I disease are rare. There is an approximately 10% 5-year survival after treatment.

Large-cell Carcinoma

Large-cell carcinomas are large necrotic masses, which frequently invade the overlying pleura and grow into adjacent structures. This is a diagnosis of exclusion since no acinar or squamous differentation or mucin production are seen. Large-cell carcinomas have sheets and nests of large cells with prominent vesicular nuclei and nucleoli (**Figure 26; see colour plate section**). The cell borders are easily visualized. Necrosis and haemorrhage are frequent and there may be acute and/or chronic inflammation.

There are several variants. These include clear cell carcinoma, composed entirely of clear cells, basaloid carcinoma and lymphoepithelioma (lymphoepithelial-like carcinoma). Lymphoepithelial-like carcinoma has a marked lymphoplasmacytic infiltrate admidst large cell carcinoma. Epstein–Barr virus genome may be identified in this tumour. It is commoner in the Far East.

The differential diagnosis is similar to pleomorphic carcinoma, described below. It includes all the carcinomas, primary or secondary, described above, as well as large cell neuroendocrine carcinoma, primary or metastatic sarcoma, anaplastic large cell lymphoma and melanoma. The prognosis is grim, with an approximately 1% 5-year survival.

Adenosquamous Carcinoma

These are usually peripheral tumours. There must be definite squamous and adenocarcinomatous components, one of which must comprise more than 10% of the tumour.

Pleomorphic Carcinoma of Lung

This entity was first described by Fishback *et al*. (1994), who studied 78 cases of this entity, with a male to female ratio of 2.7 : 1. About 80% of the patients had thoracic pain, cough and haemoptysis but 18% were asymptomatic. Foci of squamous cell carcinoma were present in 8%, large cell in 25% and adenocarcinoma in 45%. The remaining 22% of tumours were completely spindle and/or giant-cell carcinomas. Spindle and giant-cell carcinomas were found together in 38% of patients. Nodal metastases were the most significant single prognostic factor. Only 1% of cases had SCLC foci; such cases are classified as combined small-cell carcinoma.

The malignant spindle cell component consists of fusiform cells with eosinophilic cytoplasm (**Figure 27; see colour plate section**). The spindle cells vary from an epithelioid appearance to a slender banal morphology. The chromatin is frequently vesicular or coarse and hyperchromatic. Giant-cell carcinoma, with multiple tumour giant cells, covered by neutrophil polymorphs (emperopolesis), is a variant of pleomorphic carcinoma. However, tumour giant cells are an integral component of pleomorphic carcinoma. There is a mild to moderate inflammatory component, predominantly lymphocytes. As in mesothelioma, separation of spindle cell carcinoma from a desmoplastic stroma is often difficult.

Molecular techniques have shown that K-*ras*-2 showed mutations in fewer pleomorphic carcinomas than adenocarcinomas. Pleomorphic carcinomas also showed fewer *p53* point mutations than adenocarcinoma or squamous carcinoma. The *p53* point mutations in pleomorphic carcinoma were commoner exon 7 than those of squamous and adenocarcinoma, which were exon 8 (Pryzygodzki *et al*., 1996).

The differential diagnosis is largely described above under Large-cell Carcinoma. In addition, any spindle cell lesion has to be considered. The WHO recommends epithelial markers, such as keratin or epithelial membrane antigen, as useful in confirmation of the carcinomatous differentiation in the spindle cell component. However, if these markers are negative the tumours are classified as pleomorphic carcinomas, assuming the neoplasm does not have the immunoprofile of one of the tumours noted in the differential diagnosis, especially a soft tissue sarcoma. Heterologous elements are necessary for the diagnosis of carcinosarcoma.

Keratin may be positive in synovial sarcoma, mesothelioma, epithelioid haemangioendothelioma, rarely in leiomyosarcoma, anaplastic large cell lymphomas and occasionally in malignant peripheral nerve sheath tumours. Therefore, reliance on a single cytokeratin will cause problems. In such problematic pulmonary tumours a battery of immunostains may be useful.

Pulmonary Lymphomas

Pulmonary lymphoid proliferations may occur in rheumatoid disease and in Sjogren's syndrome. Lymphomas in the lung are classified as lymphoid interstitial pneumonia, non-Hodgkin lymphomas, primary pulmonary Hodgkin disease, plasma cell neoplasms and leukaemic infiltration.

Lymphoid interstitial pneumonia (LIP) consists of a diffuse infiltrate with small lymphocytes and plasma cells affecting large areas of the lung. There may be reactive lymphoid follicles and it may be associated with autoimmune disease and poly- or monoclonal gammopathy. Cases with monoclonal light-chain restriction probably represent 'MALT' lymphomas. Some cases have polytypic light-chain expression and have been classed as non-neoplastic.

Primary Pulmonary Non-Hodgkin Lymphoma

These are similar to non-Hodgkin lymphomas elsewhere, most being B cell in origin. Patients are middle-aged with respiratory and systemic symptoms. There may be lobar consolidation as well as skin, renal and central nervous system involvement. About 50% of patients are asymptomatic, the rest having a variety of respiratory and constitutional symptoms. In the T cell lymphomas there is often a vasculitic (**Figure 28; see colour plate section**) and granulomatous pattern, with necrosis.

Sarcomas

Primary pulmonary sarcomas are rare. Epithelioid haemangioendothelioma is a low-grade vascular sarcoma, occurring more commonly in young women. There is chest pain, dyspnea, mild cough and multiple bilateral nodules. It can also affect liver, bones and soft tissues. The tumour may resemble a mesothelioma. Kaposi sarcoma may be seen in HIV-positive patients and is described elsewhere.

Carcinosarcomas have a mixture of sarcomatous and carcinomatous elements.

Secondary Tumours

The lung is frequently the site of secondary adenocarcinomas and sarcomas, especially osteogenic and chondrosarcoma. Because of advances in chemotherapy it may be beneficial to treat these patients with localized resection. In the case of secondary adenocarcinomas, as mentioned above, TTF-1 and cytokeratins can be useful in determining whether a tumour is primary or secondary.

REFERENCES

Aozasa, K., *et al*. (1995). Polymorphic reticulosis is a neoplasm of large granular lymphocytes with CD3 plus phenotype. *Cancer*, **75**, 894–901.

Armas, O. A., *et al*. (1995). Diffuse idiopathic pulmonary neuroendocrine cell proliferation presenting as interstitial lung disease. *American Journal of Surgical Pathology*, **19**, 963–970.

Auerbach, O., *et al*. (1957). Changes in the bronchial epithelium in relation to smoking and cancer of the lung. *New England Journal of Medicine*, **256**, 97–104.

Batsakis, J. G., *et al*. (1963). Adenocarcinoma of the nasal and paranasal cavity. *Archives of Otolaryngology*, **77**, 625–633.

Bellacosa, A., *et al*. (1996). Cyclin D1 gene amplification in human laryngeal squamous cell carcinomas; prognostic significance and clinical implications. *Clinical Cancer Research*, **2**, 175–180.

Bradford, C. R., *et al*. (1997). P53 mutation as a prognostic marker in advanced laryngeal carcinoma. Department of Veterans Affairs Laryngeal Cancer Cooperative Study Group. *Archives of Otolaryngology and Head Neck Surgery*, **123**, 605–609.

Brambilla, C. and Brambilla, E. (1999). *Lung Tumors; Fundamental Biology and Clinical Management* (Marcel Dekker, New York).

Cecchi, F., *et al*. (1980). Adenocarcinoma of the nose and paranasal sinuses in shoemakers and woodworkers in the province of Florence, Italy (1963–77). *British Journal of Industrial Medicine*, **37**, 222–225.

Digby, K. H., *et al*. (1941). Nasopharyngeal malignancy. *British Journal of Surgery*, **28**, 517–537.

Dokiya, F., *et al*. (1998). Retinoblastoma protein expression and prognosis in laryngeal cancer. *Acta Otolaryngolica*, **118**, 759–762.

Dosaka-Akita, H., *et al*. (2000). Differential retinoblastoma and p16 (INK 4A) protein expression in neuroendocrine tumors of the lung. *Cancer*, **88**, 550–556.

Fishback, N. F., *et al*. (1994). Pleomorphic (spindle/giant cell) carcinoma of lung: A clinicopathologic correlation of 78 cases. *Cancer*, **73**, 2936–2945.

Furuta, Y., *et al*. (1992). Detection of human papilloma virus DNA in carcinomas of the nasal cavities and paranasal sinuses by polymerase chain reaction. *Cancer*, **69**, 353–357.

Gown, A. M., *et al*. (1993). Androgen receptor expression in angiofibromas of the nasopharynx (abstract). *Modern Pathology*, **6**, 81a.

Hasleton, P. S. (1996). *Spencer's Pathology of the Lung*, 5th edn. 896–897 (McGraw-Hill, New York).

Hasleton, P. S and Roberts, T. E. (1999). Adult respiratory distress syndrome: an update. *Histopathology*, **34**, 285–294.

Kerr, K. M., *et al*. (1994). Atypical alveolar hyperplasia: relationship with pulmonary adenocarcinoma, p53 and *C-erb* B2 expression. *Journal of Pathology*, **174**, 249–256.

Kitaguchi, S., *et al*. (1998). Proliferative activity, p53 expression and loss of heterozygosity on 3p, 9p and 17p in atypical adenomatous hyperplasia of the lung. *Hiroshima Journal of Medical Science*, **47**, 17–25.

Kitamura, H., *et al*. (1996). Atypical adenomatous hyperplasia and bronchoalveolar lung carcinoma: analysis of morphometry and the expressions of p53 and carcinoembryonic antigen. *American Journal of Surgical Patholology*, **20**, 553–562.

Kitamura, H., *et al*. (1999). Atypical adenomatous hyperplasia of the lung. Implications for the pathogenesis of peripheral lung adenocarcinoma. *American Journal of Clinical Pathology*, **111**, 610–622.

Krecicki, T., *et al*. (1999). c-erb B-2 immunostaining in laryngeal cancer. *Acta Oto-Laryngolica*, **119**, 392–395.

Kurasono, Y., *et al*. (1998). Expression of cyclin D1, retinoblastoma gene protein and p16 MTS-1 protein in atypical adenomatous hyperplasia and adenocarcinoma of the lung: An immunohistochemical analysis. *Virchows Archiv*, **432**, 207–215.

Lantuejoul, S., *et al*. (1997). Multifocal alveolar hyperplasia associated with lymphangioleiomyomatosis in tuberous sclerosis. *Histopathology*, **30**, 570–575.

Lam, S. T., *et al*. (1998). Localisation of bronchial intraepithelial neoplastic lesions by fluorescence bronchoscopy. *Chest*, **113**, 696–702.

Lewis, J. S. and Castro, E. B. (1972). Cancer of the nasal cavity and paranasal sinuses. *Journal of Laryngology and Otology*, **86**, 255–262.

Logan, P. M., *et al*. (1998). Bronchogenic carcinoma and coexistent bronchoalveolar cell adenomas: assessment of radiological detection and follow-up in 28 patients. *Chest*, **109**, 713–717.

Lucchi, M., *et al*. (1997). Surgery in the management of small cell lung cancer. *European Journal of Cardiothoracic Surgery*, **12**, 689–693.

Madri, J. A. and Carter, D. (1984). Scar cancers of the lung: origin and significance. *Human Pathology*, **15**, 625–631.

Miller, M. A., *et al*. (1978). Multiple peripheral carcinoids and tumorlets of carcinoid type, with restrictive and obstructive lung disease. *American Journal of Medicine*, **65**, 373–378.

Miller, R. R. (1990). Bronchioloalveolar cell adenomas. *American Journal of Surgical Pathology*, **14**, 904–912.

Miller, R. R. and Muller, N. L. (1995). Neuroendocrine cell hyperplasia and obliterative bronchiolitis in patients with peripheral carcinoid tumours. *American Journal of Surgical Pathology*, **19**, 653–658.

Mills, S. E., *et al*. (2000). *Tumors of the Upper Aerodigestive Tract. Atlas of Tumor Pathology* (Armed Forces Institute of Pathology, Bethesda, MD).

Mountain, C. F. (1988). Prognostic implications of the of the International Staging System for Lung Cancer. *Seminars in Oncology*, **15**, 236–245.

Nadav, Y., *et al*. (1998). Multiple synchronous lung cancers and atypical adenomatous hyperplasia in Li Fraumeni syndrome. *Histopathology*, **33**, 52–54.

Narayana, A., *et al*. (1998). Is p53 an independent prognostic factor in patients with laryngeal carcinoma? *Cancer*, **82**, 286–291.

Pryzygodzki, R. M., *et al*. (1996). Pleomorphic (giant and spindle cell) carcinoma cell carcinoma by K-ras-2 and p53 analysis. *American Journal of Clinical Pathology*, **106**, 487–492.

Sheppard, M. N. (1997). Nuclear pleomorphism in typical carcinoid tumours of the lung: problems in frozen section interpretation. *Histopathology*, **30**, 478–480.

Simons, M. J., *et al*. (1976). Immunogenetic aspects of nasopharyngeal carcinoma in young patients. *Journal of the National Cancer Institute*, **57**, 977–980.

Suzuki, K., *et al*. (1998). Loss of heterozygosity in a tuberous sclerosis gene-associated region in adenocarcinoma of the lung accompanied by multiple atypical adenomatous hyperplasia. *International Journal of Cancer*, **79**, 384–389.

Tirelli, U., *et al*. (2000). Lung carcinoma in 36 patients with human immunodeficiency virus infection. The Italian Cooperative Group on AIDS and Tumors. *Cancer*, **88**, 563–569.

Tomita, Y., *et al*. (1996). Non-Hodgkins lymphoma of the Waldeyer's ring as the manifestation of human T-cell leukaemia virus type I-associated lymphoproliferative diseases in South West Japan (abstract). *Modern Pathology*, **9**, 12a.

Travis, W. D., *et al*. (1995). Lung cancer. *Cancer*, **75**, 191–202.

Travis, W. D., *et al*. (1999). *Histologic Typing of Lung and Pleural Tumors*, 3rd edn, (World Health Organization, Springer, Berlin).

Witsuba, I. I., *et al*. (1999). Allelic losses at chromosome 8p 21–23 are early and frequent events in the pathogenesis of lung cancer. *Cancer Research*, **59**, 1973–1979.

FURTHER READING

Brambilla, C. and Brambilla, E. (eds) (1999). *Lung Tumors: Fundamental Biology and Clinical Management* (Marcel Dekker, New York).

Churg, A. and Green, F. H. Y. (1998). *Pathology of Occupational Lung Disease*, 2nd edn. (Williams and Wilkins, Baltimore).

Colby, T. V., *et al*. (1995). *Tumors of the Lower Respiratory Tract. Atlas of Tumor Pathology* (Armed Forces Institute of Pathology, Washington, DC).

Dail, D. H. and Hammar, S. P. (1988). *Pulmonary Pathology*, 2nd edn (Springer, New York).

Gregor, R. T. (1998). *Diseases of the Head and Neck, Nose and Throat* (Arnold, London).

Hasleton, P. S. (1994). Histopathology and prognostic factors in bronchial carcinoid tumours. *Thorax*, **49**, (Suppl.), S56–S62.

Hasleton, P. S. (ed.) (1996). *Spencer's Pathology of the Lung*, 5th edn (McGraw-Hill, New York).

Parkes, W. P. (1994). *Occupational Lung Disorders*, 3rd edn. Park, W. P. (ed.) (Butterworth-Hernemann Ltd, Oxford).

Roggli, V. L., *et al*. (1992). *Pathology of Asbestos-associated Diseases* (Little Brown, Boston).

Samet, J. M. (1994). *Epidemiology of Lung Cancer* (Marcel Dekker, New York).

Websites

The websites given are those of predominantly pulmonary journals, since there is no purely pulmonary pathology journal. These pulmonary journals feature good review articles, especially the *European Respiratory Journal* and the *American Review of Critical Care and Respiratory Care Medicine*. Pathology journals also carry original pulmonary articles. The recommended ones are *Modern Pathology*, *American Journal of Pathology*, *Laboratory Investigation*, *American Journal of Clinical Pathology*, *Histopathology* and *Journal of Pathology*. The *New England Journal of Medicine* has an excellent weekly clinicopathological conference with many pulmonary topics. Most of these journals require subscriptions/passwords, etc., to obtain full text but most libraries will be able to provide access.

American Journal of Respiratory and Critical Care Medicine http://intl-ajrccm.atsjournals.org/.

Chest http://www/chestjournal.org/.

European Respiratory Journal http://195.226.52.174/ers/issue-list.html.

Thorax http://thorax.bmjjournals.com/.

Cancer http://www3/interscience.wiley.com/cgi-bin/jtoc?ID=28741.

Chapter 37

Upper Gastrointestinal Tract

Grant N. Stemmermann, Amy E. Noffsinger and Cecilia M. Fenoglio-Preiser
University of Cincinnati School of Medicine, Cincinnati, OH, USA

CONTENTS

- Introduction
- Normal Development and Structure
- Epidemiology
- Aetiology
- Screening and Prevention
- Pathology
- Treatment
- Molecular Genetic Findings
- Conclusion

INTRODUCTION

Worldwide, oesophageal and gastric cancers are very common tumours, accounting for 6–34% of cancer-related deaths (Aoki *et al*., 1992). Variations in carcinogen exposures account for striking geographical differences in the frequency of both tumours. Oesophageal and gastric cancers have decreased in both incidence and mortality rates in economically prosperous populations but their incidence remains high in less prosperous countries. This chapter will summarize the current status of our knowledge of the origin, pathology, molecular biology and behaviour of upper gastrointestinal tumours. We will also describe methods used to prevent and treat these malignancies.

NORMAL DEVELOPMENT AND STRUCTURE

General Organisation

The gastrointestinal tract is divided into several major anatomical regions, each with its own unique physiological functions and structure. The various regions share a basic structure. The inside has a mucosal lining beneath which is a thin muscular layer known as the muscularis mucosae. The mucosa serves many of the digestive and absorptive functions of the gastrointestinal tract. Beneath the muscularis mucosae is the submucosa. The next layer is the muscularis propria, which is responsible for propelling gastrointestinal contents forward throughout the gut.

Oesophagus

The oesophagus develops from the cranial portion of the primitive foregut, becoming recognizable at the 2.5-mm stage of development (approximately the third gestational week) as an annular constriction located between the stomach and pharynx (Fenoglio-Preiser *et al*., 1999). The oesophagus elongates, growing in a cephalad direction and becoming increasingly tubular. Early in development, the cephalad portions of both the oesophagus and trachea lie within a common tube. When the oesophagus and trachea divide, the oesophagus comes to lie dorsal to the trachea. The oesophageal mucosal lining progresses through a series of epithelial changes before attaining the appearance of the glycogenated, nonkeratinized, stratified squamous epithelium present in the adult.

The mucosal lining of the oesophagus consists of stratified layers of squamous cells resembling those found in the skin (**Figure 1**). These stratified cells show an orderly progression of differentiation from the bottom of the epithelium to its surface. This squamous cell lining regularly renews itself. New cells form from progenitor cells at its base just above the basement membrane. The basement membrane is a linear structure lying under the epithelium; it serves as a boundary between the epithelium and its underlying tissues. Old, nonfunctional cells are shed from the mucosal surface. Newly formed epithelial cells pass upward in the mucosa, becoming increasingly more mature as they do so. At the surface they are held tightly together by intercellular junctions. This mucosal layer protects underlying tissues from damage by abrasion from food passing over it and from damage induced by the chemical contents of material in the oesophageal lumen.

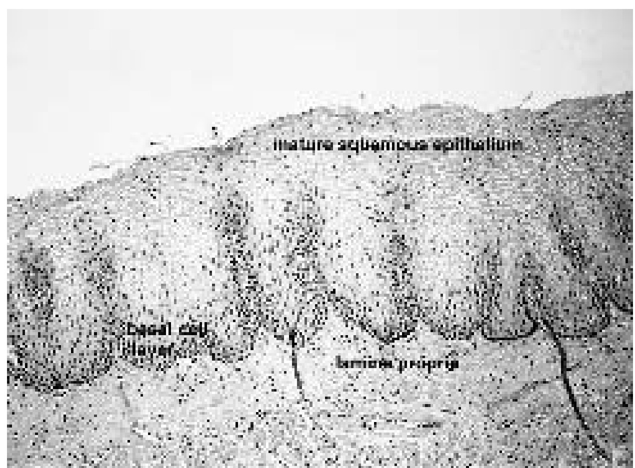


Figure 1 The normal histology of the oesophagus. The oesophagus consists of squamous epithelium arranged in stratified layers. The cells at the bottom (basal layer) are small and these cells represent the proliferative compartment. They give rise to daughter cells which pass upward into the overlying epithelium, becoming increasingly mature as they do so. As a result, the amount of cytoplasm relative to the nucleus, the nuclear : cytoplasmic ratio decreases. The basement membrane separates the basal cell layer from the underlying lamina propria.

The oesophageal lining is lubricated by mucus derived from the salivary glands and from oesophageal submucosal mucus glands. Passage of food and liquid through the oesophagus into the stomach is accomplished by contraction of its thick muscular coat. A muscular sphincter, known as the lower oesophageal sphincter, acts as the gateway to the stomach, controlling the passage of food forward and preventing reflux of gastric acid and digestive enzymes back into the oesophagus.

Stomach

The stomach is a bag-like structure lying between the oesophagus and the intestines. It develops from a fusiform swelling of the foregut at approximately 4 weeks gestation. During the sixth to seventh fetal week the gastric curvature develops. In the ninth week, a diverticulum appears in the upper stomach, which subsequently merges with, and lengthens, the greater curvature. The stomach is initially lined by stratified or pseudostratified epithelium. Later, it is replaced by cuboidal cells. As secretions accumulate, droplets and vacuoles coalesce to form the gastric lumen. The first differentiated cell types to appear are mucous neck cells that act as progenitors for other cell types. Gastric glands grow by progressively branching, a process that continues until birth.

Food reaching the stomach is mixed with acid and digestive enzymes to accelerate the digestive process. The muscles of the stomach wall help mix and churn the gastric contents increasing the contact between food and the digestive mixture and grinding the food into smaller particles.

The stomach has three functionally and anatomically distinct parts: the cardia, the corpus and the pyloric antrum (**Figure 2**) (Fenoglio-Preiser *et al.*, 1999). These three areas are covered by a mucosa that contains a lining epithelium composed of gastric pits and gastric glands. The histological features of the lining epithelium of the gastric pits (**Figure 3**) is similar in all three areas, although the lengths of the pits differ in different gastric regions. In contrast, gastric glands differ in their histological features in the three anatomical regions. The cardia, a 1-cm segment of the stomach lying at its junction with the oesophagus, consists of mucus-producing glands resembling oesophageal mucus glands. The corpus normally constitutes approximately two-thirds of the gastric surface area. Its glands are lined by two major cell types. Chief cells produce pepsin, a digestive enzyme, from its precursor, pepsinogen. Parietal cells have several functions. They act as proton pumps that secrete hydrochloric acid into the stomach lumen, lowering the pH of gastric juice to 1.5. They also serve as the source of intrinsic factor which modulates small intestinal absorption of vitamin B_{12}. Autodigestion of the stomach by pepsin and hydrochloric acid is prevented by a thick mucus coat elaborated by mucosal surface cells. The pyloric antrum constitutes the distal third of the stomach. As in the corpus, its surface cells produce a thick layer of protective mucus. Antral glands contain chief cells that, like those in the corpus, secrete pepsin. Pepsinogen occurs in two forms. Pepsinogen group I (PGI) occurs only in the corpus, while pepsinogen group II (PGII) is made in all three parts of the stomach. The gastric epithelium in all three regions of the stomach maintains a dynamic equilibrium between cell production and cell loss. Cell proliferation increases any time that there is an excessive loss of gastric cells due to any mucosal injury (**Figure 4**).

The gastric mucosa also contains a diverse endocrine cell population. The predominant antral endocrine cell is the gastrin-producing G cell. The G cell is part of a physiological negative feedback loop. When acid secretion is low, gastrin stimulates acid secretion by parietal calls, followed by inhibition of gastrin release. Acid secretion also follows neural stimulation. Prolonged acid suppression induced by drugs, vagus nerve denervation or atrophy of the corpus mucosa leads to an increase in the G cell population and hypergastrinaemia (**Figure 5**). Hypergastrinaemia stimulates the growth of corpus ECL cells which may develop into small localized growths. Gastrin also functions as a general mucosal growth factor. Other gastric endocrine cells produce substances such as serotonin and somatostatin. Hormones produced by all gastric endocrine cells integrate physiological communications between the

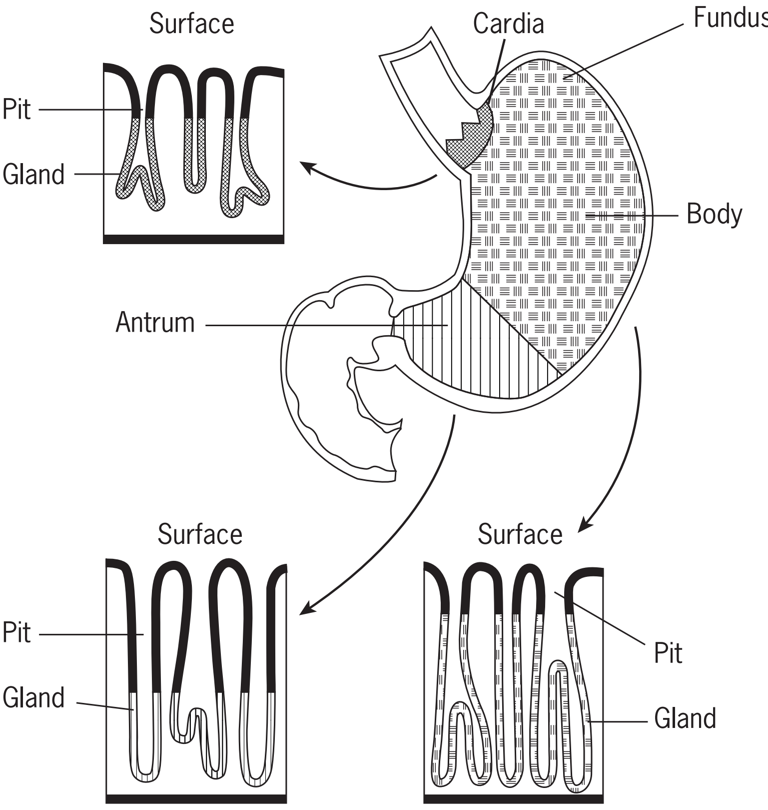


Figure 2 Diagrammatic representation of the stomach. The stomach is divided into four areas: the cardia, fundus, body and antrum. In each of these areas, the stomach is lined by surface epithelium which extends down into the gastric pits. Underlying the gastric pits are the gastric glands. The gastric glands of the fundus and body resemble one another and are sometimes referred to as oxyntic glands. These glands differ from those seen in the cardia and the antrum.

central and peripheral nervous system and the cells in the gastric mucosa and the muscle layers.

EPIDEMIOLOGY

Oesophageal Cancer

There are two major types of oesophageal cancer, squamous cell carcinoma (SCC) and adenocarcinoma, each with different risk factors and epidemiologies. SCC arises from the squamous cells lining the oesophagus. Adenocarcinomas arise in areas of metaplasia known as Barrett oesophagus.

Squamous Cell Carcinoma

The geographical distribution of oesophageal cancer shows wide variations within and between countries, in both mortality and incidence (Munoz and Day, 1996). As much as a 500-fold incidence difference can exist between the most and least affected areas of the same country and between different countries. SCC is virtually nonexistent in western and northern Africa. It also remains relatively rare in most of North America and Western Europe. In contrast, a high-risk zone extends from eastern Turkey through the southern former Soviet Union, Iraq, and Iran to northern China. High-risk areas also include Chile, the Transkei region of South Africa, Japan and regions of France and Brazil. Even in China, where 60% of oesophageal cancers develop, widespread differences in incidence and mortality exist. These high-risk foci have been explained on the basis of local food preservation practices that favour the generation of carcinogenic nitroso compounds from mould growing in pickled vegetables. Familial clusters of SCC of the oesophagus have been observed in Chinese high-risk areas. This can be attributed to shared environmental hazards, or to a common genetic influence.

Adenocarcinoma

Adenocarcinomas, cancers that form glands, constitute a relatively small proportion of oesophageal cancers worldwide, but among white males in the United States

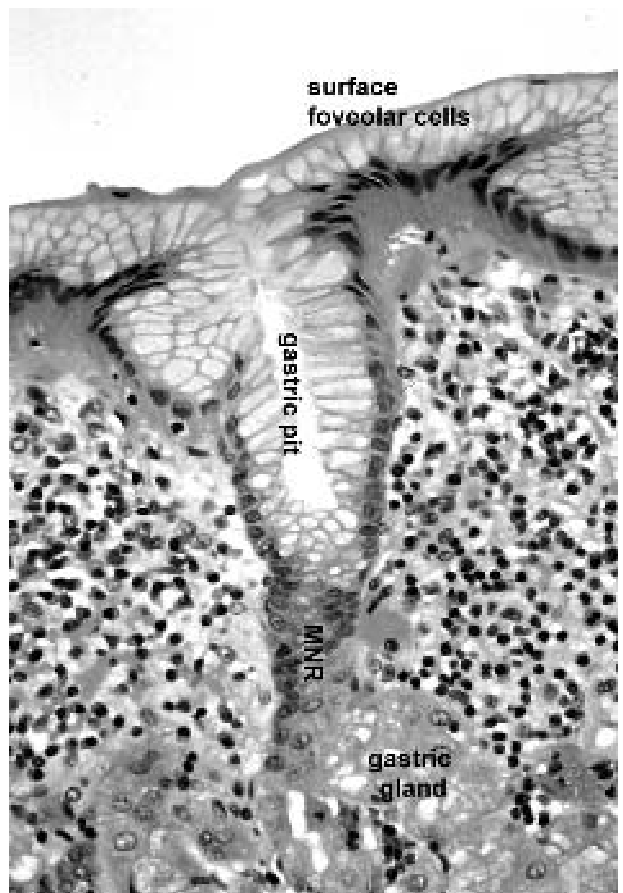


Figure 3 Photograph of the superficial portion of the gastric mucosa showing the surface foveolar epithelium extending into the gastric pit. Cell replication occurs in the mucous neck region (MNR), which is the area of junction betweem the pit and the underlying glands.

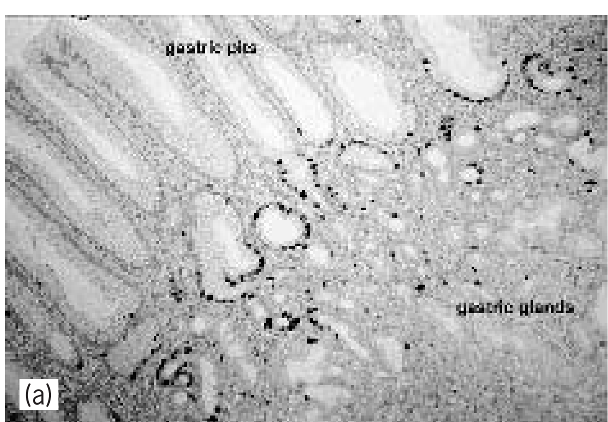
they are now more frequent than SCC (Zheng *et al*., 1993). Indeed, most affected patients are white males. Oesophageal carcinomas share many clinical features with cancers of the gastric cardia, and it may be difficult to determine whether a cancer involving the gastro-oesophageal junction originates from the oesophagus or the stomach. Thus, it is likely that epidemiological studies of oesophageal adenocarcinomas include some cancers of the cardia, and vice versa. This may explain why both cancers share similar time trends and a predisposition to affect white men. Both tumours have increased in incidence between 1975 and 1995, but the increased incidence of oesophageal adenocarcinoma in white males is especially strong, rising from 0.7 to 3.2 per 100 000, an increase of more than 350%. A diet characterized by a high intake of meat, fat and calories, but a low intake of fruits and vegetables, has also been identified as a risk factor for this tumour.

Stomach Cancer

As in the oesophagus, there is more than one form of gastric cancer and the epidemiology of the various forms differs.

Intestinal-type Gastric Cancer

The most common form of stomach cancer in high-risk populations is a gland-forming tumour that arises in the pyloric antrum and is preceded by the appearance of mucosal glands that resemble those in the intestine – a condition called intestinal metaplasia. The tumours that develop in the intestinalized mucosa are termed intestinal-type gastric cancers. Intestinal-type tumours account for almost two-thirds of the gastric cancers in the high-risk areas of northeast Asia, central America and eastern Europe. They arise after long-term infection with



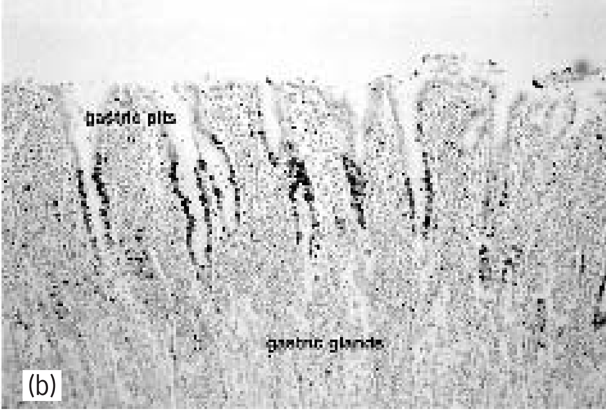


Figure 4 The area of the gastric pits which lies between the surface cells and the underlying gastric glands represents the proliferative zone of the stomach. (a) Under normal circumstances, isolated replicating cells are present, as indicated by the darkly stained cells. (b) When the mucosa becomes damaged, the number of brown staining cells increases and the gastric pits shorten. These changes indicate increased proliferation in response to mucosal injury.

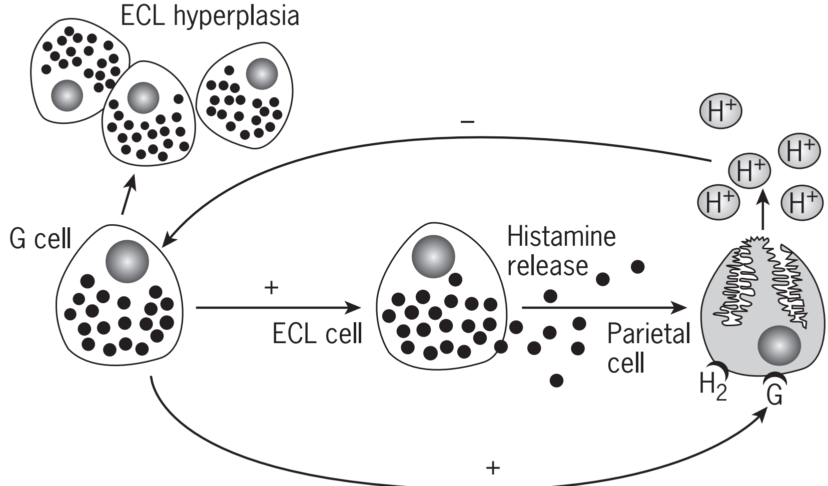


Figure 5 A feedback loop exists between parietal cells and G cells that secrete gastrin. When acid is secreted, the acid suppresses G cell function. In situations of hypochlorhydria, G cells increase their gastrin output. This serves directly to stimulate gastrin receptors on the surfaces of parietal cells. Gastrin also acts indirectly by stimulating ECL cells to release histamine. The histamine then binds to histamine receptors on the parietal cells. The binding of both gastrin and histamine to their respective receptors on parietal cells causes the acid secretion. In situations of prolonged G cell secretion, ECL hyperplasia develops.

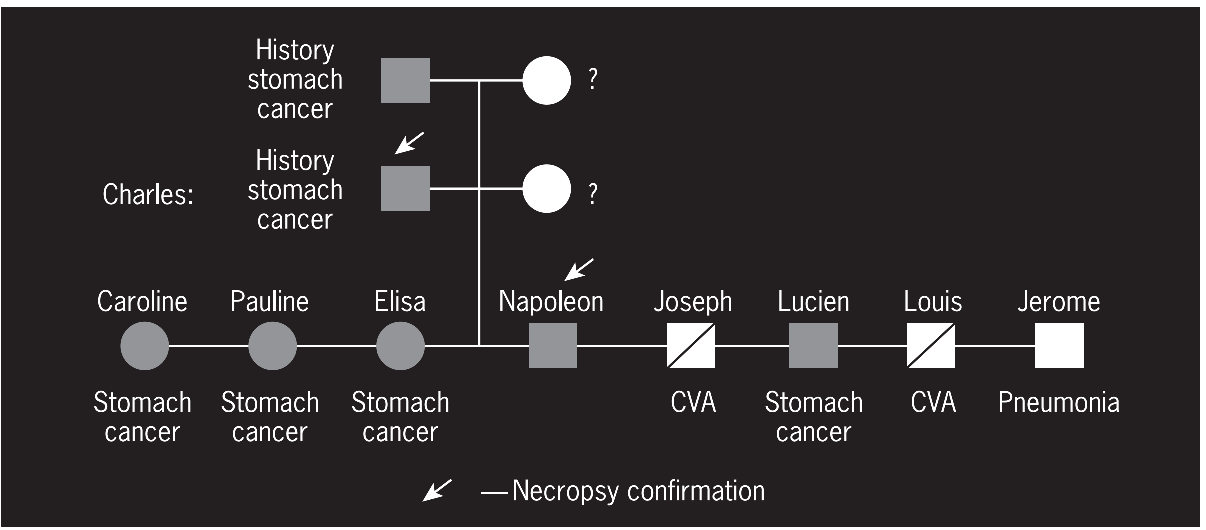


Figure 6 Diagram of the Bonaparte family tree indicating the presence of gastric cancer in several generations of the family.

Helicobacter pylori in persons whose diet is deficient in antioxidant vitamins (Correa, 1988; Nomura, 1996). Intestinal-type gastric cancer has shown a dramatic decrease in frequency in Western Europe and North America since 1940. This rapid decline in tumour incidence in westernized countries, and the persistence of high incidence rates in developing countries, suggest that gastric cancer risk is closely tied to socioeconomic status.

It is well recognized that the first-degree relatives of stomach cancer patients are at increased risk of developing gastric cancer, perhaps because the patients share common risk factors such as diet and *H. pylori* infections or because they share genetic factors that increase their risk of developing gastric cancer, or both. The Bonaparte family is a good example of a familial gastric cancer cluster (**Figure 6**) (Sokoloff, 1939). Napoleon and his father had autopsy-confirmed antral cancers. Four of his seven siblings were diagnosed with gastric cancer. Subsequently, there were 30 offspring from Napoleon's generation, and none developed gastric cancer. This and other similar familial kindreds suggest a pattern of weak genetic penetrance modulated by environmental factors. Napoleon's

brothers and sisters were raised in rural Corsica under less than ideal living conditions imposed by the hand to mouth existence of parents who, for many years, were active participants in a guerilla war against the French. The small size of the Bonaparte families after the move to the European mainland is typical of migrants who improved their economic status. The children were now housed in palaces and probably escaped *H. pylori* infection. They would have had an improved diet as well, that could have accounted for the absence of stomach cancer in subsequent generations.

Diffuse Gastric Cancer

Diffuse tumours preferentially arise in the corpus of patients <50 years of age who have severe superficial gastritis due to *H. pylori* infection (Nomura and Stemmerman, 1993). The appearance of these cancers in younger individuals without atrophic gastritis suggests the presence of a genetically driven increase in vulnerability to environmental carcinogens. Two observations support this concept. The first is the fact that patients with diffuse gastric cancers are more likely than the general population to have blood type A (Correa *et al*., 1973). Second, families with clusters of stomach cancer usually have diffuse-type carcinomas (Lehtola, 1978).

AETIOLOGY

Oesophagus

Squamous Cell Carcinoma

Oesophageal SCCs can be divided into two major risk groups: those tumours associated with tobacco and alcohol consumption and those occurring in populations consuming a poor, generally monotonous diet that lacks green, leafy vegetables, citrus fruits, micronutrients such as zinc, riboflavin and vitamin A and other unknown factors. These circumstances render the oesophageal mucosa more susceptible to injury by various carcinogens, including mycotoxins in the Transkei, substituted hydroxyphenanthrenes (a strongly mutagenic form of opium) in Iran, and *N*-nitroso compounds in China.

Heavy alcohol and cigarette consumption are recognized risk factors for oesophageal SCC, especially in Western populations. The increased risk with increasing alcohol consumption is exponential while the increase in risk from tobacco is linear (Tyuns *et al*., 1977). The strong association of oesophageal cancer with alcohol helps explain the especially high rates of oesophageal cancer in calvados-producing regions of northern France. Alcohol may contribute to an increased oesophageal cancer risk by reducing nutrient intake and by displacing protective dietary micronutrients (Ziegler, 1986). In addition, alcoholic beverages may contain carcinogens, or facilitate the transport of tobacco-associated carcinogens across the oesophageal lining. It may also impair the ability of the liver to detoxify carcinogens. In contrast, there appears to be a strong protective effect between the consumption of antioxidant vitamins and fresh fruits in areas at high risk for SCC.

Human papillomavirus (HPV) associates with the development of squamous cell carcinomas at many sites, including the oesophagus. There are approximately 70 distinct HPV strains, two of which, types 16 and 18, have been identified in oesophageal SCC. The frequency of HPV in oesophageal cancer varies from country to country, and appears to be most common in South Africa (Lewin and Appleman, 1996). HPV is probably not involved in the evolution in all oesophageal SCCs since it is absent from the majority of oesophageal cancers in Asia and Europe. When present, viral genetic sequences are present in nonneoplastic, dysplastic, invasive and metastatic lesions.

Motility disorders that delay oesophageal emptying increase the exposure of oesophageal squamous cells to ingested carcinogens. One such disorder, achalasia, results from impaired relaxation of the lower oesophageal sphincter and absent motility in the oesophageal wall. It is usually diagnosed after age 60 years. A population-based study showed that during the 24 years of follow-up, the risk increased more than 16-fold (Sandler *et al*., 1995). Fortunately, achalasia is uncommon.

Adenocarcinoma

The lining of the lower oesophagus can be injured if it is exposed to prolonged reflux of gastric acid and digestive enzymes. As a result, the normal squamous epithelium may change into a glandular lining resembling that of the stomach or intestines, producing the lesion known as Barrett metaplasia or Barrett oesophagus (Jankowski *et al*., 2000). Barrett oesophagus is a precursor to the development of oesophageal adenocarcinoma. Barrett oesophagus associates with decreased oesophageal sphincter resistance and increased gastric acid production, both of which lead to severe reflux oesophagitis. Obesity is also a risk factor, since it too predisposes to reflux disease.

Stomach Cancer

Intestinal-type Cancers

Intestinal metaplasia first appears in foci of atrophic mucosa at the antral–corpus junction (**Figure 7**). These foci ultimately fuse as the process progresses along the lesser curvature (Stemmerman, 1994). Further progression may result in the replacement of the entire antral mucosa by intestinal type glands and by proximal extension into the corpus. In late stages of the process all but a small portion of the corpus mucosa may be replaced by intestinal-type glands. In high-risk populations intestinalization of the stomach begins in adolescence, initiating a

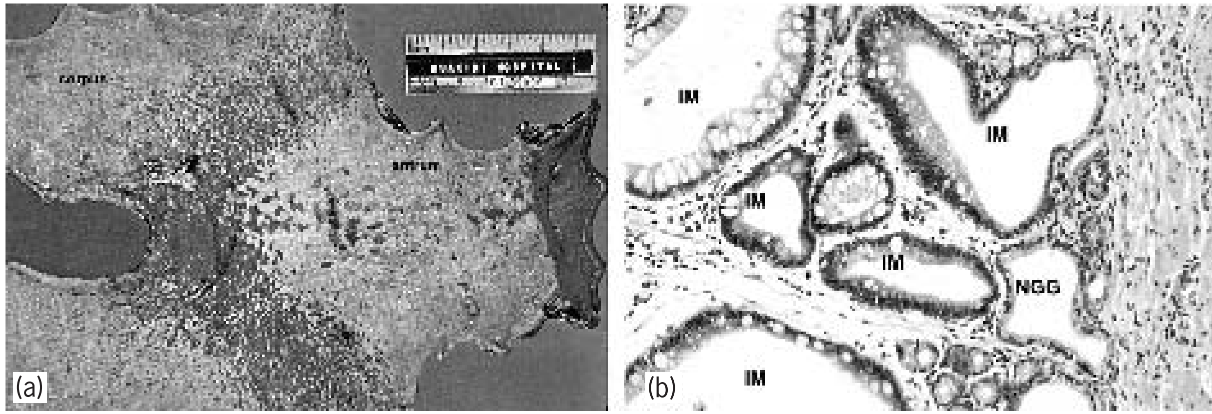


Figure 7 Intestinal metaplasia in the stomach. (a) This gross photograph of the stomach shows areas of dark staining. The stomach was immersed in a solution of alkaline phosphatase. This preferentially stains areas of intestinal metaplasia, resulting in the dark, inverted V-shaped pattern at the junction of the corpus and antrum. The duodenum located at the right hand edge of the photograph also intensely stains with the enzyme. (b) Histological features of intestinal metaplasia (IM) showing replacement of the normal gastric glands (NGG) with intestinalized glands.

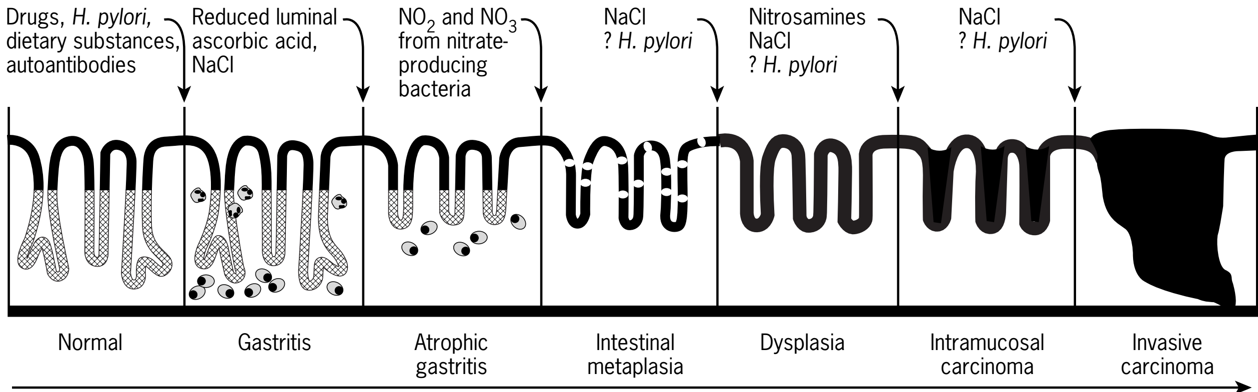


Figure 8 Diagram of the sequential abnormalities occurring in the development of intestinal-type cancers in the stomach. The individual steps are listed along the arrow at the bottom of the photograph. Substances thought to play a role in the progression of these steps are indicated above the diagram. This process is multifactorial and involves the presence of *H. pylori*, dietary substances and autoantibodies which produce gastritis and atrophy of the gastric glands that then lead to intestinal metaplasia. As a result of the atrophic gastritis and the replacement of gastric mucosa by mucosa resembling the intestine, the parietal cells decrease in number and as a result hydrochloric acid production decreases or ceases, allowing bacterial growth. These bacteria then serve to metabolize gastric contents into additional carcinogens. Additionally, the bacteria elicit an inflammatory response that generates genotoxic damage, further contributing to the molecular alterations occurring within the stomach and the genesis of a progressively abnormal gastric mucosa.

progressive expansion of the metaplasia. By the seventh decade, the metaplastic process may reduce the parietal cell volume to a level insufficient for maintaining gastric acid production. Loss of corpus chief cells results in low serum levels of PG I. As a result, tests that measure gastric acid production and serum PG I levels can identify persons who are at an increased risk for developing stomach cancer. Unfortunately, these tests do not identify individuals destined to develop stomach cancer before the intestinal metaplasia is sufficiently advanced to affect gastric function.

The sequential, multifactorial steps in the development of intestinal-type cancers seen in high-risk populations are shown in **Figure 8**. The gastritis results from the combined effects of *H. pylori* infection (**Figure 9**), high salt and nitrate intake, smoking and a diet deficient in fresh vegetables and protective antioxidant vitamins. Prospective studies have shown that 95% of persons who

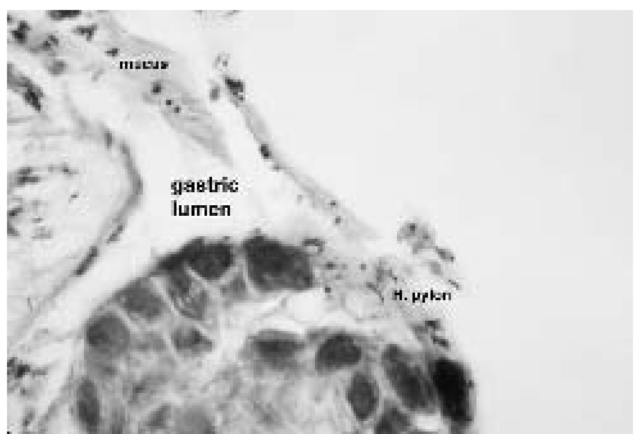


Figure 9 *H. pylori* infection of the stomach showing the presence of numerous *Helicobacter* attached to the surface epithelium. Alcian yellow stain of *H. pylori* infection showing the presence of numerous corkscrew-shaped organisms in the mucus overlying the gastric epithelial cells.

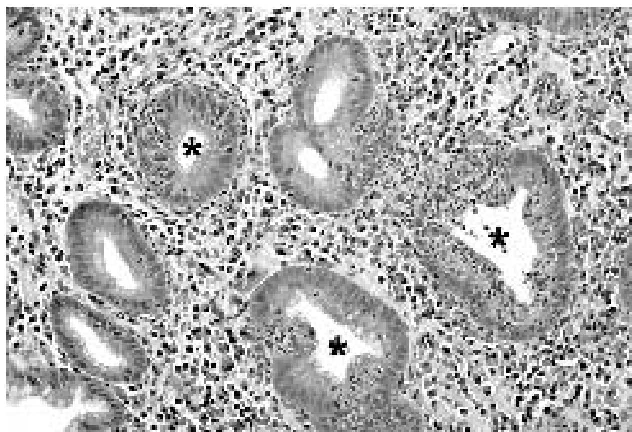


Figure 10 (a) Active chronic gastritis due to *H. pylori* infection. (b) Note the presence of large numbers of acute inflammatory cells in the starred glands.

develop cancer were infected with *H. pylori.* However, although at least 75% of persons in high-risk areas are infected with *H. pylori*, only 5% actually develop stomach cancer (Nomura and Stemmerman, 1993). Whether or not an infected person develops cancer probably depends upon their genetic predisposition to develop cancer, their level of exposure to other risk factors and possibly to the strain of *H. pylori* infecting them.

H. pylori Infection

H. pylori infection is most frequent in large families occupying crowded living quarters and lacking hot water. The infection is usually acquired during childhood, is related to birth order; the highest incidence affects the youngest children of large families (Goodman and Correa, 1995). *H. pylori* infection is most common among those with the shortest duration of schooling. The smaller families and improved housing that typify a prosperous economy in developed countries may explain, at least in part, the decrease in their gastric cancer rates.

H. pylori infection increases gastric cancer risk via several pathways. The infection causes epithelial cell damage and the rate of cell proliferation increases in response to the cell loss. As the cells proliferate, the DNA unwinds, increasing the number of gastric cells vulnerable to genotoxic damage and subsequent genetic instability. Additionally, *H. pylori* infections generate a brisk inflammatory response with invasion of the replicating epithelium by white blood cells (**Figure 10**). Two types of inflammatory cells, neutrophils and monocytes, generate the toxic molecule, nitric oxide (NO) (**Figure 11**). In addition, other toxic molecules can be generated from NO including nitrosamines. Nitrosamines induce mutations in bacteria and are gastric carcinogens in many animal models. The local production of these compounds near the replicating gastric epithelium exposes them to an increased risk of DNA damage. The inflammatory cells also induce free radical damage due to the formation of NO and its oxidative byproducts. Vitamin C (ascorbic acid), an antioxidant, protects the gastric mucosa against oxidative stress. However, *H. pylori* infections decrease ascorbic acid concentrations in the gastric juice. Not all *H. pylori* are equally efficient in inducing the changes noted above. The most virulent of the bacteria contain a gene known as *cag* (Blaser *et al*., 1995).

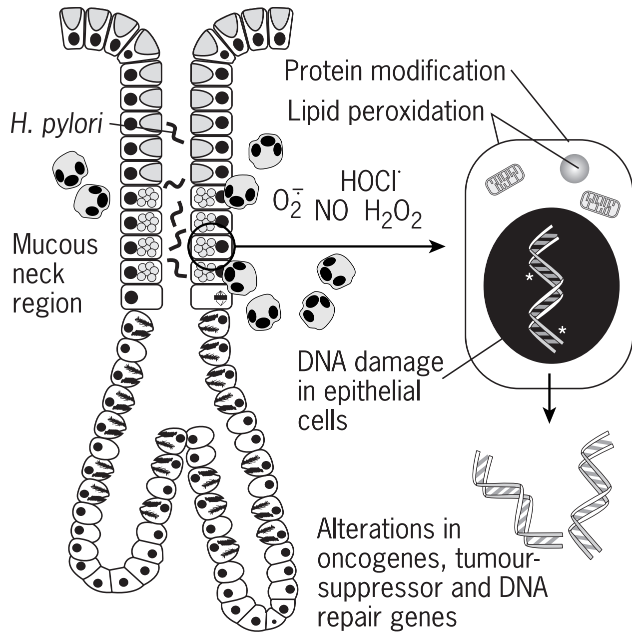


Figure 11 Diagram of the injury induced by neutrophils in the mucous neck region.

Diet and Stomach Cancer

Dietary patterns were once thought to be *the* most important basis of gastric cancer induction, and although diet has recently conceded the pride of place to *H. pylori* infection, diet still retains a major role in the cause and prevention of this tumour. Strong experimental and epidemiological evidence links salt intake to gastritis, intestinal metaplasia and gastric cancer. Evidence that dietary nitrite and nitrate play a direct role in gastric carcinogenesis is inconclusive, but these substances do induce intestinal metaplasia. Foods with high salt and nitrate concentrations, such as dried and salted fish or pickled vegetables, predispose to stomach cancer in many studies. Such foods constituted a major portion of the diet prior to the almost universal ownership of household refrigeration in Western countries and the availability of fresh or frozen fruits, vegetables, fish and meat. The decrease in gastric cancer rates in Western countries since 1945 is attributable to altered dietary habits with decreased consumption of the potential carcinogens in smoked, salted and pickled foods.

Autoimmune Gastritis

Autoimmune gastritis is a precursor to both benign and malignant stomach tumours. It is less common than multifocal gastritis and it develops when patients generate antibodies against their own parietal and chief cells. Cell proliferation in the corpus increases in response to the antibody-mediated cell loss. However, because cell destruction proceeds faster than cellular replacement, the mucosa of the corpus becomes atrophic. Loss of parietal cells results in a reduction in gastric acid production and a compensatory increase in antral gastrin secretion. The increased gastric cancer risk in this condition results from several synchronous events: (1) a bacterial flora emerges in the stomach which can generate potential carcinogens from dietary amines; (2) endogenous production of mutagens may follow NO production by reacting white cells in the stomach lining; (3) the increased cell turnover puts more cells at risk of accumulating genetic damage; and (4) gastrin functions as a growth promoter of gastric lining cells.

The diagnosis of autoimmune gastritis is confirmed by demonstrating low or no stomach acid and very low serum levels of pepsinogen group I (Samloff *et al*., 1975). As in multifocal gastritis, the gastric mucosa of patients with autoimmune gastritis may develop intestinal metaplasia. The increased gastrin production stimulates gastric mucosal growth, sometimes contributing to the development of gastric polyps and gastric carcinoma (**Figure 12**). The risk of developing carcinoma with autoimmune gastritis varies from 2.1 to 5.6 times that of persons without the condition. The cancers arise in the antrum as well as in the corpus.

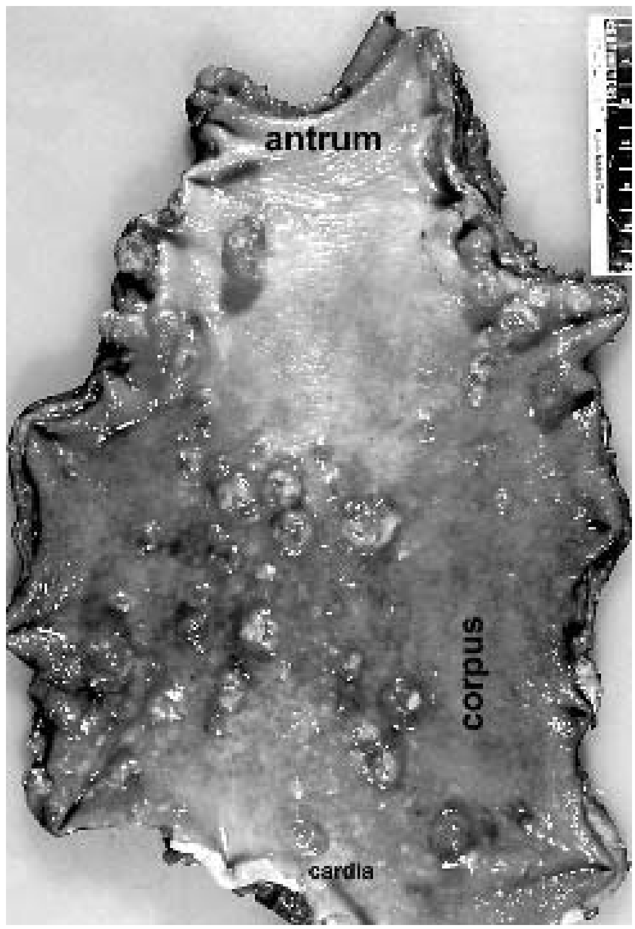


Figure 12 Gross photograph of a stomach from a patient with autoimmune gastritis showing the presence of numerous gastric polyps, as well as a gastric carcinoma. They mainly arise in the corpus. The surrounding mucosa is atrophic.

Previous Surgery

The risk of developing stomach cancer complicates previous ulcer surgery (Fisher *et al*., 1993). Bleeding from antral or duodenal peptic ulcers historically required the removal of these segments of the gastrointestinal tract, and also the nerves that stimulate acid secretion. The gastric remnant was often joined to a small intestinal loop, resulting in reflux of intestinal and pancreatic juices into the gastric remnant. This reflux of digestive enzymes into the denervated gastric remnant causes a gastritis resembling multifocal or autoimmune gastritis, increasing the gastric cancer risk. The elevation of risk becomes apparent 20 years after the resection and increases thereafter. The younger the patient at the time of surgery, the greater the risk of ultimately acquiring a carcinoma.

Radiation

The stomach is sometimes in the radiation field in patients with abdominal tumours, especially certain forms of lymphoma occurring in young people. A small number of these patients have an increased risk of developing stomach cancer. The tumours derive from X-ray-induced mutations in the gastric mucosa. Japanese victims of the

Hiroshima/Nagasaki atomic bomb explosions also had an increased incidence of gastric cancer.

Diffuse Cancer

The aetiology of diffuse gastric cancer is poorly understood, although *H. pylori* infections appear to play a role in their genesis.

SCREENING AND PREVENTION

Oesophagus

Squamous Cell Carcinoma

Patients with SCC present with dysphagia (difficulty swallowing) or with odynophagia (pain). Initially, symptoms only occur with the consumption of solid foods, but as the tumour progresses even liquids are swallowed with difficulty. Severe weight loss results from decreased food intake. Many cancers ulcerate, resulting in upper gastrointestinal bleeding. The bleeding is usually inconspicuous and is not noticed by the patient. Fatal internal bleeding may occur if a deeply invasive cancer erodes a major blood vessel. The only way to avoid these dire consequences is to screen populations at risk for the disease. The intent of screening programmes is to detect tumours while they are still curable. This is possible because cancers develop over a period of time, usually evolving over a decade or more.

The study of oesophageal lining cells that are dislodged by the passage of a brush or a balloon (cytological examination) allows the discovery of cancerous cells in early phases of tumour progression (**Figure 13**). Cytology screening programmes form the basis of community screening programmes in the high-risk areas of central China where up to 50% of the resected oesophageal cancers are noninvasive or early invasive tumours (Qui and Yang, 1988). In contrast, SCC is too uncommon in Western populations to justify community screening programmes. Rather, diagnostic procedures in most Western, low-risk countries are designed to explain the symptoms of late-stage disease and to estimate the extent of the cancer that caused them. The diagnosis is usually made on tissue obtained from an endoscopic biopsy.

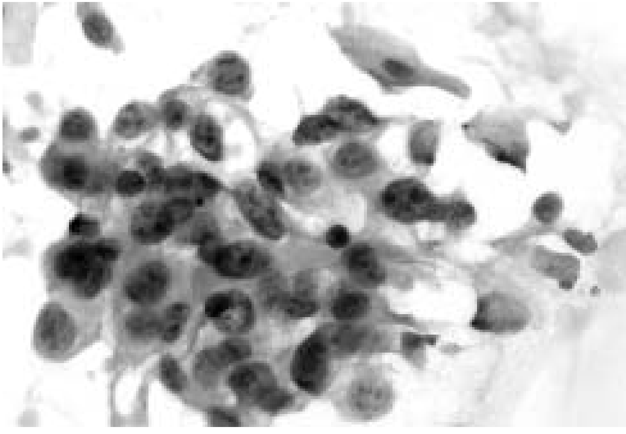


Figure 13 Cytological preparation of a carcinoma of the oesophagus. These cells were brushed off the surface of the oesophagus. They preferentially dislodge because of their malignant qualities. They have enlarged nuclei and small amounts of cytoplasm. There is cytologic atypia.

Suggested guidelines for endoscopic surveillance of the oesophageal mucosa in patient populations with a high risk of developing squamous cell cancer areas are as follows: biopsy specimens should be obtained from all plaque-like, nodular, erosive and friable areas; normal areas, including mildly wrinkled mucosa and isolated small white patches, need not be sampled; and biopsy specimens should be obtained from diffusely irregular (prominently wrinkled) mucosa and distinct areas of focal reddening at the endoscopist's discretion.

It remains unclear as to whether these recommendations should be applied to patients in non-high-risk areas, although they may be a desirable addition to the examination of patients entering an alcohol abuse treatment programme.

Adenocarcinoma

The symptoms associated with progression of oesophageal adenocarcinomas are essentially similar to those associated with SCC – difficulty swallowing, obstruction of the passage of solid foods, weight loss and dissemination to other sites. As is the case of SCC, long-term survival requires diagnosis at a preinvasive or early invasive stage. This is most likely to occur if patients known to have Barrett change are screened on a regular basis in order to detect progression of the process to the level of high-grade dysplasia.

The following management approach is that recommended by the 1990 Barrett's Esophagus Working Party of the World Congress of Gastroenterology (Dent, 1989):

- A programme of regular endoscopic surveillance for dysplasia and early carcinoma is recommended for patients with Barrett oesophagus unless contraindicated by comorbidity. For patients who have no dysplasia or cancer, endoscopy (with procurement of biopsy and brush cytology specimen(s) from the Barrett's epithelium) is performed every other year.
- If dysplasia is detected, the finding should be confirmed by at least one other expert pathologist. If any doubt remains, the endoscopic examination is repeated immediately to obtain more biopsy and cytology specimens for analysis.
- For patients confirmed to have multiple foci of high-grade dysplasia, surgery is advised to resect all of the oesophagus lined by columnar epithelium.

- For patients confirmed to have low-grade dysplasia, intensive medical antireflux therapy (including omeprazole) should be given for 8–12 weeks, at which time endoscopic examination is repeated to obtain multiple oesophageal biopsy and cytology specimens.
- For patients whose specimens show histological improvement, intensive surveillance (e.g. endoscopic examination every 6 months) is recommended until at least two consecutive examinations reveal no dysplastic epithelium.
- For patients with persistent low-grade dysplasia, continued intensive treatment and surveillance are recommended.

Stomach

The nature of the symptoms depends upon the site and growth pattern of the tumour. Bulky tumours may erode and cause bleeding, while diffuse tumours may restrict the expansion of the stomach, causing early satiety and severe weight loss due to decreased food intake. It is not unusual for the first symptoms to be caused by metastases in sites remote from the stomach. In young women, the spread of gastric cancer to the ovaries produces so-called Krukenberg tumours that may reach the size of a grapefruit. Other symptoms caused by metastases include spread of the cancer to the lymph nodes of the lower neck and to the accumulation of excess fluid in the abdominal cavity. The configuration and mode of progression of stomach cancer depend upon the age and sex of the patient, the subsite of its origin within the stomach and its histological type.

The Japanese, in response to the high incidence of gastric cancer in their country, have introduced community screening to detect cancers early in their development. As a result, 50–60% of stomach cancers are discovered while still limited to the stomach lining or to the tissues immediately beneath it. Recent Japanese studies have also assessed the effectiveness of radioimmunoassays to determine the serum pepsinogen group I level (PG I), and the ratio between PG I and PG II in predicting the presence of stomach cancer in participants in an endoscopic screening programme (Kitahara *et al*., 1999). However, the sensitivity of this approach is not sufficient to detect all tumour cases. Additional variables must be identified before a focused screening programme can be used to identify high-risk subjects in low-risk Western settings. These might include two or more of the following: (1) first-generation migrants from countries with high rates of gastric cancer; (2) a high *H. pylori* antibody level; (3) persons with a strong history of stomach cancer among close family relatives; (4) smokers with a history of stomach ulcer; (5) persons who have had a partial resection of the stomach 17 or more years previously; and (6) persons who have received radiation treatment to the upper abdomen prior to age 30. Japanese studies of patients with untreated early cancer suggest that re-examination need not be performed more frequently than every 5 years.

PATHOLOGY

Oesophageal Carcinoma

Squamous Cell Carcinoma

As the oesophageal lining passes through a series of sequential steps that eventually result in the development of invasive SCC, it becomes progressively abnormal, a change that can be seen both by gross examination and by examination under the microscope. These steps include the presence of varying degrees of intraepithelial (non-invasive) neoplasia, termed dysplasia, finally ending in an invasive carcinoma that has the ability to spread and kill the patient (**Figure 14**).

Gross Appearance

The gross appearance of oesophageal neoplasia varies with the stage of neoplasia present. The earliest lesions may

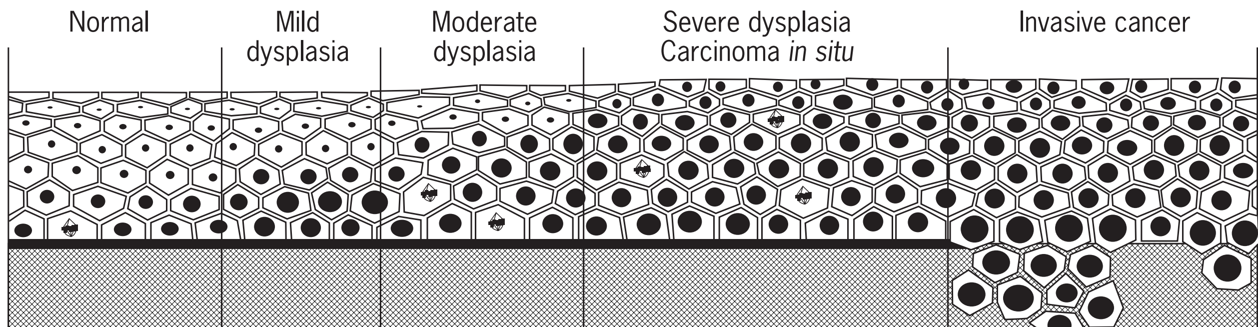


Figure 14 Diagrammatic summary of the progression of squamous cell neoplasia. As the disease progresses, more and more of the thickness of the oesophageal mucosa becomes replaced by neoplastic cells. These eventually invade through the area of the basement membrane into the underlying lamina propria and then into the submucosa. From here they can metastasize to sites distant from the oesophagus.

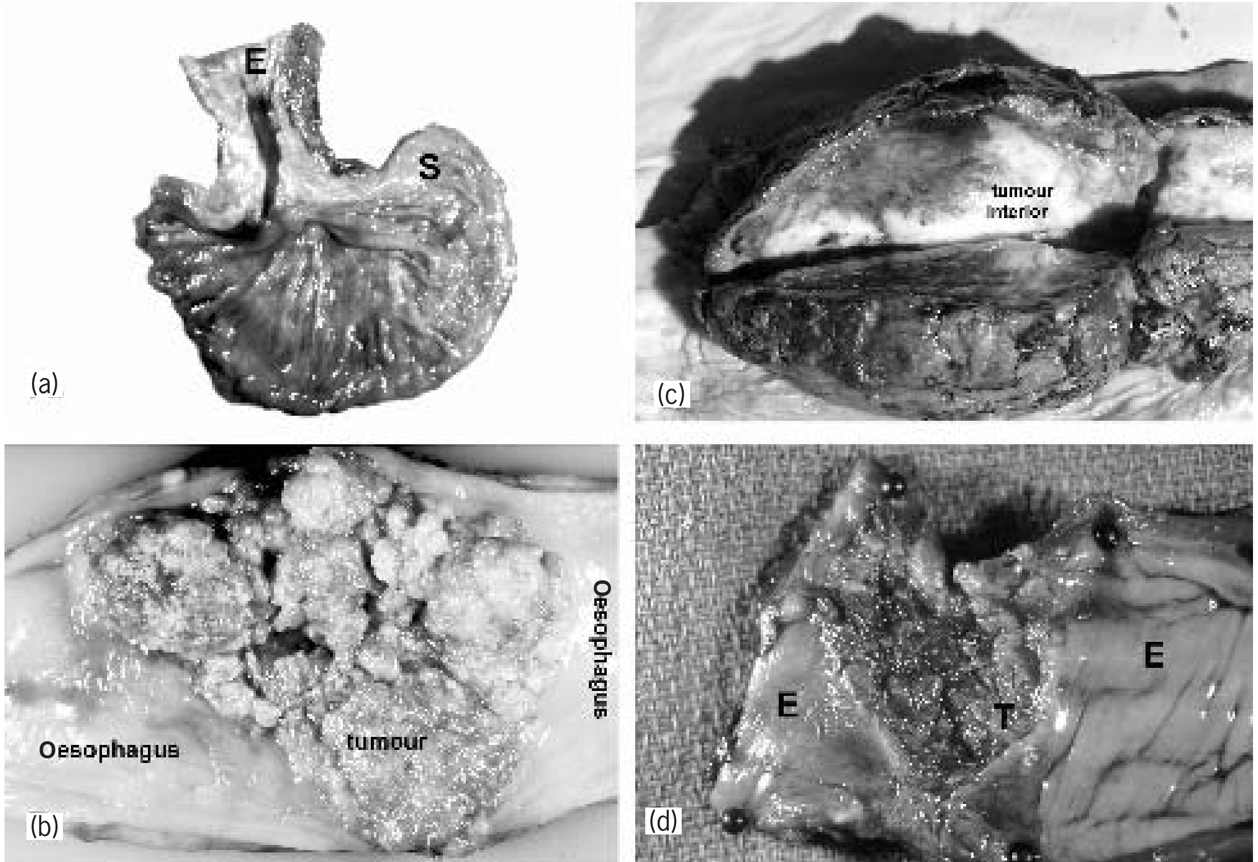


Figure 15 Gross appearances of oesophageal cancers. Oesophageal squamous cell cancers may assume a variety of gross appearances. (a) Diffuse infiltration of the oesophageal wall causes thickening of the distal end of the oesophagus (E). The line shows the area of the tumour. The stomach (S) is also present. (b) An exophytic papillary tumour rises above the surrounding oesophageal mucosa. (c) A large, bulky, polypoid growth. The tumour has been cut in half. (d) An ulcerating oesophageal cancer (T) is surrounded by normal oesophagus (E).

appear completely normal to the naked eye or to the endoscopist. (The endoscopist is the clinician who examines the oesophageal mucosa with a magnifying camera inserted into the oesophageal lumen). Alternatively, the oesophageal mucosa may appear reddened owing to inflammation and vascular congestion induced by the presence of neoplastic cells. Early lesions (intraepithelial or minimally invasive lesions) may also appear as areas of superficial erosion and/or as whitish plaques. Such lesions may be single or multiple. Invasive carcinomas usually arise from these erosive or plaque-like lesions. Invasive SCC usually arise in the distal half of the oesophagus, although they can occur anywhere. Grossly, invasive oesophageal squamous cell cancers assume various growth patterns (**Figure 15**). They may be large bulky polypoid lesions growing into and obstructing the oesophageal lumen. Other tumours grow as flat lesions or they grow into the oesophageal wall, creating oesophageal ulcers. Still other tumours diffusely infiltrate the oesophageal wall, causing oesophageal strictures. These growth patterns can occur alone, or they can associate with one another. The extent and configuration of an invasive tumour partly determines how a patient presents. Patients with bulky polypoid lesions, or with areas of stenosis, are more likely to experience difficulty in swallowing than those with ulcerating or flat tumours.

Microscopic Appearance

As indicated in an earlier section, the oesophageal mucosal lining consists of an orderly arrangement of squamous epithelial cells that regularly renew themselves. Progenitor cells lying in the basal zone give rise to new cells that become increasingly mature as they pass upward in the mucosa. As part of their maturation, the cells progressively enlarge acquiring a smaller nuclear–cytoplasmic ratio as they do so. When the cells become neoplastic, they lose this orderly pattern of growth and maturation and proliferation no longer remains restricted to the basal cell layer. The cells appear disorganized and jumbled, and depending on the degree of change that is present, mitotic activity appears at various levels in the epithelial layer. This earliest form of neoplasia is called dysplasia and

implies the presence of an unequivocally neoplastic epithelium.

Historically, dysplasia came in several grades, ranging from mild to moderate to severe (**Figure 16**). When the entire epithelial thickness is replaced by neoplastic cells, the process may be termed carcinoma *in situ* (CIS). Today there is a tendency to place intraepithelial neoplasias into two grades, high- and low-grade dysplasias. Using such a two-tier system, CIS falls into the category of severe dysplasia. Progression of squamous cell neoplasia beyond CIS is recognized by the presence of disruption of the basement membrane and penetration of the tumour into the underlying tissues. The term microinvasive carcinoma can be used if the invasive tumour tongues are minute, penetrating only a few millimetres into the underlying tissues. Once invasion has occurred, the tumour may spread horizontally beneath the intact squamous cell lining adjacent to the cancer, or it may variably penetrate through the oesophageal wall, eventually reaching its external surface and potentially entering contiguous structures. If that structure is the trachea, ingested food can be aspirated into the lungs via the fistula created by the tumour. If the adjacent structure is the aorta or pulmonary artery, massive and fatal internal haemorrhage will result. If the patient is spared either of these devastating events, the tumour may find its way into the lymph channels that drain the wall of the oesophagus. Tumour cells in these lymphatics may then spread to regional lymph nodes and then disseminate to more distant sites causing metastases in other organs.

Invasive squamous cell carcinomas exhibit varying degrees of cellular differentiation. If they closely resemble the squamous cells from which they arose, they are termed well-differentiated squamous cell carcinomas. Tumours composed of cells difficult to recognize as being squamous in nature are termed poorly differentiated. Tumour cells whose degree of differentiation is intermediate between these two extremes are said to be moderately differentiated.

Adenocarcinoma of the Oesophagus

As indicated above, oesophageal adenocarcinomas do not arise from the native squamous epithelium. Rather, they arise from areas of Barrett oesophagus which is a metaplastic epithelium consisting of glandular cells. As with squamous cell carcinomas, the epithelial lining of Barrett oesophagus becomes progressively abnormal as it passes through a series of sequential steps that eventually result in the development of invasive adenocarcinoma (**Figure 17**). These steps include the development of glandular dysplasia followed by the appearance of an invasive malignancy.

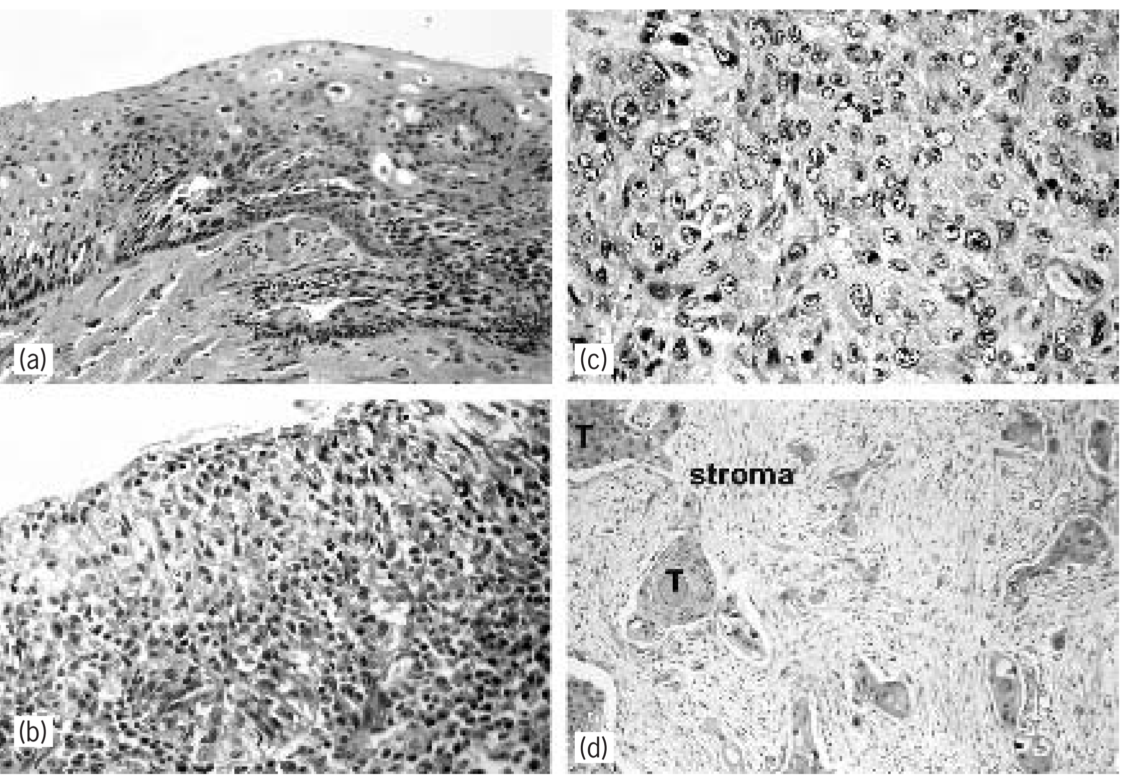


Figure 16 Histological features of the progression of oesophageal squamous cell neoplasia. (a) Low-grade dysplasia. (b) High-grade dysplasia. (c) Carcinoma *in situ*. (d) Invasive cancer. The tumour nests (T) are surrounded by a desmoplastic stroma.

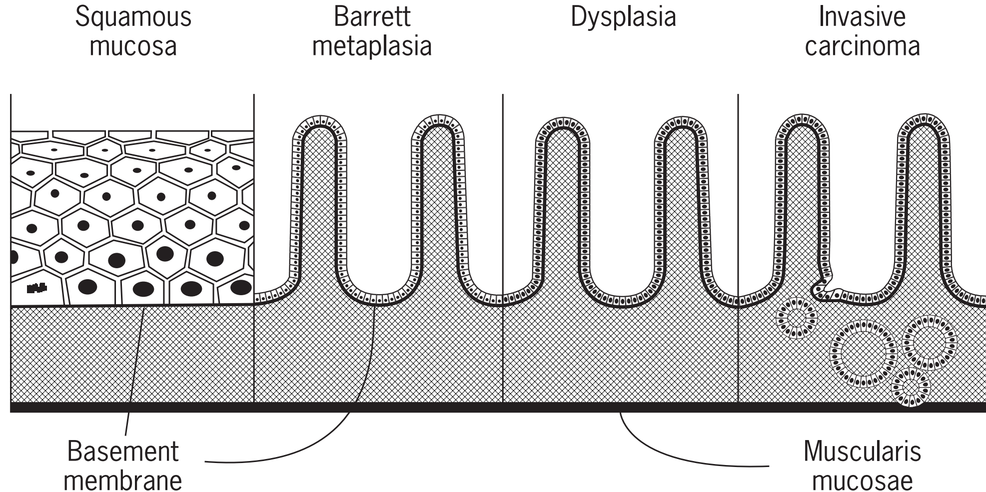


Figure 17 Diagrammatic representation of the progression of the normal oesophagus to an adenocarcinoma arising in the background of Barrett oesophagus.

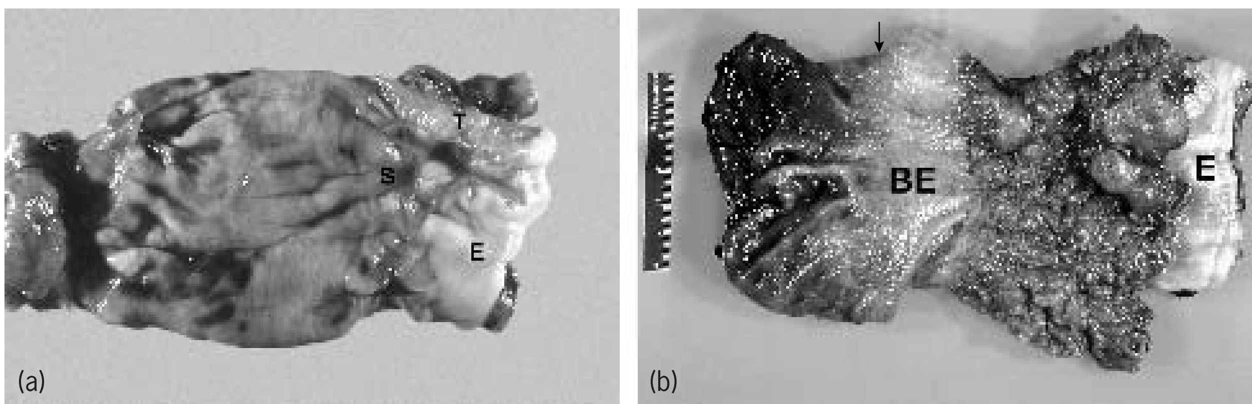


Figure 18 Gross appearance of adenocarcinoma arising in Barrett oesophagus. (a) Note the irregularity of the squamocolumnar junction. Normally, there should be a sharp demarcation with the distal stomach. The stomach (S) ends near the S. The smooth, lighter coloured mucosa represents the normal colour of the oesophagus (E). A small tumour is present beneath the T. (b) This photograph shows a large, fungating neoplasm arising in an area of Barrett oesophagus. The gastric folds terminate at the area of the arrow. The area above this is replaced by Barrett oesophagus (BE). Above that is the tumour, and the smooth, lighter coloured area of the mucosa represents normal oesophagus (E). The oesophagus has been extensively replaced by glandular epithelium and by the neoplasm.

Gross Appearance

The gross appearance of oesophageal glandular neoplasia varies, depending on the stage in the process at which it is examined (**Figure 18**). The lower end of the oesophagus loses its normal pinkish tan smooth appearance, typical of the native squamous mucosa, and instead it becomes reddened and less smooth. This is the earliest change and it is seen in patients with Barrett oesophagus. The mucosal alterations may make it difficult to determine where the original gastro-oesophageal junction lay before the Barrett oesophagus developed. The best way of determining the proximal extent of the stomach is to identify the termination of the gastric folds. The dysplasias that develop in the areas of Barrett oesophagus may be flat or polypoid. Those that are flat may be completely invisible to the naked eye and to the endoscopist. Some areas of dysplasia may appear slightly depressed or ulcerated. As invasive cancers develop, they tend to invade both laterally and into the oesophageal wall. Such lesions may be more visible when the cut surface of the oesophagus is examined or if the tissues are palpated for areas of firmness. It is very common for adenocarcinomas to extend underneath adjacent normal tissues. When the tumours become large, bulky polypoid or ulcerating lesions, it may be impossible

to tell whether they arose in the distal oesophagus or in the cardia of the stomach. These are sometimes referred to as tumours of the oesophageal–gastric junction The majority of these cancers are flat or ulcerated, although in a third of the cases they may form elevated masses – polypoid cancers.

Microscopic Appearance

Disordered cell growth, called dysplasia, constitutes the first step in the progression of Barrett change into cancer. Dysplasia is recognized by the appearance of an increased nuclear–cytoplasmic ratio and loss of nuclear polarity. Initially, the metaplastic glands are lined by cells that become pseudostratified, somewhat resembling colonic adenomas. As the degree of dysplasia progresses, the cells become increasingly disorganized and the nuclei lose their pseudostratified appearance. Invasive cancer becomes recognizable once the tumour cells extend through the muscularis mucosae into the underlying submucosa. The evolution from low-grade dysplasia through moderate- to high-grade dysplasia and invasive cancer is shown in **Figures 17 and 19**. Overall, many invasive tumours appear to be very well differentiated because they make clearly identified glands. Other tumours are poorly differentiated and it is difficult to identify glandular structures. Even well-differentiated tumours have the tendency to undermine adjacent non-neoplastic tissues and to invade the lymphatics of the oesophageal wall at an early stage. As a result, most cancers arising at the cardio-oesophageal junction are late-stage tumours, with dissemination to lymph nodes on both sides of the diaphragm at the time of diagnosis. Spread of these tumours to the liver is also fairly common. Oesophageal cancers that have progressed to the point of causing symptoms carry an extremely poor prognosis. The majority of symptomatic oesophageal cancers are unresectable, and the average life expectancy of untreated patients is less than 10 months.

Stomach Cancer

Gross Appearance

The gross appearance of early gastric cancers tumours is shown in **Figure 20**. Protruded, type 1, early cancers are most likely to occur in the antrum of older male patients with diffuse intestinal metaplasia of the stomach. Flat or ulcerated type 2 and 3 early tumours are more likely to occur in younger patients and may arise in the mucosa at the junction of the antrum and corpus of a stomach showing minimal or no intestinal metaplasia.

The appearance of advanced tumours is shown in **Figures 21 and 22**. Cancers that arise in the corpus are usually type IV tumours, as are most cancers that occur in women <50 years old. All other growth patterns preferentially affect antral tumours, and they are most frequent in older men from high-risk populations.

Microscopic Appearance

Gastric cancer presents a complex variety of growth patterns, and one cancer may assume several forms. Lauren, a Finnish pathologist, devised a simplified classification of these tumours (Lauren, 1965), which subsequently proved helpful to epidemiologists studying the origins of stomach cancer. This system is based on the observation that many stomach cancers form rudimentary glands that superficially resemble intestinal glands; these are termed intestinal in type (**Figure 23**). Others consist of discohesive cells that form no distinct structures and stimulate an overgrowth of the supporting connective tissues (**Figure 24**). These are termed diffuse tumours and their cells may be so widely separated from one another that that they are difficult to identify. The antral tumours in older men with extensive multifocal gastritis generally are intestinal in type, whereas the cancers that arise in the corpus of younger men and women are usually diffuse in type. In addition, most stomach cancers arising in high-risk areas are intestinal in type. It is not unusual for gastric cancers to show both patterns of growth and these can be called mixed tumours. A committee of World Health Organisation pathologists devised another classification scheme, that identifies the many variants of stomach cancer (**Table 1**). Classifying gastric cancer is difficult by either system, and interobserver differences lessen their value as prognostic markers for individual cases.

TREATMENT

Oesophagus

Squamous Cell Carcinoma

Surgical treatment of SCC is most likely to be successful when the tumour is confined to the lining of the oesophagus or when the invasive tumour is limited to the most superficial portion of the oesophageal wall. The extent of the tumour may be estimated through the use of endoscopic sonography, which uses echoes of ultrasonic pulses at the time of endoscopy to measure the depth of tumour penetration into the oesophageal wall and to identify metastases into the regional lymph nodes. This procedure allows the physician to assign a stage to the tumour and to plan patient treatment. The standard staging method employed is termed the TNM system, where T indicates the extent of tumour penetration, N indicates the presence or absence of lymph node metastases and M indicates the presence or absence of distant metastases. The TNM staging criteria for oesophageal cancer are shown in **Table 2**.

The treatment of the late-stage oesophageal cancer that typically confronts the physician in most Western countries is more likely to be palliative than curable. Surgery may resect or bypass the tumour, ionizing radiation may be used to shrink or destroy it and chemotherapy may be used

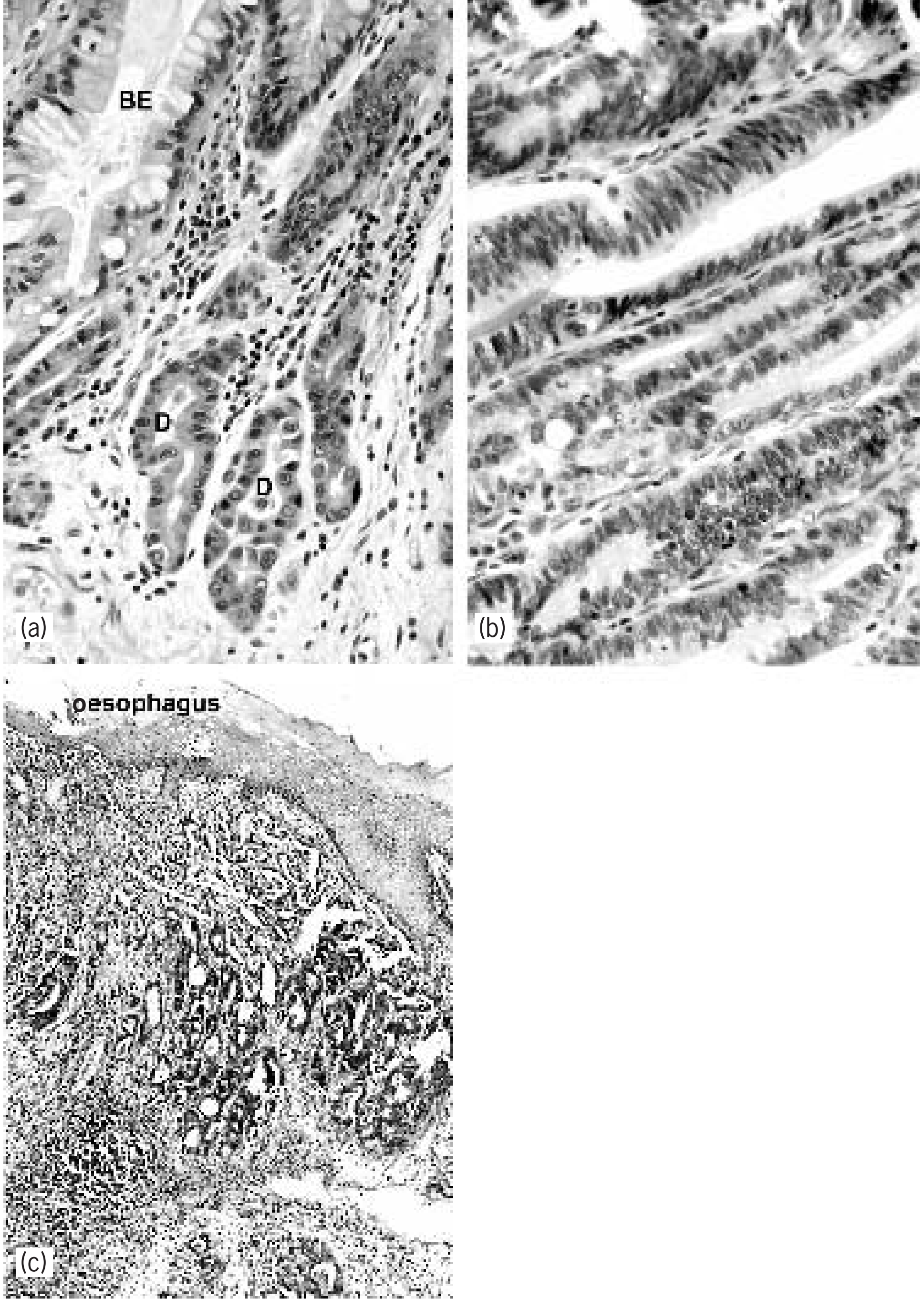


Figure 19 Evolution of the neoplastic changes in Barrett oesophagus. (a) The squamous epithelium has been replaced by glandular epithelium. One gland is a non-neoplastic gland representing Barrett oesophagus (BE). The majority of the glands that are present show evidence of low-grade dysplasia (D). (b) Further progression of the dysplasia with the presence of large numbers of cells showing nuclear palisading. (c) An invasive adenocarcinoma undermining the normal squamous mucosa of the oesophagus.

to control its growth. These approaches may be used alone or in combination with one another, depending upon the stage of the tumour and the physical condition of the patient. For example, the presence of cirrhosis of the liver, a not uncommon condition among heavy alcohol users with advanced oesophageal cancer, may preclude the use of many chemotherapeutic agents. In contrast, an early cancer in an otherwise healthy patient may be treated by surgical resection alone. Radiation or chemotherapy may be used prior to surgery (so-called neoadjuvant therapy) in

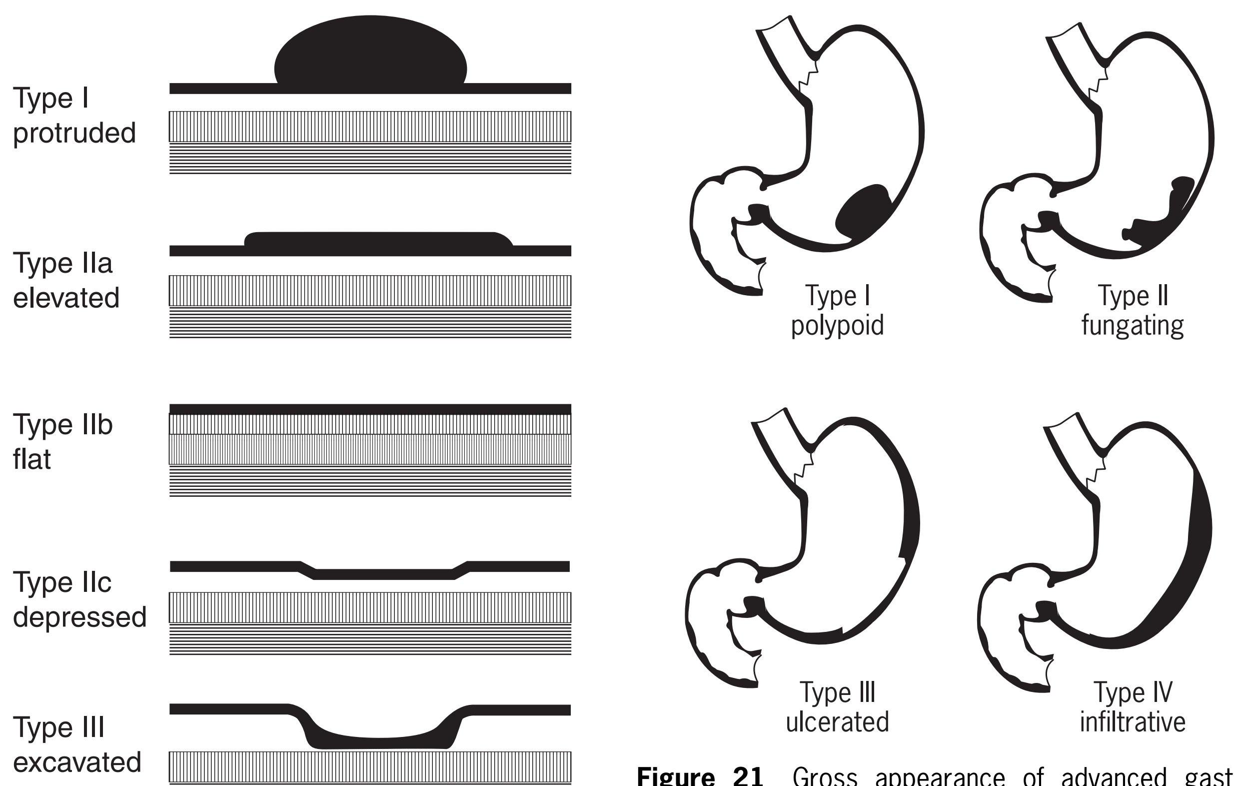


Figure 21 Gross appearance of advanced gastric cancer.

Figure 20 Gross appearance of early gastric cancer.

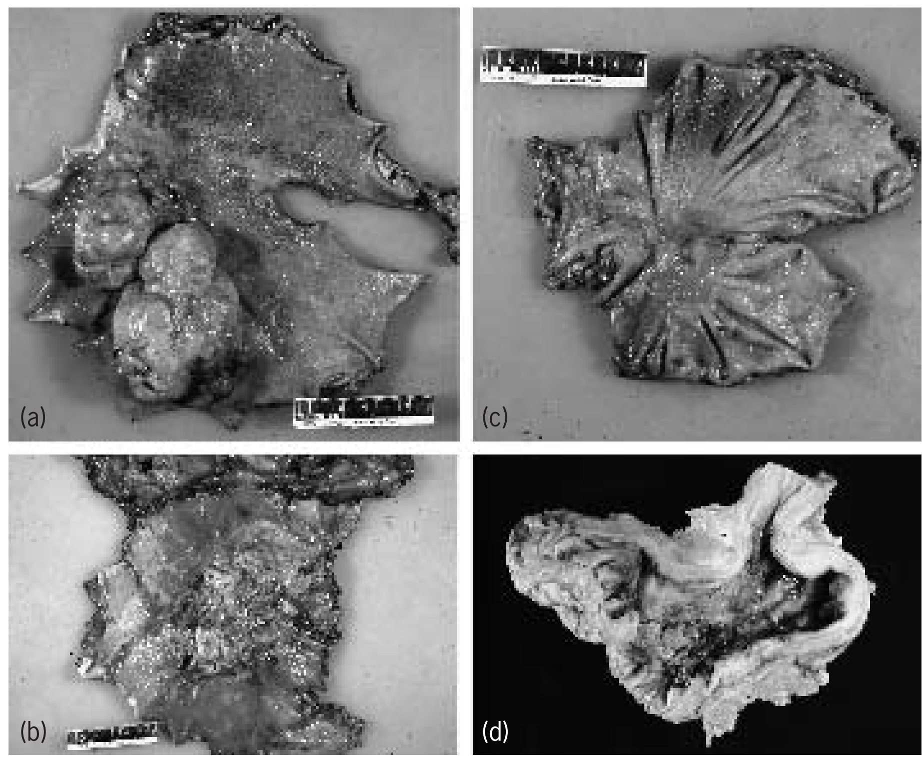
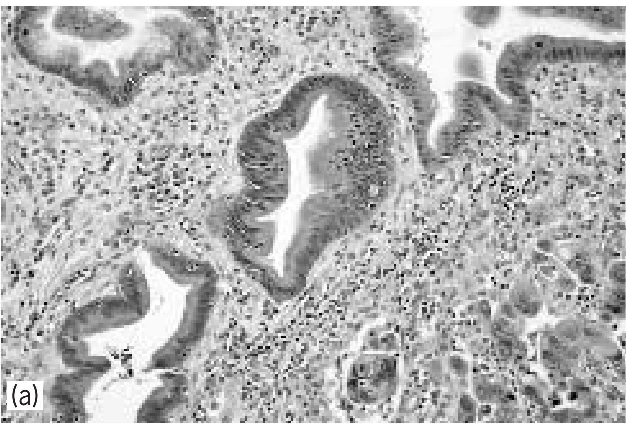


Figure 22 Gross appearance of gastric cancer. (a) Borman type I gastric carcinoma is a bulky tumour protruding into the gastric lumen. (b) A Borman type III lesion shows a heaped-up, ulcerated lesion. (c) A Borman type II gastric carcinoma shows a fungating tumour. (d) Linitis plastica with diffuse thickening of the gastric wall.



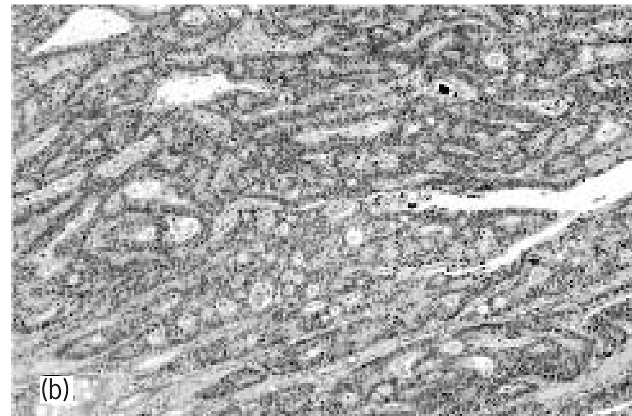
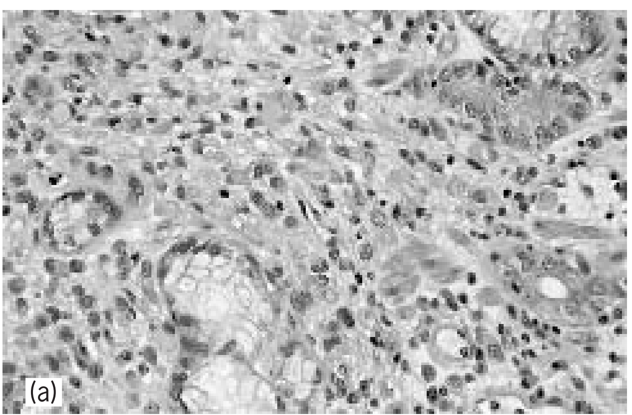


Figure 23 Gastric cancer of the intestinal type. (a) A combination of well- and poorly-differentiated carcinoma. The tumour forms reasonably intact glands everywhere except in the lower right-hand corner where they are discohesive. (b) Back-to-back glands in a moderately differentiated intestinal-type carcinoma.



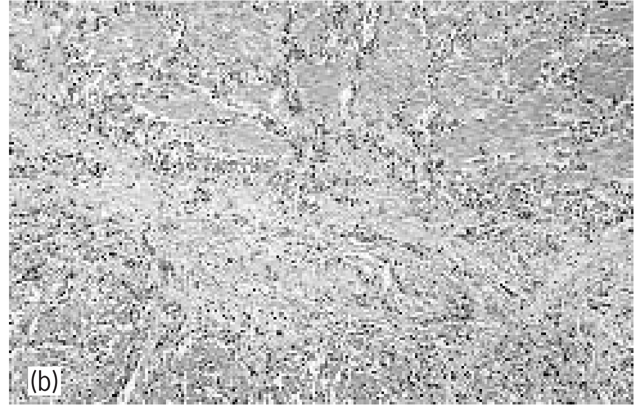


Figure 24 Diffuse gastric carcinoma. In contrast to the cancers illustrated in Figure 23, glandular formations are not present. (a) The tumour cells diffusely infiltrate the intestinal wall. (b) Infiltration of these same cells through the deeper layer of the stomach.

Table 1 World Health Organization classification of histological types of epithelial tumours of the stomach

Adenocarcinoma	Adenosquamous carcinoma
Papillary	Squamous cell carcinoma
Tubular	Undifferentiated carcinoma
Mucinous	Unclassified carcinoma
Signet ring cell carcinoma	

order to reduce the tumour to manageable proportions or may be used to supplement a surgical procedure so as to diminish the risk of early recurrence, so-called adjuvant therapy. Radiation may also be used alone to palliate symptoms in treating patients with inoperable disease.

The relative 5-year survival with oesophageal cancer in the United States for the year ending 1988 was only 8.1%. This is a measure of the grim prognosis that accompanies the late-stage cancers that prevail in that country. Fewer than 3% of patients with late-stage tumours survive for 5 years after diagnosis.

Table 2 TNM classification of oesophageal tumours

Primary tumour (T)	
T1	No obstruction or circumferential involvement
T2	Obstruction and/or circumferential involvement
T3	Extraesophageal spread
Nodal involvement (N) (surgical evaluation)	
N0	No involvement
N1	Positive nodes
Distant metastases (M)	
M0	None
M1	Present
Stage I	T1, N0, M0
Stage II	T1, N1, M0T2, N0, M0
Stage III	Any T3
	Any M1

Adenocarcinoma

Adenocarcinomas extensively infiltrate the oesophageal wall and often show lymphatic and vascular involvement,

as well as direct extension through the oesophageal wall. Both regional and distant metastases develop. Primary oesophageal adenocarcinomas rarely spread distally into the stomach, whereas proximal spread commonly occurs.

About 33% of intramucosal tumours, 67% of intramural and 89% of transmural oesophageal adenocarcinomas metastasize to the regional lymph nodes; 60% of patients with nodal recurrence have recurrence at sites outside the resection margins. Patients with lymphatic metastases can be cured, particularly if fewer than four nodes are involved.

Stomach

The microscopic measurement of the depth of tumour penetration of the stomach wall and quantification of the number of regional lymph nodes involved by cancer are the most consistent methods for predicting the subsequent course of gastric cancer. This method is called the TNM system, as in the oesophagus, and is used to identify the stage of the tumour at the time of diagnosis. This staging system is summarized in **Table 3**. A British study of 31 716 gastric cancers showed that as stomach cancers advance from stage I to IV, the 5-year survival drops from 80% to less than 3%. These data reflect the experience of gastric cancer patients in most Western countries prior to the development of modern diagnostic and therapeutic methods (Craven, 1993). The cancers in this series were so far advanced that only 20% were considered suitable for curative resections. More recent Japanese experience reflects the advantage of making a diagnosis in early stage disease, of more extensive surgery and of adjuvant chemotherapy for advanced disease. A report of 2824 patients treated between 1979 and 1990 noted that 95% were considered suitable for surgical resection, that half of the patients were at Stage 1 and that their overall 5-year survival was 72% (Kinoshita *et al*., 1993).

Surgery remains the principal treatment for gastric cancer; the extent of surgery depends upon tumour stage. Japanese surgeons who have had the opportunity to treat many early cancers have used sonographic techniques to assess the depth of tumour invasion and have excised the tumour from the stomach lining endoscopically, leaving the stomach intact. Unfortunately, this procedure is not applicable to the late-stage tumours typically found in Western patients. Resection of all, or part, of the stomach is required for all advanced, operable tumours. As in the oesophagus, surgery may be supplemented by neoadjuvant or adjuvant radiation and chemotherapy. The relative 5-year survival of stomach cancer patients in the United States, treated in the years 1983–1989 was only 16.9%, reflecting the late stage of the great majority of these cancers at the time of presentation.

Table 3 TNM classification of gastric cancers

1. Primary tumour (T)

The principal factor is the degree of penetration of the stomach wall by carcinoma.

TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma *in situ*: intraepithelial tumour without invasion of the lamina propria
T1	Tumour invades lamina propria or submucosa
T2	Tumour invades the muscularis propria or subserosa
T3	Tumour penetrates the serosa (visceral peritoneum) without invasion of adjacent structures
T4	Tumour invades adjacent structures (spleen, transverse colon, liver, diaphragm, pancreas, abdominal wall, adrenal gland, kidney, small intestine and retroperitoneum)

2. Nodal involvement (N)

The regional nodes include the perigastric nodes along the lesser and greater curvatures and the nodes along the left gastric, common hepatic, splenic and coeliac arteries. Involvement of other intraabdominal nodes represents distant metastasis.

NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in perigastric lymph nodes within 3 cm of the edge of the primary tumour
N2	Metastasis in perigastric lymph nodes greater than 3 cm from the edge of the primary tumour, or metastasis in lymph nodes along the left gastric, common hepatic, splenic and coeliac arteries

3. Distant metastasis (M)

MX	Presence of distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

MOLECULAR GENETIC FINDINGS

Oesophageal Cancer

Squamous Cell Carcinoma

A condition known as tylosis accounts for some forms of familial SCC. This autosomal dominant genetic disorder associates with oesophageal cancer in 90% of affected persons. The cancers usually appear during middle age, but are presaged in early life by marked thickening and fissuring of the palms of the hands and soles of the feet. Tylosis results from a mutation in the *TOC* (tylosis oesophageal cancer) gene on chromosome 17.

Molecular alterations are more commonly investigated in patients with sporadic oesophageal SCC and a number of abnormalities have been detected. These include the accumulation of p53 protein in aggressive cancers. The p53 protein is overexpressed in more than 75% of oesophageal SCC and it may associate with dietary factors. It may also

predict responsiveness to chemotherapy. (The *p53* gene is important in regulating the cell cycle and it plays a critical role in protecting the cell from acquiring genetic damage that can be passed on to daughter cells. When the gene is overexpressed it is usually due to the presence of a mutation that potentially makes the protein nonfunctional and therefore predisposes the altered cells to genetic instability.) Another gene that is commonly abnormal in oesophageal SCCs is the *Cyclin D1* gene (another cell-cycle regulator) which is frequently amplified and overexpressed. Allelic loss on chromosome 13q12–13 may associate with lymph node metastasis in these patients.

Adenocarcinoma

Oesophageal adenocarcinomas associate with a number of molecular abnormalities, some of which resemble those seen in other gastrointestinal adenocarcinomas. Abnormalities in the expression of growth factors and their receptors, especially epidermal growth factor receptor, epidermal growth factor and transforming growth factor alpha, occur commonly and may contribute to nonregulated cell growth. These three substances are overexpressed in Barrett oesophagus and in oesophageal adenocarcinomas and may play a role in the progression of Barrett oesophagus to carcinoma, perhaps by providing a proliferative advantage to metaplastic cells overexpressing the proteins. Another protein related to cell proliferation that becomes abnormal in Barrett oesophagus and the cancers associated with it is p27, which is a cyclin-dependent kinase inhibitor and a negative regulator of cell cycle division. It is found that p27 is inactivated in many Barrett oesophagus-associated cancers and its loss associates with an aggressive tumour behaviour and unfavourable outcome. Other genes that may also be abnormal in patients with oesophageal adenocarcinomas are those which are often altered in colon cancer. These include the *p53* gene mentioned above for oesophageal SCC, the *APC* gene (a gene that predisposes to colon cancer) and the *DCC* (deleted in colon cancer) gene.

Gastric Cancer

There are several forms of hereditary gastric cancer in which the genetic makeup of the patient and the patient's family strongly favour the development of gastric cancer.

Germ-line mutations account for a small proportion of familial gastric cancer clusters, some of which may also be influenced by environmental factors. The hereditary nonpolyposis colon cancer (HNPCC) syndrome is an autosomal dominant disease that carries an increased risk of stomach and colon cancer. It results from a germ-line mutation in one of a number of DNA mismatch repair genes. Affected cells are susceptible to an accelerated accumulation of mutations.

Familial adenomatous polyposis (FAP) resulting from germ-line mutations in the *APC* gene on the long arm of chromosome 5 predispose patients to intestinal polyposis and cancer. In addition, the stomach develops adenomas that may evolve into carcinomas. An environmental influence on the underlying genetic alteration is suggested by the observation that Japanese patients with FAP more frequently have gastric adenomas and carcinomas than Westerners with this trait, reflecting the higher risk of gastric cancer experienced by all Japanese (Utsonomiya *et al*., 1994).

A newly recognized autosomal dominant mutation in the E-cadherin gene was first found in a familial cluster of poorly differentiated gastric cancers in a Maori, New Zealand family. This observation was quickly followed by the observation of other gastric cancer families with germ-line E-cadherin mutations in Europe, Korea and Japan (Gayther *et al*., 1998). The gastric cancer resulting from this defect develops in early life and is diffuse in type. E-cadherin is an adhesion molecule, and loss of its function might explain the discohesive growth pattern of these diffuse tumours.

Genetic polymorphisms may also contribute to gastric cancer risk. Inherited differences in the ability to detoxify potential carcinogens, and inherited variations in the ability to induce transcription or programmed cell death might explain some familial clusters of gastric cancer. The glutathione *S*-transferase enzyme system represents an example of a polymorphism that may favour the development of gastric cancer. These enzymes catalyse the conjugation of numerous carcinogens. Persons lacking the mu form of this enzyme (GSTM1) constitute 40% of the Japanese population and absence of this form of the enzyme increases the risk of developing gastric cancer (Katoh *et al*., 1996). A similar finding has been noted in an English study. Smoking increases the risk of both gastric cancer and its precursor lesion, intestinal metaplasia of the stomach lining. The null variant of this enzyme may contribute to the increased gastric cancer risk associated with smoking due to alterations in the production of potential gastric carcinogens present in tobacco smoke. These interactions may be further modulated by dietary influences.

Polymorphisms present at codon 72 of the *p53* gene may also be important in the development of gastric cancer. Ethnic differences occur in the frequencies of different forms of this codon. The majority of people of European origin are *p53* argenine homozygotes, as are patients with cancer of the cardia of the stomach. Whether the predilection for cardia cancers to occur among white males is due to excess vulnerability derived from *p53* argenine homozygosity, or only reflects parallel, unrelated traits, remains to be determined. The same observation applies to the association of blacks with the presence of a proline allele at codon 72 and antral GC.

CONCLUSION

It is apparent that environmental hazards account for most oesophageal and stomach cancers worldwide. Both tumours are closely related to a poor economic status. Prevention of these cancers is a socioeconomic rather than a medical problem, and should include improved housing, raised living standards, better education and a food distribution system that makes fresh fruit and vegetables available on a year-round basis to all levels of society. Medical intervention is necessary when these ideals are not achieved, but this will only produce a major reduction in mortality rates if the proportion of patients who are treated in the early stages of tumour progression is greatly increased. The Japanese and Chinese community screening programmes for stomach and oesophageal cancer have achieved this result and they are clearly adaptable to other high-risk societies, and the development of focused, tumour-specific, screening programmes is necessary where these cancers are less common. These should be given high priority since the late stages of these tumours have not proved amenable to currently available therapeutic strategies.

REFERENCES

Aoki, K., *et al*. (1992). Death rates for malignant neoplasms for selected sites by sex and five year age group in 33 countries, 1953–57 to 1983–87. *70–155. UICC* (University of Nagoya Press, Nagoya).

Blaser, M., *et al*. (1995). Infection with helicobacter pylori strains possessing cag A is associated with increased risk of adencarcinomas of the stomach. *Cancer Research*, **55**, 2111–2115.

Correa, P. (1988). A human model of gastric carcinogenesis. *Cancer Research*, **48**, 3554–3560.

Correa, P., *et al*. (1973). Pathology of gastric carcinoma in Japanese populations: comparison between Miyagi Prefecture, Japan and Hawaii. *Journal of the National Cancer Institute*, **51**, 1449–1459.

Craven, J. L. (1993). End results of surgical treatment: British experience. In: Nishi, M., *et al*. (eds), *Gastric Cancer*. 341–348 (Springer, Berlin).

Dent, J. (1989). Approaches to oesophageal columnar metaplasia (Barrett's esophagus). *Scandinavian Journal of Gastroenterology*, **168**, 60.

Fenoglio-Preiser, C. M., *et al*. (1999). *Gastrointestinal Pathology, an Atlas and Text*. Oesophageal structure and development, 15–29; gastric structure and development, 133–151 (Lippincott-Raven, Philadelphia).

Fisher, S. G., *et al*. (1993). A cohort study of stomach cancer in men after gastric surgery for benign disease. *Journal of the National Cancer Institute*, **85**, 1303–1310.

Gayther, S. A., *et al*. (1998). Identification of germline E-cadherin mutations in gastric cancer families of European origin. *Cancer Research*, **58**, 4086–4089.

Goodman, K. J. and Correa, P. (1995). The transmission of *Helicobacter pylori*. A critical review of the evidence. *International Journal of Epidemiology*, **24**, 875–887.

Jankowski, J. A., *et al*. (2000). Barrett's metaplasia. *Lancet*, **356**, 2079–2085.

Katoh, T., *et al*. (1996). Glutathione S-transferase M1 (GSTM1) and T1 (GSTT1) genetic polymorphism and susceptiblity to gastric and colorectal adenocarcinoma. *Carcinogenesis*, **17**, 1855–1859.

Kinoshita, T., *et al*. (1993). Treatment results of gastric cancer patients: Japanese experience. In: Nishi, M., *et al.* (eds) Gastric Cancer. 319–330 (Springer, Berlin).

Kitahara, F., *et al*. (1999). Accuracy of screening for gastric cancer using pepsinogen concentration. *Gut*, **4**, 693–697.

Lauren, P. (1965). The two histologic main types of gastric carcinoma: diffuse and so-called intestinal type carcinoma. An attempt at a histo-clinical classification. *Acta Pathologica Microbiologica Scandinavica*, **64**, 31–49.

Lehtola, J. (1978). Family study of gastric carcinoma. *Scandinavian Journal of Gastroenterology*, **13** (Suppl. 50), 1–54.

Lewin, K. and Appleman, H. (1996). Tumors of the esophagus and stomach. In: Rosai, J. (ed.), *Atlas of Tumor Pathology*. 43–144 (Armed Forces Institute of Pathology, Washington, DC).

Munoz, N. and Day, N. E. (1996). Esophageal cancer. In: Schottenfeld, D. and Fraumeni, J. E., Jr (eds), *Cancer Epidemiology and Prevention*. 681–706 (Oxford University Press, New York).

Nomura, A. M. Y. (1996). Stomach cancer. In: Schottenfeld, D. and Fraumeni, J. E., Jr (eds), *Cancer Epidemiology and Prevention*. 707–724 (Oxford University Press, New York).

Nomura, A. M. Y. and Stemmermann, G. N. (1993). *Helicobacter pylori* and gastric cancer – a review article. *Journal of Gastroenterology and Hepatology*, **8**, 294–303.

Qiu, S. and Yang, G. (1988). Precursor lesions of esophageal cancer in high risk populations of Henan Province China. *Cancer*, **62**, 551–557.

Samloff, I. M. and Liebman, N. M. (1973). Cellular localization of group II pepsinogens in human gastric mucosa by immunofluorescence. *Gastroenterology*, **65**, 36–42.

Samloff, I. M., *et al*. (1975). A study of the relationship between serum group I pepsinogen and gastric acid secretion. *Gastroenterology*, **66**, 494–502.

Sandler, R. S., *et al*. (1995). The risk of esophageal cancer in patients with achalsia. *Journal of the American Medical Association*, **274**, 1359–1362.

Sokoloff, B. (1939). Predisposition to cancer in the Bonaparte family. *American Journal of Surgery*, **40**, 673–678.

Stemmermann, G. N. (1994). Intestinal metaplasia of the stomach, a status report. *Cancer*, **74**, 556–564.

Tyuns, A. J., *et al*. (1977). Le cancer de l'oesophage en Ille-et Vilaine en fonction des niveaux de consommation de alcool et

tabac. Des risques qui se multiplient. *Bulletin du Cancer*, **64**, 45–60.

Utsonomiya, J., *et al*. (1994). Hereditary gastric cancer. *Surgical Clinics of North America*, **3**, 545–562.

Zheng, T., *et al*. (1993). The time trend of age-period cohort effects on the incidence of adenocarcinoma of the stomach in Connecticut from 1955–1989. *Cancer*, **72**, 330–340.

Ziegler, R. G. (1986). Alcohol–nutrient interactions in cancer etiology. *Cancer*, **58**, 1942–1958.

FURTHER READING

Powell, S. M. (1997). Stomach cancer. In: Vogelstein, B. and Kinzle, K. W. (eds), *The Genetic Basis of Human Cancer*. 647–650 (McGraw-Hill, New York).

Smalley, S. R. and Williamson, S. K. (1996). Radiation and combined modality therapy for stomach cancer. In: Wanebo, H. J. (ed.), *Surgery for Gastrointestinal Cancer: a Multidisciplinary Approach*. 363–369 (Lippincott-Raven, Phildelphia).

Weber, W., *et al*. (1996). *Familial Cancer Management* (CRC Press, Boca Raton, FL).

Chapter 38

Lower Gastrointestinal Tract

Jeremy R. Jass
University of Queensland, Brisbane, Australia

CONTENTS

- Normal Development and Structure
- Tumour Pathology
- Epidemiology
- Aetiology
- Screening and Prevention
- Preinvasive Lesions
- Gross Appearances
- Histopathology
- Molecular Genetic Findings
- Prognostic Factors
- Clinical Management

NORMAL DEVELOPMENT AND STRUCTURE

Development

The foregut, midgut and hindgut are derived from the embryonic yolk sac, which is of endodermal origin. During embryogenesis, the midgut communicates with the remains of the yolk sac via the vitello-intestinal duct. The midgut is supplied by the superior mesenteric artery and from it develop the third and fourth parts of the duodenum, jejunum, ileum, caecum, appendix, ascending colon and transverse colon. As the midgut lengthens, it becomes coiled and this promotes the development of its dorsal mesentery containing the superior mesenteric artery. Eventually, the midgut can no longer be accommodated in the abdominal cavity and it is extruded into the umbilical cord as a physiological hernia. A process of rotation occurs before the gut is restored to the abdominal cavity and takes up its normal anatomical position.

The hindgut is suspended by a shorter mesentery through which passes the inferior mesenteric artery to supply the descending colon, sigmoid colon and upper rectum. The superior, middle and inferior rectal arteries are derived from the inferior mesenteric, internal iliac and internal pudendal arteries, respectively. At 5 weeks, the distal hindgut, allantois and urogenital tract end in a common cloaca. Downward growth by the urorectal septum occurs until it fuses with the cloacal membrane. Two anal tubercles arise from adjacent ectoderm to fuse with the urorectal septum, forming the proctodeum where ectoderm and endoderm are in direct contact. The proctodeum then fuses with the rectum and canalizes during the third month to become the anus.

Gross Appearance

The small intestine is approximately 500 cm in length (shorter in life owing to tone in the longitudinal muscle coat) and lacking obvious features on external examination apart from its mesentery (the duodenum is retroperitoneal). The large intestine (colon and rectum) is 150 cm in length and variably covered by peritoneum in its different regions. The longitudinal muscle is arranged in three continuous bands (taeniae coli) which fuse into a continuous sheet at the junction of sigmoid colon and rectum. The taeniae coli shorten the colon and produce the characteristic sacculations or haustra. The colon is also distinguished from the small intestine on external inspection by the rows of fat tags called epiploic appendages. The superior and inferior mesenteric arteries anastomose via a continuous marginal artery along the mesenteric border of the colon (**Figure 1**).

Histology

The layers of the intestinal wall from within outwards comprise mucosa, submucosa, muscle coat and serosa or peritoneum (**Figure 2**). The mucosa consists of columnar

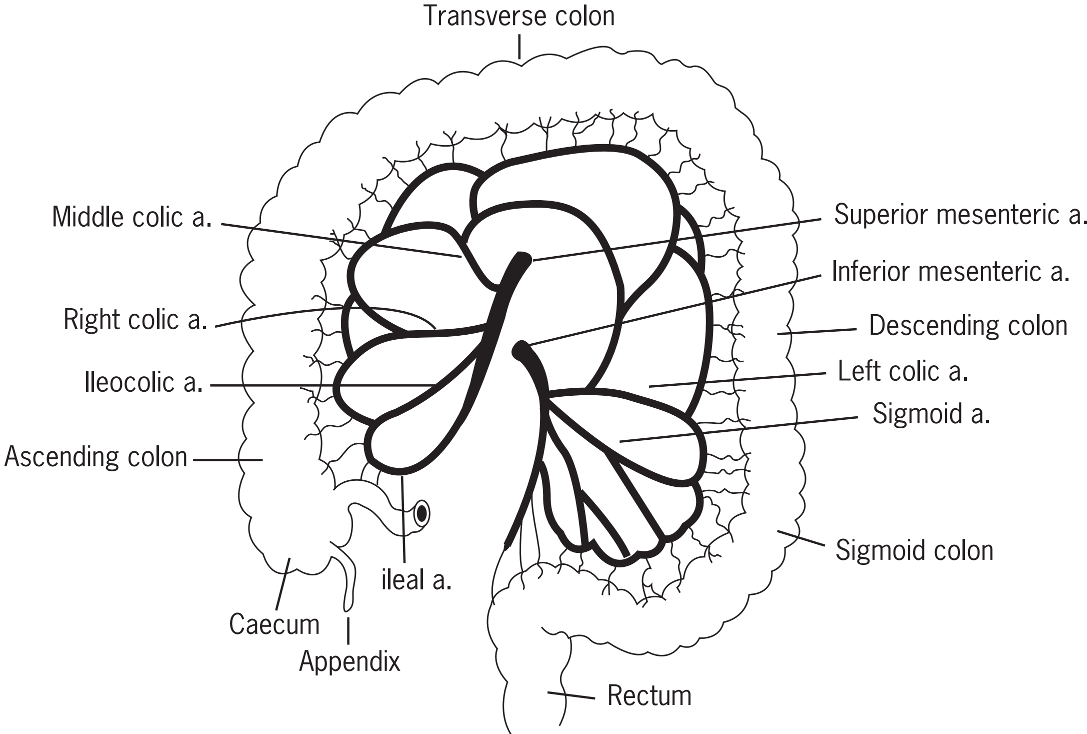


Figure 1 Transverse colon lifted to show arterial blood supply to the large intestine. (a. = artery)

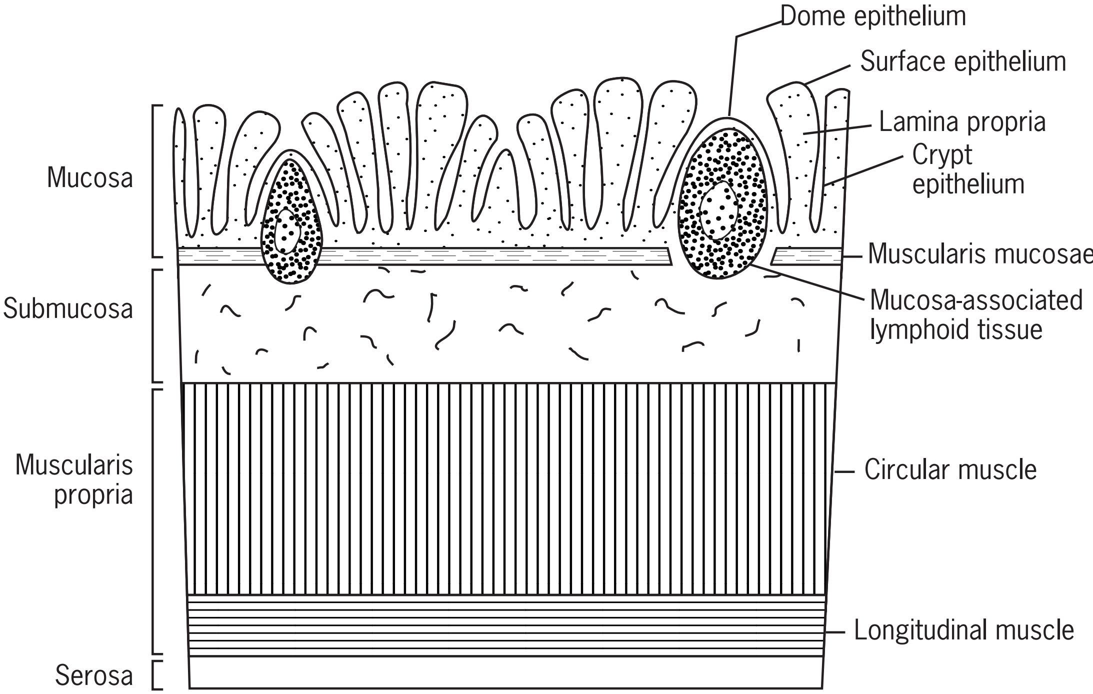


Figure 2 Layers forming wall of large intestine.

epithelium, a delicate connective tissue or lamina propria and a thin sheet of muscle called the muscularis mucosae. The epithelium comprises secretory crypts and an absorptive surface layer. In the small intestine, but not the colon or rectum, the area of the surface epithelium is greatly increased by the formation of finger-like villi.

The columnar epithelium comprises columnar cells, goblet cells, endocrine cells and Paneth cells (the latter are limited to the small intestine and proximal colon). These cells are derived from a common precursor or stem cell. The lower crypt includes the replicative compartment. Maturing cells migrate up the crypt and reach the surface epithelium within a matter of days where they undergo apoptosis.

TUMOUR PATHOLOGY

In this section, a brief overview will be given of the types of tumour that may occur in the lower intestinal tract, both cancerous and benign. The rest of this chapter will focus upon the most common and important cancer: carcinoma of the colon and rectum.

Types of Tumour

The types of tumour reflect the normal tissues that contribute to the structure of the intestinal tract (Jass and Sobin, 1989).

Epithelial Tumours

The columnar or glandular epithelium of the intestinal tract gives rise to benign and malignant neoplasms termed adenoma and adenocarcinoma, respectively. The colon and rectum are by far the most common sites of both adenoma and adenocarcinoma in the lower intestinal tract. Most adenocarcinomas arise in a pre-existing adenoma and these lesions may be viewed as a neoplastic continuum. However, only about 5% of adenomas will transform into adenocarcinomas. Multiple adenomas occur in the colon, rectum, duodenum and appendix in the rare autosomal dominant condition familial adenomatous polyposis (FAP) which is caused by mutation of the Adenomatous Polyposis Coli (*APC*) gene.

Certain types of epithelial polyp that have traditionally been regarded as non-neoplastic may occur in the intestinal tract. They are noted here (**Table 1**) because it is now appreciated that these lesions may show clonal genetic alterations and under certain circumstances may be precancerous. These polyps may be either single, occur in small numbers or be so numerous that the term polyposis is used. The commonest epithelial polyp is the hyperplastic polyp and occurs principally within the distal colon and rectum. These small polyps are characterized histologically by a saw-tooth or serrated crypt outline. While regarded as innocuous lesions, colorectal cancer has been associated with the rare condition hyperplastic polyposis in which the polyps are large and distributed throughout the colorectum.

Two types of epithelial polyp are classified as hamartomas (a maldevelopment in which epithelial and connective tissues are arranged in a haphazard manner). One of these occurs as part of the autosomal dominant Peutz–Jeghers syndrome (germ-line mutation of *STK11*). Peutz–Jeghers polyps are multiple and tend to be most numerous within the small intestine. Juvenile polyps are the second type of hamartoma and occur in the colon and rectum either singly, in small numbers or multiply in the autosomal dominant condition juvenile polyposis (germ-line mutations identified in *DPC4/Smad4*). Both Peutz–Jeghers syndrome and juvenile polyposis are associated with an increased risk of intestinal cancer. Endocrine cell tumours are considered in the chapter Systemic Oncology of the Endocrine Organs.

The lower anal canal is lined by squamous epithelium. Squamous cell carcinoma may arise from this site. Although uncommon, the frequency of this cancer is increasing in both males and females. Human papillomavirus (types 16 and 18) has been implicated in the aetiology; risk factors and mechanisms are similar to those underlying squamous cell carcinoma of the female genital tract (see the chapter *Systemic Oncology of the Female Reproductive System*).

Table 1 Nature and distribution of epithelial polyps of the intestinal tract

Type	Nature	Site[a] Small intestine	Appendix	Colon	Rectum
Adenoma Tubular Tubulovillous Villous	Neoplastic	S	S	U	U
Hyperplastic	Unknown	N	S	S	U
Peutz-Jeghers	Hamartoma	U	N	S	S
Juvenile	Hamartoma	S	N	U	U
Inflammatory	Inflammatory	N	N	U	S

[a]U = usual site; S = site of occurrence; N = not found (or uncommon).

Lymphoid Tumours

Lymphoid tissue is distributed throughout the intestinal tract with notable collections in the terminal ileum known as Peyer's patches. Mucosa-associated lymphoid tissue (MALT) is organized differently from nodal lymphoid tissue in terms of both structure and function. Lymphoid neoplasms are malignant by definition and those arising in MALT are known as MALTomas and are derived from B-lymphocytes. The stomach does not normally contain lymphoid tissue, yet MALTomas are more common in this organ than the intestine (see the chapter *Systemic Oncology of the Upper Gastrointestinal Tract*). A special type of MALToma occurs in the small intestine and affects populations around the Mediterranean and in the Middle East. The neoplastic B-lymphocytes secrete α heavy chains (a component of immunoglobulin class A) and the lymphoma is also known as α heavy chain disease or immunoproliferative small intestinal disease (IPSID). Two additional B-lymphocyte lymphomas of the intestinal tract are Burkitt lymphoma occurring mainly in Africa and principally affecting the ileocaecal region and mantle cell lymphoma presenting as multiple mucosal polyps throughout the gastrointestinal tract. There is also a rare T-cell lymphoma which mainly affects the proximal jejunum and is most common in people of Northern Europe. Many of these lymphomas are associated with gluten-sensitive enteropathy or coeliac disease.

Stromal Tumours

Despite the large amount of smooth muscle and connective tissue in the wall of the intestinal tract, stromal tumours are uncommon. Gastrointestinal stromal tumours (GISTs) were considered to be tumours of smooth muscle in the past, but most are now known to express the phenotype of a pacemaker cell found in the muscle coat (interstitial cell of Cajal). Behaviour is unpredictable but large size and a high mitotic rate are markers of malignancy. Bona fide smooth muscles occur in the rectum; most are small and benign.

Secondary (Metastatic) Cancer

The intestinal tract is not a common site of metastatic cancer. The small intestine is the main site for metastases and melanoma, breast and lung cancer are the principal primary sources.

EPIDEMIOLOGY

Colorectal cancer is a leading cause of cancer death, second only to lung cancer in the West. It is responsible for 55 000 deaths per year in the USA. There were around 875 000 new cases worldwide in 1996, accounting for 8.5% of all cancers diagnosed in that year. Incidence rates vary by at least a factor of 20 between countries, with the highest rates being found in the West and the lowest in the developing world (Potter, 1999). Although these differences are exaggerated by competing causes of mortality in developing countries, such as infectious disease and other types of cancer, age-adjusted statistics show that the low incidence is genuine. Colon and rectal cancers share many environmental risk factors, but differences exist also. Colonic cancer occurs with approximately equal frequency in males and females whereas rectal cancer is at least twice as common in males.

Time trends indicate rapid increases in incidence in countries or populations that have adopted Western lifestyles in recent years. Examples include Japan, urban China and male Polynesians in Hawaii. Colonic cancer seems to be more sensitive to these changes than rectal cancer. Migrant studies have shown that populations migrating from low-risk to high-risk countries rapidly acquire the incidence rates of the adopted country, sometimes within the migrating generation. For example, one of the highest rates in the world is seen among Hawaiian Japanese (Haenszel, 1961).

Colorectal cancer is age related, colonic cancer more so than rectal. There will be competing causes of mortality in ageing populations and the mean age at presentation in most hospital series falls within the range 65–75 years. The age relationship, common to many types of solid cancer, is a reflection of the time-related multistep evolution of cancer. The stepwise accumulation of genetic errors in somatic cells requires a particular passage of time. In individuals who inherit one of the causative genetic errors, each somatic cell is already primed (or one step ahead). Such individuals develop multiple neoplasms and at an early age. This is observed in several forms of hereditary colorectal cancer (see below).

AETIOLOGY

The aetiology of colorectal cancer is not known but the brief account of descriptive epidemiology given above indicates the involvement of both environmental influences and genetic factors. The major shifts in incidence following migration were once taken as evidence of the overwhelming importance of the environment. Nevertheless, individuals inhabiting high-risk areas may live to advanced years without even developing a minute adenoma (the earliest visible signs of colorectal neoplasia). Conversely, an inherited genetic mutation may lead to early onset disease in an inhabitant of a low-risk area. The modern approach to unravelling the aetiology of colorectal cancer integrates lifestyle factors and genetic constitution within the discipline of molecular epidemiology.

Dietary and Lifestyle Factors

Vegetables and Fibre

Particular dietary practices have been associated with colorectal cancer in defined groups of subjects (cohort studies) or comparing affected cases with matched controls (case-control studies) (Trock *et al*., 1990). However, these studies demonstrate associations rather than specific causality. Diets high in vegetables (raw, green or cruciferous) are protective although the nature of the specific agent is unclear. High on the list is fibre (the complex carbohydrate constituent of plant cell walls), which could act by diluting or binding potential carcinogens in the lumen of the bowel. Fibre is poorly digested in the small intestine but is fermented by bacteria in the proximal colon to generate short-chain fatty acids, notably butyrate. Resistant starch (that is not digested in the small bowel) is another source of butyric acid. Butyrate serves as an essential respiratory nutrient for colonic epithelium and its lack may be a factor in the aetiology of colorectal cancer. Butyrate may also serve as a protective factor in its capacity as a differentiating agent. Vegetables are high in folate and absence of this vitamin may also serve as a risk factor.

Meat and Fat

Estimates of risk of colorectal cancer have been either increased or null in cohort and case-control studies examining the role of dietary meat and fat. It appears that total protein consumption is relatively unimportant whereas meat preparation (processing or heavy cooking) may generate carcinogens such as heterocyclic amines. Additionally, saturated fat of animal origin may be associated with increased risk (Potter, 1999).

Calcium and Bile Acids

Epidemiological studies and intervention data indicate a protective role for dietary calcium. A possible mechanism may be the conversion of ionized fatty acids and faecal bile acids into insoluble salts. Deoxycholic and lithocholic acid are bacterially deconjugated bile acids that have been shown to serve as promoters or co-mutagens in experimental studies of colonic carcinogenesis. Faecal bile acid levels have in turn been correlated with meat consumption.

Selenium

This essential trace element is a component of the enzyme glutathione peroxidase which catalyses the removal of intracellular hydrogen peroxide. Deficiency of selenium occurs when diets lack whole grains and vegetables or when soil levels are low. A protective role is supported by epidemiological and intervention studies.

Smoking and Alcohol

Higher risks of colonic (not rectal) cancer and adenomas have been associated with a long history of smoking. The association with alcohol is less clear, beer being implicated in rectal cancer in males.

Non-steroidal Anti-inflammatory Drugs (NSAIDS)

NSAIDS, including aspirin, have been shown to lower the risk of colorectal cancer in cohort and case-control studies. Sulindac causes adenomas to regress in FAP and stabilizes the disordered apoptotic ratio that occurs in the normal colorectal mucosa of subjects with FAP. (See the chapter *Dietary Genotoxins and Cancer*.)

Genetic Factors–High-prevalence Polymorphisms

N-Acetyltransferases (NAT1, NAT2) and Cytochrome P450 (CYP) Enzymes

The weak effects of diet upon cancer risk might be increased by genetic factors. For example, at least three enzymes (NAT1, NAT2 and CYP1A2) influence the metabolism of heterocyclic amines produced in the course of cooking meat at high temperatures and hydrocarbons contained in tobacco smoke. Particular genetic polymorphisms affect the activities of these enzyme systems, but consistent correlations between genotypes, lifestyle factors and risk have not been demonstrated to date (Potter, 1999).

Methylenetetrahydrofolate reductase (MTHFR)

Folate may serve as a protective factor through influencing methylation of DNA and the size of the nucleotide pool required for DNA synthesis and repair. The polymorphic enzyme MTHFR appears to increase the risk of colorectal neoplasia in subjects with a particular genotype (TT) and a low-folate diet (Potter, 1999). Some tumour suppressor genes are silenced through hypermethylation of the promoter region. However, this mechanism applies mainly to the subset of colorectal cancers showing DNA microsatellite instability (MSI).

Genetic Factors–Rare Inherited Syndromes

Familial Adenomatous Polyposis (FAP)

About 1 in 8000 individuals carries a mutation in the tumour suppressor gene *APC*, either inherited (75%) or acquired as a new mutation (25%). The disease manifests as an autosomal dominant trait in which affected individuals develop many hundreds and usually thousands of colorectal adenomas in their teens. By the age of 50, one or

more of these polyps is likely to have transformed into a carcinoma. The severity of the disease in terms of polyp numbers depends upon the nature of the germline mutation and modifying influences of other genes. Extracolonic manifestations occur, notably duodenal adenoma and carcinoma, fibromatosis (aggressive but non-metastasizing tumour composed of fibroblasts), benign bone tumours, sebaceous cysts of the skin, pancreatic cancer, thyroid cancer, hepatoblastoma (primary liver tumour recapitulating embryonic liver) and tumours of the central nervous system.

Inactivation of the normal or wild-type *APC* gene, by either mutation or loss, is sufficient to initiate and allow further growth of adenomas. The germ-line mutation leads to subtle growth disturbance of the normal colorectal mucosa by a dominant negative effect (not requiring mutation or loss of the normal copy of the gene). These changes include an increased rate of crypt fission and a disordered distribution of cells undergoing apoptosis (more in the crypt base and less on the epithelial surface) (Wasan *et al*., 1998; Keller *et al*., 1999). The *APC* gene mediates its normal function (the control of growth, differentiation and proliferation) through the *wnt* signalling pathway.

I1307K Mutation of APC

This mutation, found only in Ashkenazi Jews, creates a repetitive poly-A sequence that is subject to replication error, in turn causing protein truncation. Small numbers of adenomas and cancers occur in affected family members.

Hereditary Non-Polyposis Colorectal Cancer (HNPCC)

This autosomal dominant condition is caused by mutation of one of a family of DNA mismatch repair genes. The most frequently implicated are *hMLH1* and *hMSH2* followed by *hMSH6*, *hPMS1* and *hPMS2*. The germ-line mutation is nearly always inherited rather than new to an individual. Nevertheless, an obvious family history may not be apparent in small families as penetrance (the extent to which a mutated gene is expressed phenotypically) is not 100%. The frequency of disease causing mutations in particular populations may be exaggerated by founder effects (where a particular affected individual has been succeeded by multiple, large generations within a relatively sparsely populated region). Approximately 2% of individuals developing bowel cancer in the West do so on the basis of HNPCC. Since around one in 20 individuals will develop colorectal cancer, the population frequency of disease-causing mutations must be around 1 in 1000.

The first step leading to neoplasia in HNPCC is somatic mutation or loss of the second (wild-type) mismatch repair gene. Some genes, such as *TGFβRII* and *BAX*, contain short repetitive tracts that serve as targets for mismatch errors occurring during DNA replication. If the DNA repair mechanism is not operating adequately, the mismatch mutations will not be repaired. Particular mutations, for example in *TGFβRII*, will initiate the development of a neoplastic clone. Most mismatch errors, however, occur in non-coding repetitive tracts or microsatellites. The resulting microsatellite instability does not drive carcinogenesis but is a useful biomarker for HNPCC.

Approximately 70% of subjects carrying an HNPCC germ-line mutation will develop colorectal cancer by the age of 65 years and at a mean age of 45 years. Extracolonic cancers may develop in the uterus (endometrium), ovary, stomach, small intestine, central nervous system and urinary tract (pelviureter). Benign sebaceous gland adenomas may occur in the facial skin (the combination of sebaceous adenoma and colorectal cancer has been called Muir–Torre syndrome but is part of the HNPCC spectrum). The majority of colorectal cancers arise in adenomas. Although adenomas do not occur in large numbers in HNPCC, an individual adenoma carries a high risk of progressing to cancer and does so within a short timeframe. This accounts for cancer multiplicity and the clinical impression of *de novo* origin of cancer. Paradoxically, the colorectal cancers in HNPCC are relatively non-aggressive in terms of their potential for metastasis. Around 60% of colorectal cancers arise in the proximal colon (caecum, ascending or transverse colon) (Lynch *et al*., 1996).

Germline Mutation of TGFβ Type II Receptor

This is a rare cause of familial colorectal cancer.

Chronic Inflammation

The risk of colorectal cancer is increased in subjects with chronic or longstanding inflammation of the large intestine, notably due to ulcerative colitis and Crohn's disease. A possible mechanism may be genetic damage by reactive oxygen species (superoxide, H_2O_2 and hydroxyl radicals), in turn generated by the inflammatory mediators released by inflammatory cells such as neutrophils and macrophages. Cycles of inflammatory damage and healing may also increase the rate of epithelial proliferation or change the microenvironment so that cells are more susceptible to the effects of luminal carcinogens.

SCREENING AND PREVENTION

The prevention of colorectal cancer may be achieved by primary means including lifestyle adjustments or taking preventive medication (chemoprevention) or by screening asymptomatic subjects for risk factors that can be reduced or neutralized. The concept of prevention can be broadened to include the prevention of radical surgery (since

early stage colorectal cancer can be cured by local or conservative surgery) or prevention of death (since advanced cancer can be cured if it is treated at a sufficiently early stage). Screening can result in unnecessary investigation and promote needless anxiety if the test lacks specificity and is therefore associated with a high false-positive rate. Furthermore, if a cancer is detected earlier than it would otherwise have been but not early enough to allow spread and death to be prevented, the test will not have improved mortality rates. The resultant increase in survival is apparent only. Bringing forward the diagnosis of cancer in this way is known as lead time bias. The main approaches to screening include: (1) testing faeces for occult blood, (2) endoscopic examination of the mucosal lining of the large bowel and (3) demonstration of a high-risk genetic mutation.

Faecal Occult Blood Testing

The premise underlying this test is that larger adenomas and early cancers may bleed to a minor extent that goes unnoticed but is identified by a sensitive test for haemoglobin. A positive result is followed by a colonoscopy and removal of small lesions. Larger lesions may require surgery. Cancers detected in asymptomatic subjects are relatively early (in terms of their limited extent of spread) and the net result of removing adenomas and early cancers is a reduction in mortality in the study group as compared with a matched unscreened group. The reduction in mortality from colorectal cancer in randomized controlled trials has ranged from 15% to 33%. Amongst fully compliant participants the mortality reduction is of the order of 40% (Hardcastle *et al*., 1996; Kronborg *et al*., 1996).

Occult blood is not specific for neoplasia and neoplasms may not necessarily bleed. Specificity and sensitivity might be increased by testing for other products such as mutated DNA. K-*ras* mutations can be detected in stool samples but only a subset of colorectal cancers has K-*ras* mutation. One would need to test for a range of mutations to pick up all cancers, but the laboratory costs would be considerable.

Endoscopy

There is retrospective evidence as well as evidence from prospective adenoma follow-up trials that adenoma removal prevents colorectal cancer. Endoscopy allows precancerous adenomas to be visualized and removed before they can transform into cancers. Since no more than 5% of adenomas will become cancerous, it is necessary to destroy or remove 20 to prevent one cancer. The cost effectiveness is increased by offering endoscopy only to subjects at moderate to high risk (e.g. by virtue of a positive family history). However, most colorectal cancers occur in subjects without a strong family history. At least 50% of colorectal cancers occur within the range of the flexible sigmoidoscope. It has been suggested that a single sigmoidoscopic examination between the ages of 55 and 64 years would have an important impact on morbidity and mortality due to cancer of the rectum and sigmoid colon (Atkin *et al*., 1998). Retrospective data suggest that subjects with small adenomas in the distal colon and rectum are at little risk of developing further neoplasms or of having more proximal adenomas. However, subjects with larger and/or multiple adenomas would require a full colonoscopy and additional colonoscopic surveillance (Atkin *et al*., 1992).

Some would argue that screening only the rectosigmoid region is unacceptable and that a full screening colonoscopy should be offered as the first step in screening. The drawbacks are that colonoscopy is more expensive and technically demanding and proximal cancers are less aggressive and more age-related. An alternative approach is 'virtual colonoscopy' achieved by computed tomography colonography which can detect remarkably small lesions. Colonoscopy would then be required to biopsy and treat the lesion.

Colonoscopic surveillance is also offered to subjects with chronic inflammatory bowel disease, particularly ulcerative colitis. The precancerous lesion is known as dysplasia, which appears macroscopically as a flat, velvety area or a raised sessile mass or may be invisible. The risk of associated or subsequent cancer is then high and colectomy is indicated.

Genetic Screening and Predictive Testing

This is offered to subjects with autosomal dominant syndromes predisposing to colorectal cancer: familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC). The steps are similar in both conditions. After counselling and obtaining consent the causative mutation is detected in an affected family member. This is achieved in about 70% of FAP families and 50% of HNPCC families. After counselling, DNA samples from at-risk family members are screened for the mutation. Mutation-negative subjects are at no increased risk compared with the general population and require no further preventive management. Prophylactic surgery is required for FAP and this is generally performed in the mid-teens. The operative approach is either total colectomy with ileorectal anastomosis or proctocolectomy with construction of an ileo-anal pouch anastomosis. The former is preferred unless there is an extensive carpet of rectal adenomas, but the rectum will then require endoscopic surveillance. Rectal excision with ileo-anal pouch anastomosis will often be performed in later years. At-risk or mutation positive members of HNPCC families require one- to two-yearly colonoscopic surveillance from the third decade. If surgery is undertaken for a large adenoma or cancer, a total colectomy is the preferred

option because of the high risk of developing additional cancers.

Dietary Intervention

A measure that is likely to be associated with general compliance is the addition to the diet of a cheap and palatable ingredient, particularly if this can be added safely to a product that is widely consumed anyway. Of the possible protective dietary factors that have been identified so far, resistant starch would be one of the easiest to accommodate in the diet. A rich natural source is unripe bananas, but resistant starch can be added to bread or sprinkled on food without altering its taste.

Chemoprevention

Cohort and case-control studies have demonstrated the protective role of aspirin and other non-steroidal anti-inflammatory drugs (NSAIDS) (Potter, 1999). These are believed to act through the inhibition of cyclo-oxygenase (COX) which is involved in the synthesis of prostaglandin (converts arachidonic acid to prostaglandin G_2). The more specific COX antagonists (sulindac) have been shown to cause adenomas to regress in FAP and to restore the altered apoptotic index in the normal-appearing mucosa of FAP subjects to normal. The molecular mechanisms underlying these interesting observational data are unknown. A drawback of NSAIDs is their gastrointestinal side-effects. COX-2 is not expressed by normal colonic epithelial cells but is upregulated in cancer cells. The latest generation of selective COX-2 antagonists was developed in order to achieve therapeutic or preventive effects without accompanying gastrointestinal disturbance. At this stage it is not known if the more selective components will be as effective as their broad-spectrum counterparts, but preliminary data are encouraging.

The most obvious use of chemoprevention is in subjects at high risk of colorectal cancer. Cost, safety and efficacy will determine if such an approach to cancer prevention can be generalized. Subjects with FAP have been observed to develop bowel cancer while taking sulindac.

PREINVASIVE LESIONS

Adenoma

The most important preinvasive lesions in the colorectum in terms of frequency and potential for malignant change are the benign neoplasms or adenomas (Muto *et al*., 1975). The evidence is as follows: (1) adenomas show a spectrum of changes ranging from low-grade dysplasia through to high-grade dysplasia (or carcinoma-*in-situ*), (2) longitudinal studies demonstrate malignant transformation with time (in situations where the adenoma is not removed), (3) adenoma and carcinoma share similar demographic data and risk factors, (4) removal of adenomas reduces the frequency of cancer and (5) the genetic changes in adenomas are also present in carcinomas.

Macroscopic Features

Adenomas mainly present as polypoid growths that may be sessile or pedunculated. They are usually sessile elevations when less than 5 mm but increasing growth is associated with the formation of a stalk composed of normal mucosa and submucosa. A minority remain as flat or even depressed lesions which may be difficult to detect at colonoscopy without the use of dye spraying and high-resolution magnification. The head is darker than the surrounding normal mucosa in larger adenomas and becomes lobulated, resembling a baby cauliflower. A rare presentation is as a large sessile (broad-based) mass with a soft, shaggy surface. These are described as villous adenomas, although the finger-like villi seen in two-dimensional sections are in reality leaf-like folds.

Microscopic Features

Adenomas are typed as tubular, tubulovillous and villous according to the predominant architectural pattern (**Figure 3**) (Jass and Sobin, 1989). Tubules are lined by columnar epithelium and embedded within lamina propria where they proliferate by branching. Villi (in reality leaves) comprise a covering of columnar epithelium and a core of lamina propria. By forming complex, brain-like folds of epithelium, the surface area of a villous adenoma may be considerable and lead to significant loss of fluid and electrolytes. Tubulovillous adenomas combine both architectural patterns.

The morphological changes that distinguish adenoma from normal include disordered architecture, cytological atypia and abnormal differentiation. The combination of changes has been described as dysplasia and may be graded as mild, moderate and severe or (in a two-grade system) low grade and high grade (**Figure 4**). Mild or low-grade dysplasia deviates little from the normal whereas severe or high-grade dysplasia approximates to carcinoma-*in-situ*. The terms intra-epithelial neoplasia and dysplasia are synonymous. Because of its aggressive connotation the term 'carcinoma-*in-situ*' is generally omitted from diagnostic reports. The risk of cancer developing in an adenoma is associated with architecture (extent of villosity), grade of dysplasia and size of the adenoma (Muto *et al*., 1975).

For the purposes of drawing a diagnostic distinction between adenoma and adenocarcinoma, a rule has been developed that is unique to the lower intestinal tract. In order to diagnose cancer there must be invasion across the line of the muscularis mucosae into the underlying submucosa. For other epithelial surfaces, cancer is diagnosed when there is invasion across the basement membrane. The reasons for this difference in approach are pragmatic and practical. First, the potential to metastasize is not realized

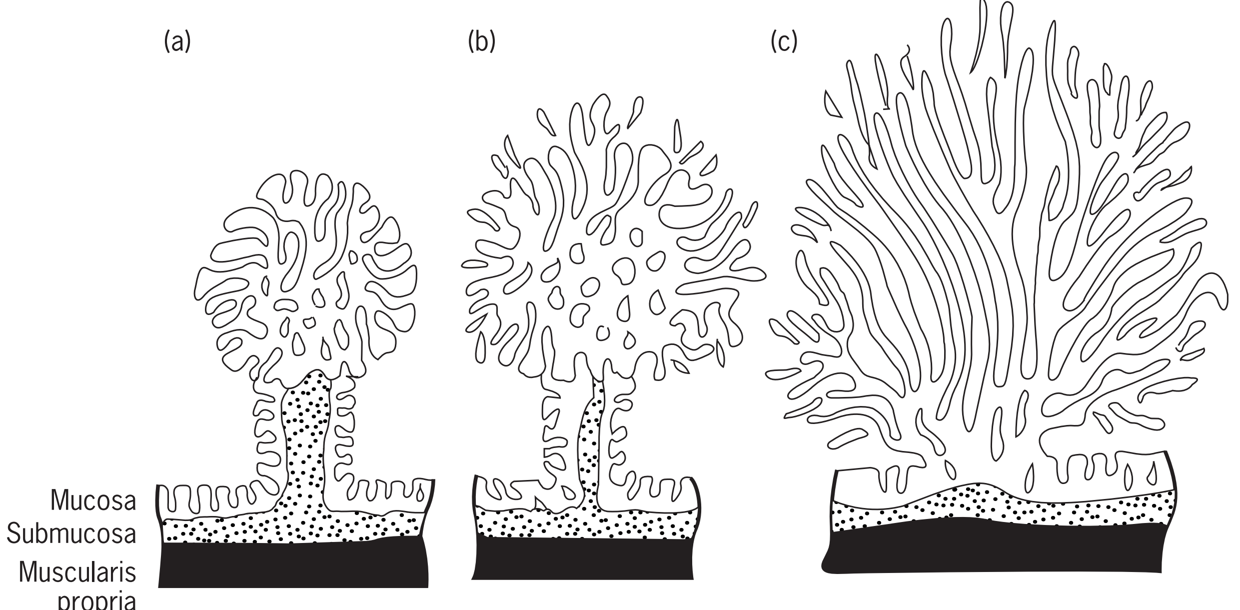


Figure 3 (a) Tubular, (b) tubulovillous and (c) villous adenoma of large intestine.

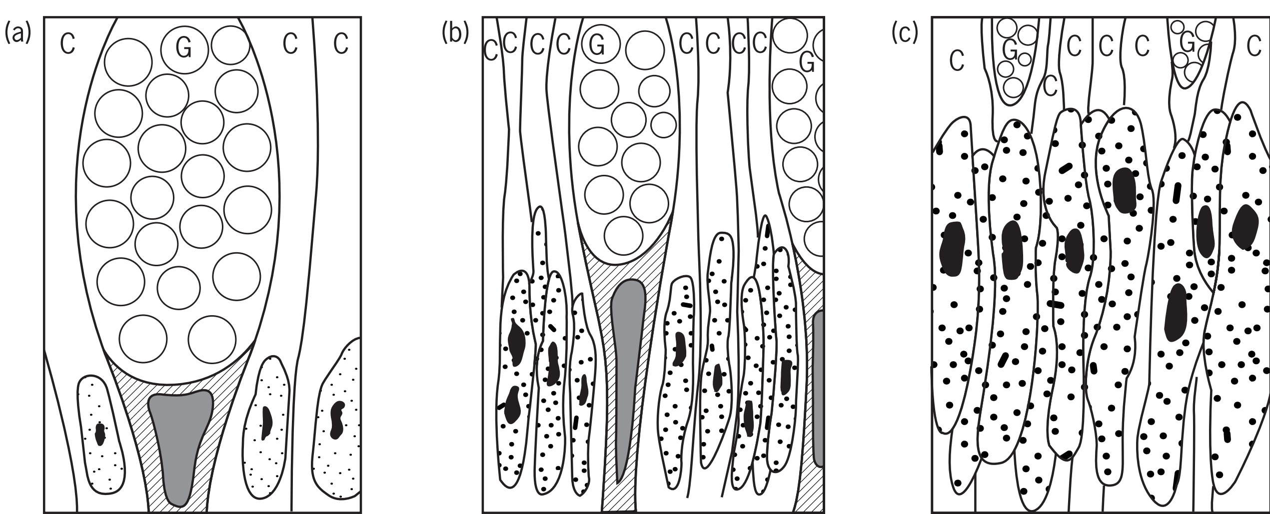


Figure 4 Normal epithelium of (a) large bowel, (b) low-grade dysplasia and (c) high-grade dysplasia. There is progressive loss of differentiation of goblet cells (G) and columnar cells (C) accompanied by nuclear enlargement, crowding, stratification and hyperchromatism.

until there has been submucosal invasion. Second invasion across the basement membrane into surrounding lamina propria is both unusual and difficult to diagnose at the light microscopic level. Invasion across the basement membrane is generally apparent only in the case of poorly differentiated adenocarcinoma, but it is unusual for early cancer arising in an adenoma to be poorly differentiated.

Early Morphogenesis

In squamous and transitional epithelial surfaces, a neoplastic clone is conceived as arising within and replacing the normal epithelium by lateral spread. While such a process may also occur within columnar epithelium, at least some if not the majority of colorectal adenomas originate through a different mechanism. Micro-reconstruction studies in the normal-appearing mucosa from subjects with familial adenomatous polyposis have demonstrated neoplastic crypts occurring as a bud from the side of a normal crypt, in other words recapitulating the normal process of crypt fission (Nakamura and Kino, 1984). The bud forms a tubule as it migrates up the crypt column to open finally on to the epithelial surface. This unicrypt adenoma then undergoes fission to form a microadenoma (**Figure 5**).

(a) (b) (c) (d) (e)

Figure 5 Formation of adenoma as an outpouching from normal crypt.

It is possible to identify microscopic epithelial lesions by their abnormal crypt openings using a combination of dye and magnification of the surface epithelium. The choice of dye will depend on whether this is performed by dissecting microscope *in vitro* (e.g. methylene blue) or during colonoscopy (e.g. indigo carmine). These microscopic lesions are known as aberrant crypt foci. A minority are microadenomas; most are minute hyperplastic polyps.

Serrated Adenomas

This subgroup of adenomas is distinguished from the usual adenoma by combining the serrated or saw-tooth crypt outline of the hyperplastic polyp with the cytological changes of an adenoma. Additionally, the proliferative compartment of a serrated adenoma remains in the lower crypt whereas proliferative cells are found within the upper crypt and surface epithelium in the usual adenoma. Serrated adenomas are also characterized by an overexpression of secretory or gel-forming mucins including MUC2 (intestinal) and MUC5AC (gastric). Serrated adenomas are probably more closely related to hyperplastic polyps than adenomas. Mixed polyps including both hyperplastic and serrated adenomatous components may be found. Furthermore, microsatellite markers have demonstrated clonal relationships between the two components. The high frequency of DNA microsatellite instability also distinguishes serrated adenomas from traditional adenomas (Iino *et al*., 1999).

Dysplasia

In inflammatory bowel disease, dysplasia occurs as an ill-defined or diffuse lesion that may be flat or raised as a sessile mass but lacking the sharp demarcation of an adenoma. Although the term dysplasia has been applied to both the diffuse lesions found in inflammatory bowel disease and the circumscribed neoplasia of an adenoma, differences between the two forms of dysplasia exist with respect to histological appearances, molecular genetics and natural history.

GROSS APPEARANCES

Colorectal cancers are relatively well circumscribed with little growth beyond their macroscopically visible borders. A cancer may appear as a mass protruding into the bowel lumen, as an ulcer with a raised edge or as a band-like stricture causing narrowing of the bowel. Protuberant masses are more common in the caecum and ascending colon. The bowel contents are fluid in this region and obstruction is uncommon. Chronic bleeding from the ulcerated surface leads to the symptoms of iron deficiency anaemia and palpation of a mass in the right iliac fossa will point to the likely diagnosis. Cancers arising in the splenic flexure and left colon are often associated with stricturing that leads to the symptoms of partial or complete bowel obstruction. Cancers of the rectum are often ulcerating and may present with the passage of bright red blood per rectum or the sensation of incomplete evacuation. On section the cancer appears as a mass of relatively firm, pale tissue replacing the normal bowel wall structures. The mucinous subset has a grey, gelatinous cut surface reflecting the accumulated secretory mucin.

HISTOPATHOLOGY

Type

About 90% of colorectal cancers are adenocarcinomas composed of glandular structures containing variable amounts of mucin. About 10% secrete large amounts of mucin (constituting at least 50% of the tumour volume) and are known as mucinous adenocarcinomas. The signet ring cell carcinoma is a rare form of mucinous carcinoma in which the cells are discohesive and contain abundant intracellular mucin which pushes the nucleus towards the cell membrane (giving a signet ring appearance) (**Figure 6**). Loss of the adhesion molecule E-cadherin is responsible for the discohesion of signet ring cells. Undifferentiated carcinoma is exceedingly uncommon.

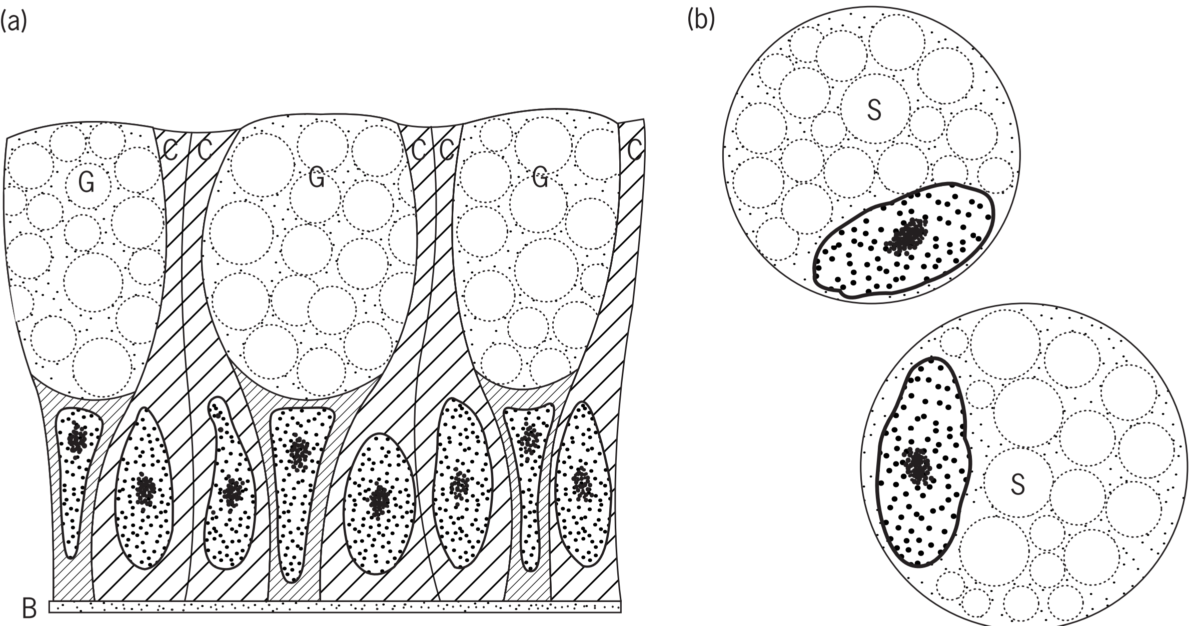


Figure 6 (a) Normal colorectal epithelium and (b) signet ring cell carcinoma. G = goblet cell and C = columnar cell in normal epithelium; S = signet ring cell carcinoma cells; B = basement membrane.

Grading

The grading of adenocarcinoma as well (grade 1), moderately (grade 2) and poorly (grade 3) differentiated is based primarily on the extent to which recognizable glands are formed (**Figure 7**). In well-differentiated adenocarcinoma, the glands are regular and the epithelium resembles adenomatous tubules. In moderately differentiated adenocarcinoma, the glands show complex budding, irregular outpouching or gland-within-gland structures. In poorly differentiated adenocarcinoma, glands are highly irregular or distorted. Cells may be arranged in solid clusters or cords. Undifferentiated carcinomas are graded as 4.

Invasive Margin

About 80% of colorectal cancers have a reasonably well circumscribed invasive margin whereas 20% show widespread dissection of normal structures and often extensive invasion around nerves and within small vessels (**Figure 8**).

Venous Invasion

The presence of tumour within large venous channels increases the risk of metastatic spread to the liver via the portal vein.

Immune Response

Lymphocytes may be conspicuous as part of a band-like arrangement of inflammatory cells at the growing margin

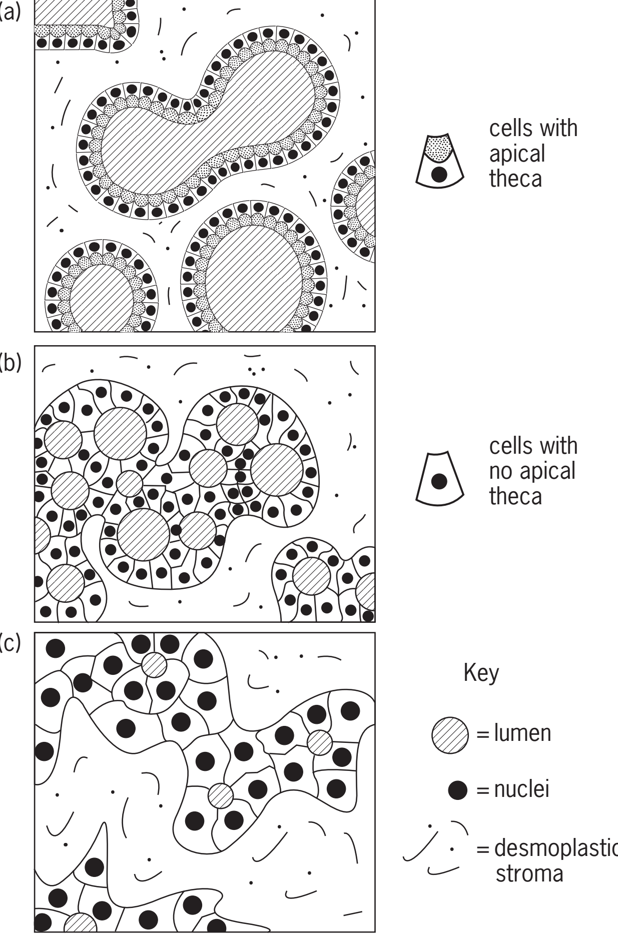


Figure 7 (a) Well, (b) moderate and (c) poor differentiation in adenocarcinoma of the large intestine.

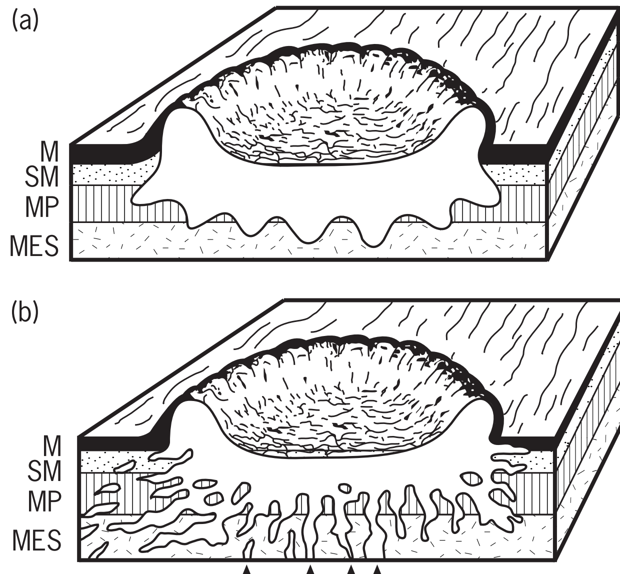


Figure 8 (a) Expanding versus (b) infiltrating adenocarcinoma of the large intestine. M = mucosa; SM = submucosa; MP = muscularis propria; MES = mesentery. Arrows indicate points of surgical transection of cancer.

(B and T cells), as nodular aggregates (B cells with a cuff of T cells) in the surrounding submucosa or serosa, within the tumour stroma (B and T cells) or infiltrating the malignant epithelium (tumour infiltrating T lymphocytes). These patterns often co-exist and are associated with hyperplasia of the regional lymph nodes.

Ultrastructure and Immunohistochemistry

In general these do not provide diagnostically or prognostically useful information but immunohistochemistry combines with the molecular genetic approach (see below) to generate key insights into basic biological mechanisms. The diagnostic use of antibodies to the DNA mismatch repair proteins is described below.

MOLECULAR GENETIC FINDINGS

Colorectal cancer has provided more fundamental insights than any other solid tumour in the search to uncover the genetic basis of cancer. The well-documented adenoma–carcinoma sequence provides a simple model that is accessible to observation and investigation. Additionally, the two major forms of hereditary colorectal cancer, FAP and HNPCC, were directly responsible for the discovery of important cancer genes through classical linkage studies and positional cloning (*APC* and the DNA mismatch repair genes, respectively). (See also the chapter *Inherited Predispositions to Cancer*.)

Genetic Instability

Molecular genetics is a rapidly expanding field. One of its major contributions to date has been the realization that colorectal cancer is not a single disease, but multiple diseases driven by separate molecular pathways. The impact of this realization upon epidemiology, aetiology, prognosis and treatment is only just being appreciated. The pathways are at least in part determined by the nature of the primary fault that drives carcinogenesis: the state of genetic instability. This state is caused by the disruption of checkpoint mechanisms in the cell cycle that normally ensure the maintenance of genomic fidelity during cell division. Genetic instability occurs in at least two major and distinct forms in colorectal cancer: chromosomal instability and DNA instability (Lengauer *et al*., 1998).

Chromosomal Instability

The majority (70%) of colorectal cancers arise through this pathway. An early genetic change is disruption of the Wnt signalling pathway, generally through biallelic inactivation of *APC* which explains the initiation and early growth of adenomas. K-*ras* mutation occurs in a subset and drives further clonal expansion. The key transition from adenoma to carcinoma is driven in many cases by biallelic inactivation of *TP53*. Other tumour suppressor genes on 18q, 1p and 8p are implicated (hence suppressor pathway). Aneuploidy is acquired around the transition from adenoma to carcinoma suggesting that chromosomal instability is the mechanism for overcoming this rate-limiting step in the evolutionary sequence. DNA microsatellite instability is not present, hence these cancers have been termed microsatellite stable (MSS).

DNA Instability and MSI-H Cancers

This mechanism is implicated in HNPCC in which there is a germ-line mutation of a DNA mismatch repair gene (see above). DNA instability also occurs on a sporadic basis in about 15% of colorectal cancers. The usual mechanism is silencing of the DNA mismatch repair gene *hMLH1* through methylation of its promoter region. The biomarker for this subtype is widespread DNA microsatellite instability (MSI) or MSI-high (MSI-H). Other genes are methylated also, but the mechanism underlying the methylator phenotype is unknown. Breakdown of the DNA mismatch repair mechanism results in mutation of genes with short mononucleotide repeats in their coding sequences: *TGFβRII*, *BAX*, *IGF2R*, *caspase 5*, *Tcf-4*, *axin*, *CDX-2* and *BCL-10* (hence mutator phenotype). Involvement of *APC*, K-*ras* and

TP53 is uncommon in sporadic MSI-H cancers. MSI-H cancers show a predilection for the proximal colon and a tendency to occur in elderly women. They are more likely to be mucinous and to express the secretory mucin core proteins MUC2 and MUC5AC (like serrated adenomas). They are also more likely to be poorly differentiated, have a well circumscribed invasive margin and show an enhanced immune reaction including tumour-infiltrating lymphocytes (TIL) (Jass *et al.*, 1998).

Loss of expression of DNA mismatch repair proteins can now be demonstrated with antibodies to hMLH1, hMSH2, hMSH6 and hPMS2. This will greatly assist in the management of HNPCC and identification of MSI-H cancers.

Mild Mutator Pathway

It has been suggested that all colorectal cancers will show minor microsatellite instability if a sufficient number of microsatellite markers is used to test for the mutator phenotype. Nevertheless selected dinucleotide and tetranucleotide markers are sensitive to cancers with low levels of MSI (MSI-L). These cancers differ from MSS cancers in their lower frequency of BCL2 expression, nuclear expression of β-catenin and 5qLOH and higher frequency of K-*ras* mutation, DNA methylation (for example the DNA repair gene *MGMT*) and lymphocytic infiltration. MSI-L cancers are in other respects like MSS cancers, showing high frequencies of APC mutation and allele loss at 17p and 18q (Jass *et al.*, 1999). In view of the multiple differences between MSS, MSI-L and MSI-H cancers, it is prudent to distinguish them carefully as a primary step in the characterization of colorectal cancer. Indeed, the failure to do so has led to oversimplification and confusion that is only now beginning to be resolved.

Pathways in Inflammatory Bowel Disease

Mutation of *TP53* and aneuploidy occur in early stages of neoplastic progression in ulcerative colitis (Potter, 1999).

PROGNOSTIC FACTORS

The outcome for a patient with colorectal cancer will depend upon (1) clinical factors such as age, gender, health status and mode of clinical presentation, (2) the extent of spread of cancer at the time of diagnosis and (3) biological properties of the cancer, both histopathological and molecular. Of these, staging classifications based upon the extent of spread cancer (**Figure 9**) provide the most reproducible and important guide to prognosis.

TNM Classification of Colorectal Cancer

Definitions of T, N and M categories are shown in **Table 2**. The data may be converted into four stages (with the equivalent Dukes pathological stage) as follows:

1. Stage I (Dukes A, Fig. 9(a)): cancer confined to bowel wall (T1 or T2) with no spread to lymph nodes (N0) or distant spread (M0).
2. Stage II (Dukes B, Fig. 9(b)): cancer extends beyond muscle coat (T3 or T4) with no spread to lymph nodes (N0) or distant spread (M0).
3. Stage III (Dukes C, Fig. 9(c)): cancer (any T) spreads to lymph nodes (N1 or N2) with no distant spread (M0).
4. Stage IV (no Dukes equivalent but called D subsequently): cancer (any T and any N) involving distant sites (M1).

Table 2 TNM Classification (American Joint Commission on Cancer/Union Internationale Contra Cancre, 1997)

Primary tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma-*in-situ* or high-grade dysplasia
T1	Tumour invades submucosa
T2	Tumour invades muscularis propria
T3	Tumour invades through the muscularis propria into subserosa or into non-peritonealized pericolic or perirectal tissue
T4	Tumour directly invades other organs or structures and/or perforates visceral peritoneum
Regional lymph nodes (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in 1–3 regional lymph nodes
N2	Metastasis in four or more regional lymph nodes
Distant metastasis (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

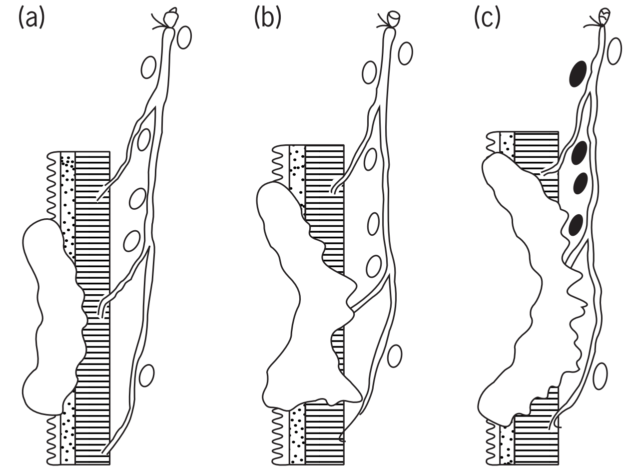


Figure 9 Dukes classification of rectal cancer.

The 5-year survival (non-cancer deaths excluded) is around 100% for stage I, 80% for stage II, 50% for stage III and 10% for stage IV. Distribution of cancers by stage is around 15% stage I, 35% stage II, 40% stage III and 10% stage IV. Prognosis is therefore uncertain (50% 5-year survival) for 40% of patients with stage III disease.

Histopathological Factors

Tumour type, grade, invasive margin, venous invasion and immune response (see above) have all been shown to be of prognostic significance. In multivariate (Cox regression) analyses only invasive margin, venous invasion and immune response have been shown to provide additional (independent) prognostic data beyond that achieved by staging alone. The Jass prognostic classification developed for curative cases of rectal cancer (**Table 3**) utilizes invasive margin and peritumoral lymphocyte infiltration in addition to direct spread and lymph node spread. The Jass system minimizes the number of uncertain prognostic cases (Jass *et al*., 1987). However, histopathological features are subjective and prone to interobserver variation.

Molecular and Genetic Factors

Microsatellite Status

Sporadic MSI-H cancers are associated with a good outcome and an independent prognostic effect has been demonstrated for stage III colorectal cancer. This will explain, at least in part, the adverse effects associated with molecular and genetic changes found in non-MSI-H cancers. These include K-*ras* mutation, 17p and 18q loss and increased expression of COX-2 and VEGF. The classification of colorectal cancer as MSS, MSI-H and MSI-L must precede the exploration of other prognostic markers, the effects of which could well be subsidiary to or explained by the major molecular pathways.

Other Prognostic Markers

The classes of molecule that have been linked to prognosis include enzymes involved in matrix degradation and their inhibitors (cathepsin L, urokinase, tissue-type plasminogen activator, tissue inhibitors of metalloproteinases), gene products involved in apoptosis (BCL2, survivin), cell surface molecules (CD44, ICAM1, galectin 3) and a variety of metabolic enzymes (GLUT1 glucose transporter, manganese–superoxide dismutase, thymidylate synthetase, ornithine decarboxylase).

CLINICAL MANAGEMENT

Initial Consultation and Diagnosis

The management of colorectal cancer begins with a process of consultation. This will vary according to the mode of presentation. The consultation may be motivated by altered bowel habit, present as an acute abdominal emergency due to obstruction or perforation or be precipitated by the diagnosis of cancer in an asymptomatic subject participating in a screening program. Regardless of the mode of presentation, frequent, clear and instructive communication is necessary to ensure the best outcome for the patient.

The main approach to treatment is a surgical one, but this is preceded by preoperative assessment to establish the diagnosis and provide information on the stage of the disease. A mass in the lower rectum need not always be a carcinoma. It could be a variety of benign lesions, a lymphoma or an upward-spreading squamous cell carcinoma of the anal canal (lesions that are not usually treated by rectal excision). Therefore, biopsy and a histopathological diagnosis are mandatory steps in the case of rectal masses and highly desirable for colonic tumours. The methods of investigation to achieve a diagnosis include digital examination, sigmoidoscopy, barium enema and colonoscopy.

Table 3 Jass prognostic classification for curative rectal cancer (UICC/TNM stages I–III): (A) scores for variables and (B) derivation of prognostic groups from total score

A

Variable	Score
Spread beyond rectal wall (muscularis propria)	1
Infiltrative growth pattern	1
No lymphocytic response (peritumoral at deep margin)	1
Lymph node metastasis	1
>4 lymph nodes with metastasis	1

B

Total score per case (0–5)	Prognostic group (I–IV)	5-year survival (%)	Frequency (%)
0–1	I	95	30
2	II	80	30
3	III	50	20
4–5	IV	25	20

Preoperative Staging

Colonic Cancer

Most patients with a colonic cancer do not require preoperative staging; the extent of tumour spread is better evaluated during laparotomy and by pathological examination of the resected specimen. A preoperative computerized tomography (CT) scan will occasionally be useful for identifying the extent of locally advanced disease.

Rectal Cancer

Preoperative staging is important for both planning surgery and establishing the need for preoperative adjuvant chemoradiotherapy. Apart from digital examination, the most important investigative technique is endorectal ultrasound, which provides a high-resolution image of the individual layers of the rectal wall and the destruction of one or more of these layers by tumour.

Distant Metastases

Routine chest X-ray, liver scanning by CT or ultrasound and serum levels of the tumour marker carcinoembryonic antigen (CEA) are used to gauge the presence of distant spread. This may assist in determining prognosis.

Early Colorectal Cancer

There is no strict definition in terms of extent of spread. The term refers to cancer that is sufficiently small and localized to be safely managed by conservative surgery, namely local excision as opposed to the radical approach required for advanced cancer. In practice, most early colorectal cancers have spread no further than the submucosa. An early cancer may be a small focus of malignancy within an adenoma which can be treated by simple polypectomy at the time of colonoscopy. It may also be a small ulcerating cancer in the lower rectum where the alternative to a disc excision of the rectal wall is radical excision of the rectum and anus (abdominoperineal excision) and establishment of a permanent colostomy. Local excision of a cancer is a safe and curative option provided that the excision is complete (there is no cancer at the surgical margin), the cancer is not poorly differentiated and there is no invasion of vascular spaces. Malignant adenomas with unfavourable features may warrant further treatment but this is individualized on the basis of the health, age and wishes of the patient.

Surgery for Colon Cancer

The principle underlying surgical treatment is to resect the segment of bowel bearing the cancer together with the mesentery carrying the blood supply and lymphatic drainage. The feeding arteries are ligated as close as possible to their sites of origin. For a cancer of the right colon, the right colic, ileocolic and right branch of the middle colic vessels are ligated and divided close to their origins and the proximal colon is then resected. For cancers of the left colon, the inferior mesenteric and ascending left colic vessels are ligated and a left hemicolectomy is performed. Intestinal continuity is restored by anastomosis using sutures or staples.

Surgery for Rectal Cancer

In the past this was often treated by an abdominoperineal excision of rectum and anus and the construction of a permanent colostomy. It is now often possible to resect most of the rectum and anastomose the proximal colon to the distal rectum or even to the upper anal canal, a procedure called an anterior resection. When the anstomosis is very low in the pelvis, it is usual to divert the faeces from it temporarily by fashioning a proximal loop ileostomy that is closed after around 12 weeks. It is still necessary to perform an abdominoperineal excision if the cancer is poorly differentiated, a distal margin of clearance of at least 2 cm cannot be achieved or the anal sphincter mechanism is not adequate for continence.

Adjuvant Therapy for Colon Cancer

Approximately 50% of subjects with stage III or Dukes C colon cancer will relapse with distant metastases and die within 5 years of surgery. This is explained by the presence of occult hepatic metastases at the time of surgical treatment. The aim of an adjuvant approach is to destroy these metastases at a time when they are of microscopic dimensions and therefore more amenable to cytotoxic therapy. Randomized controlled trials have to date demonstrated significant improvements in survival with the combination of 5-fluorouracil and leucovorin (folinic acid) that have not been bettered by any alternative regimen. A 6-month course is recommended.

Adjuvant Therapy for Rectal Cancer

Two major adverse outcomes are observed in rectal cancer: distant metastases (as for colon cancer) and local recurrence of the disease within the pelvic cavity. Recurrent pelvic cancer is incurable and will spread within the pelvis to cause pain, bleeding and urinary tract obstruction. Factors influencing local recurrence are the extent of spread of cancer beyond the rectal wall and the competence of the surgeon. Unacceptably high levels of variation in outcome have been demonstrated between surgeons (McArdle and Hole, 1991). Sharp dissection of the mesorectum ensuring that its thin fascial (connective tissue) covering is not breached is advocated as a measure to reduce the risk of incomplete

removal of cancer. However, local recurrence will continue to be a problem for the more locally advanced cancers regardless of surgical skills. For this reason, radiotherapy is recommended for locally advanced rectal cancer and this has been shown to reduce the incidence of local recurrence in randomized controlled trials. It is therefore usual to offer combined radiotherapy and chemotherapy (e.g. 5-fluorouracil) to patients with stage II and III rectal cancer.

Radiotherapy may damage normal tissues, causing long-term complications. The risks of damaging the small intestine are reduced if radiotherapy is given preoperatively. Preoperative radiotherapy (combined with chemotherapy) can be offered selectively on the basis of preoperative staging. The advantages of this approach are that the cancer may be downstaged (partially eradicated) and sometimes completely eradicated. Preoperative radiotherapy may also render surgery less technically difficult and may even convert an apparently incurable cancer to a curable one. Pathological staging of the resected specimen will be less reliable owing to destruction of both tumour and lymphoid tissue (Wheeler *et al*., 1999).

Recurrent and Distant Disease

This is usually incurable although patients with only a few hepatic metastases have been cured by excising the deposits of tumour with a clear margin. Local recurrence is also salvageable on occasion, particularly if it is at the site of the anastomosis. Local recurrence is otherwise managed by palliation with radiotherapy and/or chemotherapy as well as adequate pain relief.

The Future

In addition to the development of more effective cytotoxic compounds (e.g. Oxaliplatin and CPT-11), immunotherapy and gene therapy offer new hope for the future. These approaches depend on an understanding of cancer biology and must overcome the problems posed by tumour heterogeneity and drug resistance. Until we learn how to prevent the disease altogether, surgery will continue as the main treatment option for advanced colorectal cancer. Despite this, modern management of colorectal cancer is increasingly viewed as a team exercise involving oncologists, radiotherapists, radiologists, endoscopists, geneticists, pathologists and stoma therapists as well as surgeons. The informed participation of the patient is also integral to successful management.

REFERENCES

Atkin, W. S., *et al*. (1992). Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. *New England Journal of Medicine*, **326**, 658–662.

Atkin, W. S., *et al*. (1998). Uptake, yield of neoplasia, and adverse effects of flexible sigmoidoscopy screening. *Gut*, **42**, 560–565.

Haenszel, W. (1961). Cancer mortality among the foreign born in the United States. *Journal of the Natural Cancer Institute*, **26**, 37–132.

Hardcastle, J. D., *et al*. (1996). Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet*, **348**, 1472–1477.

Iino, H., *et al*. (1999). DNA microsatellite instability in hyperplastic polyps, serrated adenomas and mixed polyps: a mild mutator pathway for colorectal cancer? *Journal of Clinical Pathology*, **52**, 5–9.

Jass, J. R. and Sobin, L. H. (1989). *WHO International Histological Classification of Intestinal Tumours*. (Springer, Berlin).

Jass, J. R., *et al*. (1987). A new prognostic classification of rectal cancer. *Lancet*, **i**, 1303–1306.

Jass, J. R., *et al*. (1998). Morphology of sporadic colorectal cancer with DNA replication errors. *Gut*, **42**, 673–679.

Jass, J. R., *et al*. (1999). Characterisation of a subtype of colorectal cancer combining features of the suppressor and mild mutator pathways. *Journal of Clinical Pathology*, **52**, 455–460.

Keller, J. J., *et al*. (1999). Rectal epithelial apoptosis in familial adenomatous polyposis patients treated with sulindac. *Gut*, **45**, 822–828.

Kronborg, O., *et al*. (1996). Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet*, **348**, 1467–1471.

Lengauer, C., *et al*. (1998). Genetic instabilities in human cancers. *Nature*, **396**, 643–649.

Lynch, H. T., *et al*. (1996). Overview of natural history, pathology, molecular genetics and management of HNPCC (Lynch syndrome). *International Journal of Cancer*, **69**, 38–43.

McArdle, C. S. and Hole, D. (1991). Impact of variability among surgeons on postoperative morbidity and mortality and ultimate survival. *British Medical Journal*, **302**, 1501–1505.

Muto, T., *et al*. (1975). The evolution of cancer of the rectum. *Cancer*, **36**, 2251–2276.

Nakamura, S. and Kino, I. (1984). Morphogenesis of minute adenomas in familial polyposis coli. *Journal of the Natural Cancer Institute*, **73**, 41–49.

Potter, J. D. (1999). Colorectal cancer: molecules and populations. *Journal of the Natural Cancer Institute*, **91**, 916–932.

Trock, B., *et al*. (1990). Dietary fiber, vegetables, and colon cancer: critical review and analyses of the epidemiologic evidence. *Journal of the Natural Cancer Institute*, **82**, 650–661.

Wasan, H. S., *et al*. (1998). APC in the regulation of intestinal crypt fission. *Journal of Pathology*, **185**, 246–255.

Wheeler, J. M. D., *et al*. (1999). Preoperative radiotherapy for rectal cancer: implications for surgeons, pathologists and radiologists. *British Journal of Surgery*, **86**, 1108–1120.

FURTHER READING

American Joint Committee on Cancer (1997). *AJCC Cancer Staging Manual*, 5th edn. (Lippincott-Raven, Philadelphia).

Baba, S. (1996). *New Strategies for Treatment of Hereditary Colorectal Cancer*. (Churchill Livingstone, Tokyo).

Hermanek, P. and Sobin, L. H. (1995). Colorectal cancer. In: Hermanek, P., *et al*. (eds), *Prognostic Factors in Cancer*. 64–79 (Springer, Berlin).

Isaacson, P. G. and Norton, A. J. (1994). *Extranodal Lymphomas*. (Churchill Livingstone, Edinburgh).

Kune, G. (1996). *Causes and Control of Colorectal Cancer: A Model for Cancer Prevention*. (Kluwer, Boston).

Medical Research Council Rectal Cancer Working Party (1996). Randomised trial of surgery alone versus radiotherapy followed by surgery for potentially operable locally advanced rectal cancer. *Lancet*, **348**, 1605–1610.

Medical Research Council Rectal Cancer Working Party (1996). Randomised trial of surgery alone versus surgery followed by radiotherapy for mobile cancer of the rectum. *Lancet*, **348**, 1610–1614.

Phillips, R. K. S., *et al*. (1994). *Familial Adenomatous Polyposis and Other Polyposis Syndromes*. (Edward Arnold, London).

Sircar, K., *et al*. (1999). Interstitial cells of Cajal as precursors of gastrointestinal stromal tumors. *American Journal of Surgery and Pathology*, **23**, 377–389.

Vogelstein, B., *et al*. (1988). Genetic alterations in colorectal tumor development. *New England of Journal of Medicine*, **319**, 525–532.

Website

http://www.health.gov.au/nhmrc/

Chapter 39

Liver, Gall Bladder and Extrahepatic Bile Ducts

Peter Paul Anthony

Royal Devon and Exeter Hospitals and University of Exeter, Exeter, UK

CONTENTS

- Normal Development and Structure
- Pathology: Tumours of the Liver
- Hepatocellular (Liver Cell) Carcinoma
- Hepatoblastoma
- Cholangiocarcinoma (Intrahepatic and Hilar Bile Duct Carcinoma)
- Biliary Cystadenoma and Cystadenocarcinoma
- Mixed Hepatocellular Carcinoma and Cholangiocarcinoma
- Metastatic Tumours
- Sarcomas of the Liver
- Benign Tumours and Tumour-like Lesions
- Tumours of the Gall Bladder
- Tumours of the Extrahepatic Bile Ducts

NORMAL DEVELOPMENT AND STRUCTURE

Elements of the liver first appear at the beginning of the fourth week of gestation and its development is largely complete by the twelfth week. The parenchymal component (hepatocytes, biliary system and gall bladder) derive from the liver bud or hepatic diverticulum of the endodermal intestinal canal and the connective tissue framework and vessels from the mesenchyme of the septum transversum. The liver is also a major site for haematopoiesis throughout intrauterine life. Hepatocytes produce α-fetoprotein during the first trimester, after which they gradually switch to albumin. These events are, to a varying extent, reproduced in certain tumours: the morphological changes in hepatoblastoma, haematopoiesis in angiosarcoma and, most important, a switch back to α-fetoprotein in hepatocellular carcinoma.

The liver lies almost completely under the protection of the rib cage on the right and is firmly attached to the diaphragm above and loosely to the stomach and duodenum below by the lesser omentum. It is traditionally divided into right, left, caudate and quadrate lobes, which are visually recognizable, but, if one considers vascular supply and biliary drainage, it consists of nine autonomous segments, the limits of which guide the surgeon when carrying out partial hepatectomy or resection. The liver receives a dual blood supply: venous blood from the portal vein and arterial blood from the hepatic artery. Some tumours are predominantly supplied via the portal vein (e.g. angiosarcoma) and others via the hepatic artery (e.g. hepatocellular carcinoma). Blood flows out of the liver via the hepatic veins into the inferior vena cava, close to the heart. Branches of the intrahepatic bile duct system join up to form the main hepatic duct which, when it receives the cystic duct of the gall bladder, becomes the common bile duct and this drains bile into the second part of the duodenum.

Concepts of the microscopic structure of the liver have progressed from the purely anatomical, the portal 'lobule' of Mall described in 1906, to the functional unit or 'acinus' of Rappaport which was defined in 1954; the latter makes more sense. **Figure 1** shows a simplified view of the arrangement of the various components. The centre of the acinus is the portal tract which carries the smallest branches of the portal vein and of the hepatic artery. These flow into specialized vascular channels called sinusoids which run between plates of hepatocytes and are lined by endothelium. They contain phagocytic Kupffer cells. The sinusoids are separated from the liver cells themselves by the space of Disse which, in turn, contains perisinusoidal or Ito cells and scanty Type I and III collagen which constitutes the connective tissue (reticulin) framework of the liver. The main function of the perisinusoidal cells is the production of extracellular matrix proteins. The sinusoids drain into the terminal hepatic venules which unite to form the hepatic veins. The third structure in the portal tract is the first, smallest branch of the intrahepatic bile duct

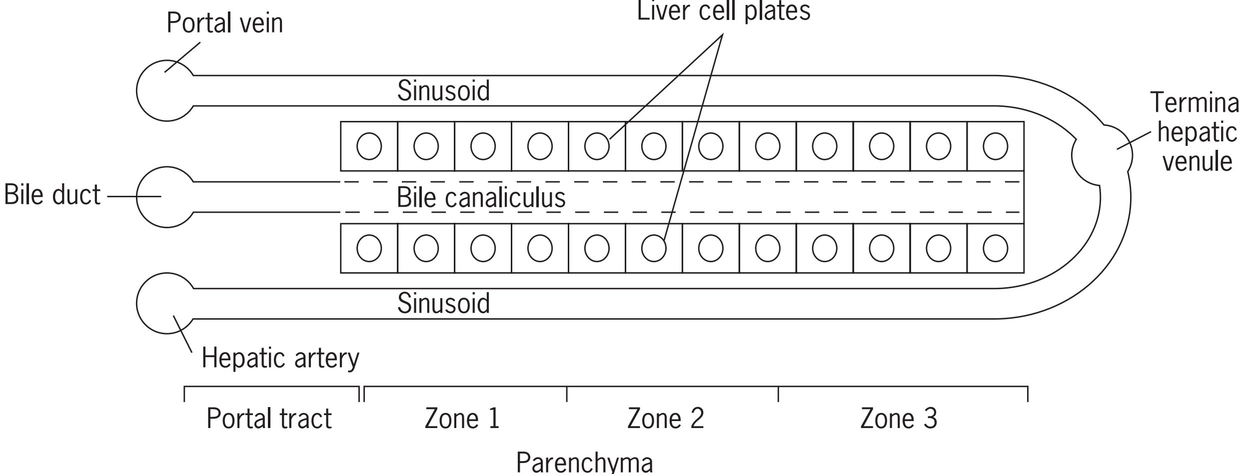


Figure 1 Microanatomy of the liver, showing portal tract with bile duct, hepatic artery and portal vein, liver cell plates with intervening sinusoids and the terminal hepatic vein.

Note: A two-dimensional diagram does not allow for the display of three-dimensional arrangements, e.g. both portal veins and hepatic arteries drain into the same sinusoids.

system which drains bile from the intercellular canaliculi between hepatocytes towards segmental bile ducts and, eventually, the common bile duct. Hepatocytes are roughly four-sided and are arranged in complex, continuous plates, being joined on two sides by other hepatocytes; the other two sides face the sinusoids or form the bile canaliculi.

Our understanding of the structure and function of the microscopic organisation of the liver is still evolving and there are several modifications and refinements. For example, the cells of the plates can be divided into Zones 1, 2 and 3 from those closest to the portal tracts (Zone 1) to those nearest to the terminal hepatic venules (Zone 3). Hepatocytes in Zone 1 are the youngest and those in Zone 3 the oldest: they proceed in this direction from birth to death by apoptosis in approximately 200 days. This is the concept of the 'streaming liver' created by Zajicek in 1985. As hepatocytes move they also display differences in function, suffer a fall in the supply of oxygen and nutrients and become more vulnerable. The site of origin of both hepatocytes and bile duct cells lies between the portal tract and liver parenchyma, the domain of stem cells.

Hepatocytes are the most complex cells in the body. They have a centrally placed nucleus and a cytoplasm rich in organelles and are capable of performing many storage, metabolic, synthetic, detoxifying and excretory functions. Bile ducts not only conduct bile but also modify it to its final effective form. The endothelial cells of the sinusoids are fenestrated, which allows the transport of nutrients to the hepatocytes. Kupffer cells and liver-specific lymphocytes are part of the mononuclear phagocytic system of the body and perisinusoidal Ito cells are capable of fibrogenesis.

This account is only a brief outline of the development, structure and function of the liver to allow the reader to relate specific tumours at this site to their cell of origin and to normal structures which they recapitulate albeit in an aberrant fashion. For a comprehensive account, refer to MacSween and Scothorne (1994).

PATHOLOGY: TUMOURS OF THE LIVER

The term 'hepatoma' is often used loosely to describe any primary tumour in the liver but it has no precise meaning. Many different types arise which are designated by their cell of origin (Anthony, 1994). The commonest is liver cell or hepatocellular carcinoma, followed by bile duct or cholangiocarcinoma and various sarcomas, of which angiosarcoma is the most important. Benign tumours are uncommon. **Table 1** shows a detailed classification of liver tumours but only those that are common and/or are of interest will be discussed and the rest referred to only briefly.

Our knowledge of the world-wide distribution of tumours is based on data derived from cancer registries in many countries. These have been published in successive volumes of the series *Cancer Incidence in Five Continents*, the latest of which appeared in 1997 (Parkin *et al*., 1997). The variability in incidence of liver tumours is almost entirely due to the large numbers of hepatocellular carcinoma in tropical Africa and South-East Asia. Cholangiocarcinoma is less common and occurs with much the same frequency everywhere except in South-East Asia. Other tumour types are rare. Overall, malignant liver tumours rank eighth in the list of all cancers, sixth in men and eleventh in women. Hepatocellular carcinoma accounts for most cases and kills around one million people in the world every year.

The outlook for hepatocellular carcinoma, cholangiocarcinoma and angiosarcoma is poor and nearly all patients die unless the tumour is detected early. Results

Table 1 Abbreviated WHO Classification of primary tumours of the liver

Type	Benign	Malignant
Epithelial tumours	Hepatocellular adenoma Bile duct adenoma Bile duct cystadenoma Biliary papillomatosis	Hepatocellular carcinoma Cholangiocarcinoma Bile duct cystadenocarcinoma Hepatoblastoma
Non-epithelial tumours	Haemangioma Angiomyolipoma Other benign tumours	Haemangiosarcoma Malignant epithelioid haemangioendothelioma Embryonal sarcoma Rhabdomyosarcoma Other sarcomas, lymphoma
Tumour-like lesions	Cysts Focal nodular hyperplasia Mesenchymal hamartoma Peliosis Inflammatory myofibroblastic tumour	

are improving with hepatoblastoma and the childhood sarcomas.

HEPATOCELLULAR (LIVER CELL) CARCINOMA

Hepatocellular carcinoma is defined by the World Health Organisation as a malignant tumour composed of cells resembling hepatocytes but abnormal in appearance; a plate-like organisation around sinusoids is common and is nearly always present somewhere in the tumour (Ishak *et al*., 1994).

Epidemiology

The remarkable geographical variability in the frequency of this tumour has attracted the attention of epidemiologists since the middle of the twentieth century. The data have been regularly summarized and updated over the years (Simonetti *et al*., 1991; Okuda, 1993; Anthony, 1994; Akriviadis *et al*., 1998; Bosch *et al*., 1999). Yearly incidence rates and mortality figures are almost the same as patients, other than those whose tumour is detected at an early stage by screening, die within a few weeks or at most months. In broad terms, countries may be divided into three groups: those with low, intermediate and high incidence rates. The highest frequencies are seen in South-East Asia and tropical Africa where the tumour is the commonest or next commonest of all cancers. The lowest rates are found in Western countries, South America and the Indian subcontinent. Intermediate rates prevail in Japan, the Middle East and the Mediterranean area. The possible role of racial and genetic factors has been examined and discounted in favour of environmental factors, notably chronic infections with the hepatitis B and C viruses and exposure to aflatoxin. The incidence of the tumour appears to be rising in many areas, e.g. USA, Japan and Southern Europe.

Males predominate over females in a ratio of 2–4 to 1; generally, the higher the incidence, the higher is the rate between the genders. The mean age in high-incidence areas is in early to middle adulthood and Africans develop the tumour earlier than Asians. This is largely due to the acquisition of hepatitis B virus infection at or near birth. Patients in low- and intermediate-incidence areas are in late adulthood or old age. Cirrhosis of the liver is often associated with the tumour and is considered to be the greatest risk factor.

Whites, wherever they live, have a low incidence, even in Africa and South-East Asia, where hepatocellular carcinoma is rampant in the native population. They appear to be protected by maintaining the life style of their home countries. The same is not true of Indians who have settled in Singapore or Hong Kong since the end of the nineteenth century: their incidence rate is roughly double that in their home country. Chinese are at high risk wherever they live, e.g. in Europe or North America, but this decreases after the second generation as they adopt the environment of their new countries. The black population of the USA, West Indies and Brazil shows an incidence that is only marginally above that of whites amongst whom they have lived since the beginning of the slave trade in the seventeenth century. All of these phenomena can be explained by the prevalence rates of chronic infection with hepatitis B and, to a lesser extent, hepatitis C viruses. Small variations in tumour incidence have been observed in racially homogeneous countries such as Greece, Spain and Switzerland, which are due to differences in rates of alcoholic cirrhosis, smoking and exposure to chemicals. It is interesting to note that, although over 100 compounds, both man-made and naturally occurring, are known to be carcinogenic to rodents,

few of these seem to affect humans, with the exception of aflatoxin, which is an important risk factor in Kenya and Swaziland.

Aetiology

The list of aetiological agents is long and is shown in **Table 2**, but some are more important than others. This is indicated in terms of the prevalence of the agent and the magnitude of the risk attached to it. The effect of most of them is to proceed from the normal state of the liver to the cancerous through cirrhosis, but the likelihood of this varies (Simonetti *et al*., 1991; Okuda, 1993; Akriviadis *et al*., 1998). Chronic hepatitis B and C account for four-fifths of cases of hepatocellular carcinoma followed by aflatoxin, chronic alcoholism, smoking and a number of other aetiologies in a minority (Anthony, 1994; Bosch *et al*., 1999). However, a few metabolic disorders such as haemochromatosis and tyrosinaemia carry a surprisingly high risk. Also, the presence of more than one agent greatly increases the likelihood of malignancy.

Hepatitis B Virus (HBV)

HBV is one of a group of viruses known as Hepadnaviruses which affect humans (HBV) and certain animals such as the woodchuck (WHV), ground squirrel (GSHV) and duck (DHV). All of these cause liver disease: acute and chronic hepatitis, cirrhosis and hepatocellular carcinoma, in their respective hosts. In humans, transmission is mainly via blood or blood products, contaminated instruments, male homosexual contact and, most important, from mother to infant at birth. HBV is a partly double-stranded DNA virus possessing a genome consisting of four open reading frames which encode for four different proteins, surface (HBsAg) and core (HBcAg) antigens, DNA polymerase and the X protein. The 'e' antigen (HBeAg) is closely associated with HBcAg. Replication takes place through reverse transcriptase and the virus is capable of integration into the host cell's genome. Most of the antigens and antibodies directed against them are demonstrable in the serum at one time or another during acute and chronic infections and the pattern of their presence is predictive of course, stage of disease and outcome. Levels of these 'marker' proteins decline in time as replication ceases and the virus becomes undetectable, but it remains in an integrated form in the nuclei of liver cells. The risk of malignant transformation relates both to the infective, replicative phase when continuing liver cell damage leads to hepatitis and cirrhosis, and to the non-replicative, integrated phase when carcinogenic events occur (Idilman *et al*., 1998; Chen and Chen, 1999; Schafer and Sorrell, 1999).

Table 2 Aetiology of hepatocellular carcinoma (male : female=2–4 : 1)

Agent	Prevalence	Magnitude of risk
HBV	High	High
HCV	High	High
Alcohol	High	Low
Chemicals and drugs	High	Low
Aflatoxin	Low	Moderate
Membranous obstruction of IVC	Low	Moderate
Haemochromatosis	Moderate	Moderate
AAT deficiency	Low	Low
Tyrosinaemia	Low	High
Glycogen storage disease	Low	Low
Porphyria	Low	Low

Chronic HBV infection is common in tropical Africa, South-East Asia and Oceania; it is uncommon in Western Europe, North America and Australia and an intermediate prevalence is seen in the Mediterranean countries, the Middle East and India. Indeed, maps can be constructed to show almost identical frequencies of HBV carriage rates and incidence of hepatocellular carcinoma, as this virus accounts for up to two-thirds of all cases of the tumour and more in some areas. It is generally accepted that a particularly high risk of chronic infection, with the subsequent development of cirrhosis and hepatocellular carcinoma, is associated with infection at or near birth or in childhood, which happens in tropical Africa and South-East Asia. Elsewhere, those most commonly affected are adults in high-risk groups, e.g. intravenous drug addicts, homosexuals, prostitutes and inmates of prisons or mental institutions. Individuals who received transfusions of blood and blood products prior to the institution of effective screening of donors in the 1980s are also at risk. Most of those who have chronic infection remain symptomless for many years and are known as carriers.

The strength of the association of HBV and hepatocellular carcinoma, however, rests on multiple lines of evidence which are listed in **Table 3**. Numerous case-control studies were carried out in the 1970s and 1980s to show a much higher prevalence of HBV markers in the blood of patients with hepatocellular carcinoma than in controls, whether these were healthy individuals or hospital patients with other types of diseases including cancer. Infection precedes the development of the tumour by two to three decades on average and long-term follow-up studies established the measure of the risk. The best known of these was carried out on government employees in Taiwan, a high-incidence area for both HBV infection and hepatocellular carcinoma. They were chosen because all had life insurance policies provided by the State and an accurate record of causes of illness and death was readily available. The relative risk of developing the tumour, by

Table 3 Hepatitis viruses and hepatocellular carcinoma

	HBV	HCV
Infection is associated with tumour (case/control)	Yes	Yes
Infection precedes tumour (follow-up)	Yes	Yes
Tumour DNA contains viral genomic element (integration)	Yes	No
Tumour in culture produces viral component	Yes	No
Virus transforms cells in culture	No	No
Virus produces tumours in transgenic mice	Yes	Yes
Virus induces tumours in animals	Yes	No
Eradication of virus leads to decrease of tumour	Yes	Not known

those who tested positive for HBV markers at the outset over those who did not, turned out to be nearly 100-fold 10 years later (Beasley, 1988), a much higher figure than the relative risk of lung cancer in smokers over non-smokers. Replicative HBV is not detectable in hepatocellular carcinoma tissue or cell lines but antigenic components such as HBsAg rarely are. The virus is not directly oncogenic in the sense that it does not transform cells in culture. However, it does produce hepatocellular carcinomas when its genome is incorporated in the germ line of mice in transgenic experiments. HBV-like hepadnaviruses also produce the tumour in their respective hosts, particularly the woodchuck. The most important event in hepatic carcinogenesis by HBV is its integration in the DNA of liver cell nuclei. No consistent site exists and it appears to be random and multiple. However, insertion of HBV in the genome leads to its destabilization and chromosomal abnormalities are common, a process known as insertional mutagenesis. In addition, two transactivating proteins have been identified; one is encoded for in the pre-S2 region of the gene coding for HBsAg and the other is the HBX protein. Of the two, the latter is of greater interest (Feitelson and Duan, 1997; Idilman *et al*., 1998; Chen and Chen, 1999). Transactivation in this context means the ability to modify the action of host genes at a distance from the integration site. Most hepatocellular carcinomas contain oncogenes (e.g. *ras*, *myc*, *fos*) and/or anti-oncogenes or tumour suppressor genes (e.g. *p53*, *Rb*) which have been made by HBX to be overexpressed or inactivated. These events, in turn, lead to changes in growth factors, cell cycle regulators and DNA repair which are relevant to carcinogenesis but no coherent pathway has yet been elucidated. Mass vaccination of infants against HBV in high-risk areas such as Taiwan and the Gambia began in the mid-1980s and has proved to be effective in preventing chronic infection, with a fall in the number of childhood hepatocellular carcinomas to a quarter of the previous level. (See the chapter *Human DNA Tumour Viruses*.)

Hepatitis C Virus (HCV)

HCV is a single-stranded RNA virus which shows marked genetic heterogeneity and at least six major subtypes are known, of which Ib is thought to be the most likely to lead to chronic liver disease. Whilst it is associated with only about one-sixth of all cases of hepatocellular carcinoma world-wide, this proportion is higher and rising in some areas, notably Japan and, to a lesser extent, Spain, Italy and the Middle East (Idilman *et al*., 1998; Bosch *et al*., 1999; Colombo, 1999). The infection is usually acquired in adult life via transfusion of blood and blood products or by the use of contaminated instruments and syringes by intravenous drug abusers. Perinatal and sexual transmission are unimportant. The onset of malignancy is preceded by cirrhosis in 90% of cases. The course is long, 20–40 years from infection to tumour, and patients are affected in late middle to old age. **Table 3** shows that HCV is definitely associated with hepatocellular carcinoma in case-control and follow-up studies, but the pathogenesis is unknown. As it is an RNA virus, which does not possess a reverse transcriptase enzyme, it cannot integrate into the nuclei of liver cells. Instead, the risk of malignancy is associated with continued viral replication, liver cell death and stimulus to proliferation which lead to cirrhosis. However, hepatocellular carcinoma develops in a minority of cases without preceding cirrhosis and, therefore, a direct oncogenic effect cannot be ruled out. Accumulation of HCV core protein is thought to be the most likely mechanism. The co-existence of chronic HBV infection and alcoholism greatly increases the risk of malignancy. Attack rates of HCV infection have been reduced by screening of blood donors and the widespread use of disposable syringes but no vaccine is available and, because of the antigenic versatility of the virus, it may take many years yet to develop.

Other Hepatotropic Viruses

Hepatitis A, E, G and TT viruses are not associated with hepatocellular carcinoma, nor are herpes and other viruses that may rarely cause hepatitis. The delta agent (hepatitis D virus) can only co-infect with HBV and has no independent role.

Alcohol

A history of chronic alcohol abuse is frequently obtained in patients with cirrhosis and hepatocellular carcinoma, particularly in Europe, North America and Japan. Alcohol is not, by itself, a carcinogenic agent and it only exerts an effect through inducing cirrhosis. The risk is increased by concomitant HBV or HCV infection and, to a lesser extent, by smoking (Simonetti *et al*., 1991; Okuda, 1993; Bosch *et al*., 1999).

Chemicals and Drugs

Hundreds of chemicals – some purely experimental, others industrial – can produce liver tumours in rodents but epidemiological studies have shown that few, if any, are relevant to humans. Those that have caused most concern are nitrites, hydrocarbons, solvents, organochlorine pesticides and polychlorinated biphenyls (International Agency for Research on Cancer, 1972–1999). Tumours that develop after the administration of high doses of chemicals in short-term animal experiments may represent false-positive results for carcinogenicity because they produce a rapid mitogenic response. Damage to nuclear DNA is not repaired in the circumstances and it is 'fixed' by mitosis, leading to permanent mutation. The same chemicals in low doses do not stimulate mitosis in humans and therefore no harm results. Studies of occupational hazards have also produced inconclusive results. Therapeutic drugs are introduced after stringent testing procedures and few have proved to be carcinogenic and only in a small number of cases. They include oral contraceptives and anabolic–androgenic steroids, danazol and cytoproterone acetate. The radiological contrast agent thorium dioxide (Thorotrast) and the industrial contaminant vinyl chloride monomer (VCM) can be added to the list.

Aflatoxins

A hot, humid climate and prolonged storage encourage the growth of moulds on foodstuffs, notably of *Aspergillus fumigatus*, which produces toxic metabolites known as aflatoxins. Dietary exposure over long periods of time has led to an increased incidence of hepatocellular carcinoma in parts of sub-Saharan Africa and Southern China. The risk is increased by concomitant infection with HBV but aflatoxins are, by themselves, carcinogenic. A specific G to T mutation at the third base of codon 249 of the *p53* gene is a hallmark of exposure and it is seen in a high proportion of cases of hepatocellular carcinoma in these areas (Simonetti *et al.*, 1991; Bosch *et al.*, 1999; Wogan, 1999). (See the chapter *Mechanisms of Chemical Carcinogenesis*.)

Membraneous Obstruction of the Inferior Vena Cava

Reports have appeared from South Africa, Japan and Taiwan of an association between abnormalities of the inferior vena cava and hepatocellular carcinoma. The incidence of the lesion, which may be either congenital or acquired, is difficult to assess as it is seldom sought and is easily overlooked (Anthony, 1994).

Congenital Abnormalities and Metabolic Disorders

A large number of inherited conditions have been recognized as possible causes of hepatocellular carcinoma. These include Alagille's syndrome, ataxia–telangiectasia, familial polyposis of the colon, hereditary haemorrhagic telangiectasia, familial cholestatic cirrhosis, neonatal hepatitis/biliary atresia, neurofibromatosis and Soto syndrome. All of these are rare and some cases may have been pure chance associations (Anthony, 1994).

Inborn errors of metabolism are not all rare and some carry a surprisingly high risk of hepatocellular carcinoma (European Association for the Study of the Liver, 1999). It is interesting that this tumour is the only malignancy that complicates these disorders regularly. Genetic haemochromatosis is an autosomal recessive iron storage disorder associated with two mutations in the *HFE* gene: *C282Y* and *H63D*. The frequency in the general population is 0.5–1.0% in Northern Europe but lower elsewhere. The accumulation of iron leads to cirrhosis which then may be complicated by hepatocellular carcinoma, especially in males. The relative risk is about 100 over that of normal individuals. α-1-Antitrypsin deficiency is associated with neonatal jaundice and cirrhosis in early childhood and with emphysema and cirrhosis in adult life. Inheritance is under the control of *Pi* (protease inhibitor) genes of which the Z variant is the most important. Male adults with cirrhosis are at an increased risk of developing hepatocellular carcinoma but the magnitude of this is debated. Tyrosinaemia carries an almost 100% risk of hepatocellular carcinoma by the age of 10 years if patients survive that long. It is now recommended that hepatectomy and liver transplantation be carried out by 2 years of age. Glycogen storage disease and porphyria syndromes are rare and the risk of malignancy is low.

Precancerous Changes in the Liver, Screening and Prevention

Cirrhosis itself is a precancerous condition but the magnitude of the risk varies with aetiology, i.e. it is high with HBV and HCV but low with alcohol. Sustained proliferation of liver cells consequent upon chronic hepatitis B and C or the daily ingestion of a hepatoxic agent such as alcohol over many years seems to be the most likely mechanism of carcinogenesis as it predisposes the cell to accumulate DNA abnormalities. Proliferation prevents the repair of any damage to DNA and it is then 'fixed' and transmitted to the progeny. The 'right' combination of sites of damage that results in neoplastic transformation is a rare event, hence it takes time for the tumour to develop.

In addition, several microscopic and architectural abnormalities have been described and assigned a precancerous role (Anthony, 1994). Cellular changes such as large-cell dysplasia (meaning liver cells with enlarged, hyperchromatic, polyploid nuclei) have been held to be associated with an increased risk of malignant change for many years, especially in studies from Africa, Italy and France. Structurally abnormal and often large liver nodules

in cirrhosis have been termed macroregenerative nodule, adenomatous hyperplasia and dysplastic nodule (**Figure 2**). These, especially when associated with cytological atypia (nuclear enlargement, prominent nucleoli, increased nucleo-cytoplasmic ratio), acinar, gland-like structures, nodules within nodules, and growth into areas of fibrosis, have been held to be directly precancerous lesions. Small-cell dysplasia has also been described in these nodules. However, dysplastic nodules may be difficult to distinguish from early, small hepatocellular carcinomas.

It must be pointed out that hepatocellular carcinoma does not always develop from cirrhosis with or without large nodules and dysplastic changes and, therefore, these abnormalities do not constitute obligate and exclusive pathways of liver carcinogenesis in humans. They are clinically useful, however, as they indicate an increased risk and the need for close supervision, which allows early intervention before overt development of tumours in cirrhotic patients (Okuda, 1993; Schafer and Sorrell, 1999; Hirohashi *et al*., 2000).

The best preventive measure for hepatocellular carcinoma is elimination of the causative agent (primary prevention). Vaccination against HBV is highly effective in reducing chronic infection with this virus and has resulted in a reduction of tumour incidence in countries with nation-wide programmes such as Taiwan. A vaccine against HCV is not yet available but the virus has now been eliminated from blood used for transfusion by screening of donors. It is also possible to reduce exposure to aflatoxin by improving conditions of food storage. However, hundreds of millions of people remain for whom such measures are too late.

Secondary prevention means the identification of individuals considered to be at highest risk. In practice, this means those with cirrhosis due to chronic infection with HBV and HCV and large nodular lesions with or without dysplastic changes. The most effective means are regular ultrasound examination and estimation of serum levels of α-fetoprotein, the role of repeat liver biopsies being more controversial. Individuals with early, small tumours can then successfully be treated by surgery: resection or transplantation.

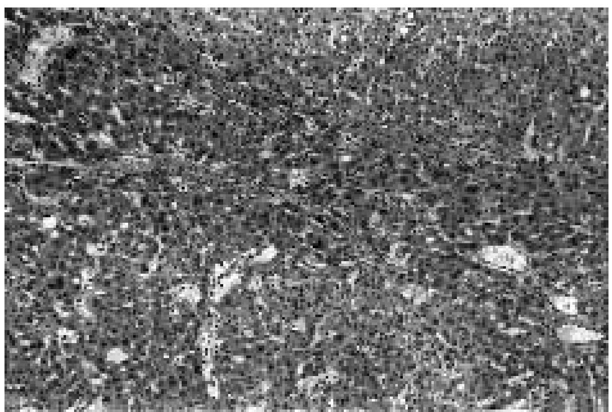


Figure 2 The interior of a cirrhotic 'dysplastic' nodule. Liver cell plates vary in thickness, their arrangement is disorganized and those in the middle of the field show nuclear enlargement.

Macroscopic and Microscopic Pathology

These have been described in detail over the years (Craig *et al*., 1989; Anthony, 1994; Ishak *et al*., 1994; Kojiro *et al*., 2000).

The Macroscopic Pathology of Hepatocellular Carcinoma

Most early tumours detected by screening with ultrasound and estimation of serum α-fetoprotein levels are small, usually <2 cm, solitary, and histologically well differentiated. Tumours that present with symptoms are >2–3 cm and usually much larger and are increasingly less well differentiated as they grow. A classification devised in the early twentieth century distinguishes three forms. The most common is multinodular in which more or less rounded, sharply demarcated, yellow–green nodules of tumour are scattered throughout the liver which is usually cirrhotic. Massive tumours are large, solitary, necrotic masses, sometimes with small, satellite nodules in the vicinity, often in an otherwise normal liver. Partial or complete encapsulation may be seen. Diffuse tumours replace the entire liver without the formation of a discrete mass. At autopsy, these patterns frequently co-exist (**Figure 3**). Rarely, the tumour is pedunculated or even outside the liver: these arise from accessory lobes or from ectopic liver tissue.

Hepatocellular carcinomas have an arterial blood supply but both portal and hepatic veins proliferate alongside, resulting in arterio-venous communications. Consequently, intra- and extrahepatic spread can take place in all directions, commonly to all parts of the liver, via the hepatic veins to the inferior vena cava and the right atrium of the heart and to the stomach and the oesophagus. This vascular spread is characteristic of hepatocellular carcinoma and is always seen in advanced cases at autopsy (**Figure 3**). Lymphatic spread to portal lymph nodes and beyond occurs in about one-third of cases. Involvement of major bile ducts and dissemination within the peritoneal cavity are much less common. Distant metastases are usually found in the lungs, adrenals, pancreas, kidney, ovary and bone.

A frequently debated issue is multicentricity of hepatocellular carcinoma, namely the simultaneous development of separate primary tumours, usually in a cirrhotic liver. This undoubtedly occurs as evidenced by studies of differences in patterns of HBV DNA integration and in

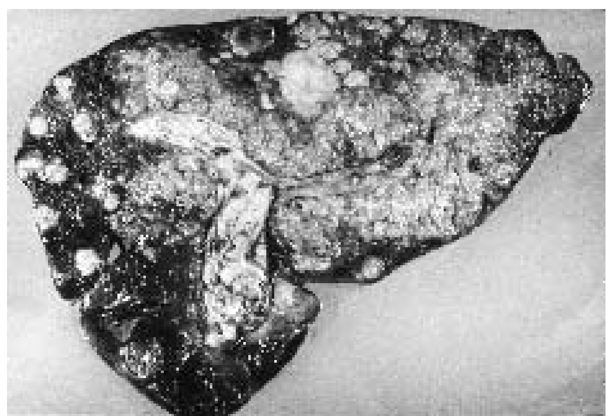


Figure 3 Cut surface of advanced hepatocellular carcinoma at autopsy. Most of the liver is replaced by tumour nodules. The triangular structure in the middle is the main portal vein and its two branches which are filled with tumour.

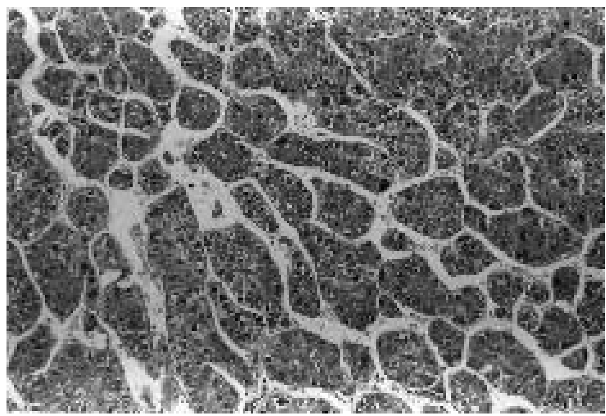


Figure 4 The microscopic appearances of hepatocellular carcinoma resemble those of the normal liver. The tumour is arranged in thick plates which are separated by sinusoids. Malignancy is manifested by the thickness of the plates and the variability in size and shape of nuclei.

dissimilar mutations of oncogenes and tumour-suppressor genes. However, separate tumours can also develop by vascular spread from the same primary source.

Microscopic Features: Histological Patterns and Cytological Variants, Ultrastructure and Special Techniques

The main microscopic characteristic of hepatocellular carcinoma is its resemblance to the normal liver, both in its plate-like growth pattern separated by sinusoids and its cytological appearances (**Figure 4**). Certain architectural and cytological variants have been recognized in the World Health Organisation classification (Ishak *et al.*, 1994) which are helpful for diagnosis but have no other, i.e. clinical or biological, significance.

The commonest architectural pattern is the plate-like or, as seen in two-dimensional histological sections, trabecular. Tumour cells grow in cords that vary in thickness from two to three to many cells. These are separated by sinusoids lined by flat endothelial cells. Kupffer cells are absent or reduced in number. Collagen fibres are increased in Disse's space surrounding the sinusoids and a basement membrane forms, i.e. they become 'capillarized'. The pseudoglandular pattern may result from dilatation of bile canaliculi or from central breakdown of cells in otherwise solid trabecula; the contents are bile or proteinaceous fluid. Compact, solid or scirrhous patterns are rare and develop from compression, scarring and chemo-radiotherapy. The term 'sclerosing hepatic carcinoma' has been applied to tumours associated with hypercalcaemia.

The commonest cytological variant is hepatic or liver-like. Tumour cells are polygonal, with vesicular nuclei and prominent nucleoli. The cytoplasm is finely granular and more basophilic than that of normal liver cells. Bile canaliculi may be seen. Pleomorphic cells show marked variation in nuclear size, shape and staining. Clear cells have a seemingly empty cytoplasm. This may be due to accumulation of water, fat or glycogen. Sarcomatoid tumours form spindle and giant cells. Recently, a small-cell variant has also been described. Recognition of these cytological variants is diagnostically important so that they can be distinguished from metastatic tumours of similar appearance: renal cell carcinoma, soft tissue sarcoma and carcinoma of the lung, respectively.

A number of intracellular inclusions may be seen. Globular hyaline bodies are small, round and often acidophilic. They usually consist of α-1-antitrypsin. Mallory bodies, similar to those seen in alcoholic liver disease, are made up of altered cytokeratins.

Grading of hepatocellular carcinoma into well-, moderately and poorly differentiated tumours is traditional but it is less important than stage for prognosis.

Special staining techniques commonly employed include silver impregnation, which demonstrates the deficiency of the reticulin framework in hepatocellular carcinoma in contrast to normal or cirrhotic liver and hepatocellular adenoma in which it is normal or increased (**Figure 5**). Others are the PAS reaction for the demonstration of glycogen and trichrome methods for cytoplasmic inclusions.

Electron microscopy is relatively little used nowadays owing to the need for fresh tissue, long time of preparation and expense. However, the ultrastructural features of hepatocellular carcinoma are useful for diagnosis. Numerous mitochondria, a well-developed endoplasmic reticulum and, most of all, the presence of intercellular bile canaliculi are pathognomonic.

Immunocytochemistry, however, has replaced electron microscopy in the investigation of liver tumours. It can be done on formalin-fixed, paraffin-embedded tissues and it

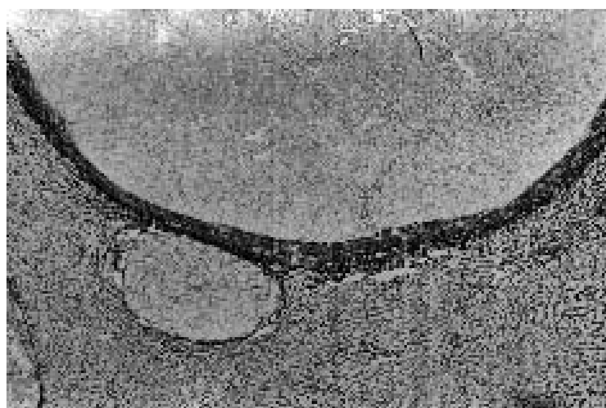


Figure 5 The curved band in the middle is the capsule of a hepatocellular carcinoma. The normal connective tissue framework of the non-tumorous liver is present below whilst it is almost completely absent in the tumour above. Note also that the tumour has broken through the capsule.

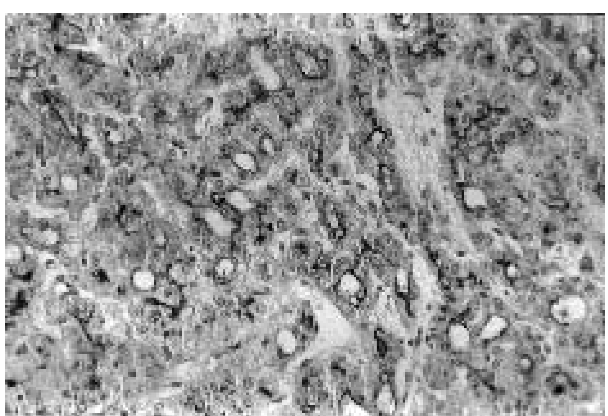


Figure 6 The darkly staining irregular circles and branching structures are bile ductules, stained by peroxidase-labelled polyclonal antisera against CEA/ biliary glycoproteins.

is quick and cheap. The demonstration of normal 'export' proteins such as α-1-antitrypsin, albumin, fibrinogen, ferritin and methallothionein is useful in identifying a tumour as being of liver cell origin. Normal adult liver cells express cytokeratins Nos 8 and 18 as defined in Moll's catalogue whereas bile duct cells contain these and Nos 7 and 19 in addition. The latest addition to liver cell markers is Hep Par1 (Leong *et al.*, 1998). The specificity of α-fetoprotein is high but its sensitivity is low. It is usually only demonstrable in tumour tissue when the serum levels are high, in excess of 5000 ng mL^{-1}. Carcinoembryonic antigen is widely distributed in glandular tissues in the body and polyclonal antibodies raised against it react with biliary glycoproteins. This allows the demonstration of bile canaliculi without recourse to electron microscopy (**Figure 6**).

Fine needle aspiration cytology is a useful diagnostic aid in expert hands, especially when smears are combined with histology of centrifuged needle washings which may contain tiny tissue fragments.

Molecular Genetic Findings

Hepatic carcinogenesis has been extensively studied for many years in experimental animals, usually rodents, by the use of chemicals. Results have led to the concept of multistep neoplastic development from initiation through promotion to progression. This is summarized in **Table 4**. Most of the chemicals used are not likely to play a part in the pathogenesis of human hepatocellular carcinoma, the experiments were short term and the phenotypic changes seen in animals are not definitely identified in humans. As evidence for the causative role of viruses began to emerge, attention has switched to mechanisms of viral carcinogenesis. These are presented in **Figure 7**. The possible pathways of hepatitis B and C viruses and of aflatoxin have been discussed already. Both **Tables 4** and **Figure 7** are somewhat wishful and idealized since such evidence as we have is patchy. Broadly speaking, the genetic changes in carcinogenesis at any site are multiple and cumulative and their phenotypic expressions are not easily identified. What we have at present is an incomplete mosaic from which most of the pieces are missing. However, they can be broken down into alterations in cell cycle regulation, growth factors, oncogenes, tumour-suppressor genes and epigenetic changes, such as telomerase activity. All of these have been investigated in hepatocellular carcinoma and found to be acting abnormally in a varying proportion of cases (Feitelson and Duan, 1997; Geissler *et al.*, 1997; Idilman *et al.*, 1998; Chen and Chen, 1999; Hirohashi *et al.*, 2000).

Completion of the cell cycle requires the successive activation of cyclin-dependent protein kinases which are opposed by their inhibitors. The latter allow for DNA repair. Altered expression leads to the uncontrolled growth of liver cells. Growth factors are generally polypeptides that act at short range via signal transduction pathways across cell membranes. Insulin-like growth factor II (IGF-II) may be an early marker of malignant transformation whilst transforming growth factor α (TGFα) may play a role in its maintenance. Transformed hepatocytes are resistant to transforming growth factor β (TGFβ), which inhibits cell proliferation. In human hepatocellular carcinoma, oncogene (*ras*, *myc*, *fos* families) activation is a rare and tumour-suppressor gene (*p53*, *Rb*) activation is a late event. The most important epigenetic change perhaps is increased telomerase activity which immortalizes the cell.

None of these changes are 'stand-alone' events: they interlock and co-act to produce genomic instability which is amply supported by many chromosomal abnormalities found in hepatocellular carcinoma.

Table 4 Chemical hepatic carcinogenesis (multistep neoplastic development)

Initiation	Metabolic conversion of proximate to ultimate carcinogen
	Damage to DNA, organelles, membranes
	Fixation of abnormality by cell proliferation
Promotion	Inhibition/selection: clonal expansion
	Altered cell populations:
	Enzyme changes
	Hyperplastic nodules
	Autonomous neoplastic nodules
Progression	Metastases and death
	Increased cell turnover all-important

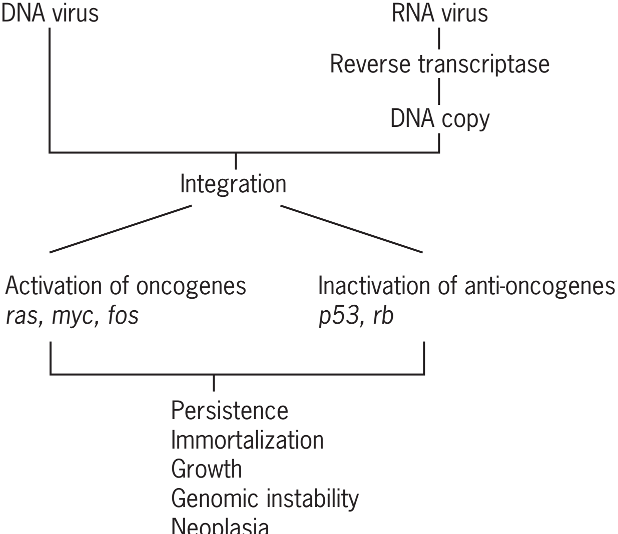


Figure 7 Viral hepatic carcinogenesis

Prognostic Factors

The outlook for hepatocellular carcinoma patients presenting with symptoms is poor, most specialized centres reporting a < 5% survival rate amongst those considered suitable for treatment. Better results can be obtained in patients with small, asymptomatic tumours. Good prognostic indicators are a serum α-fetoprotein level below 100 ng mL^{-1}, a solitary tumour, size < 5 cm and preferably < 2 cm, lack of portal or hepatic vein involvement, early stage in terms of the TNM classification and well-preserved liver function (Okuda, 1993; Akriviadis *et al*., 1998; Schafer and Sorrell, 1999).

Presentation, Clinical Diagnosis and Management

In Western countries and particularly in Japan, patients are middle aged or elderly, have had cirrhosis for years and the onset of malignancy is heralded by a sudden deterioration of their condition. In tropical Africa and Asia, patients are often young adults and cirrhosis is first discovered at the same time as the tumour, which is often large and the course is fulminant. Common presenting features are lethargy, pain, hepatomegaly or a mass, fever, weight loss and, in the case of large tumours, rupture and intraperitoneal haemorrhage. Some patients experience a variety of paraneoplastic syndromes such as hypoglycaemia, erythrocytosis and hypercalcaemia.

The most useful laboratory test is a serum α-fetoprotein level over 2–400 ng mL^{-1}, which is found in ~80% of cases. Frequently used imaging methods are ultrasound, computed tomography and magnetic resonance imaging (Okuda, 1993; Akriviadis *et al*., 1998; Curley, 1998).

Treatment modalities include surgical resection, hepatic artery embolization or chemotherapy, percutaneous injection of alcohol and total hepatectomy followed by liver transplantation. The respective roles of local attempts at tumour removal or destruction and hepatectomy with transplantation have not been finally determined yet and each has its advocates. Once the tumour has spread outside the liver, no form of therapy is effective.

An intriguing aspect of hepatocellular carcinoma is spontaneous regression which has been reported on occasions over the years.

Fibrolamellar Carcinoma

This is a distinctive type of hepatocellular carcinoma that affects adolescents and young adults of either gender. It is not associated with cirrhosis, serum α-fetoprotein is seldom raised and < 10% of patients show evidence of HBV or HCV infection. Despite all this, fibrolamellar carcinoma is of liver cell origin (Anthony, 1994; Ishak *et al*., 1994; Hirohashi *et al*., 2000).

Patients present with malaise, weight loss and a palpable mass, which is usually solitary and often large, 5–25 cm. The tumour cells are large, with vesicular nuclei and granular, pink cytoplasm, due to the presence of numerous mitochondria. Cytoplasmic globules and 'pale bodies' commonly represent α-1-antitrypsin and fibrinogen,

respectively. Bile droplets, copper and copper-associated protein may be present. Bile canaliculi may be difficult to demonstrate. An abundant fibrous stroma, arranged as lamellae of coarse collagen is characteristic and is a *sine qua non* requirement for diagnosis. This fibrous incarceration of the tumour has been held to be responsible for its slow growth and favourable prognosis.

The surgical resectability rate is high and the 5-year survival figures are in the region of 50%.

HEPATOBLASTOMA

This is the most frequently occurring liver tumour in children with a peak incidence in the second year of life. A few present at birth or develop in early adolescence. Males are twice as commonly affected as females. Hepatoblastoma consists of immature liver cells in varying stages of development and a mesenchymal component that is frequently osteoid (Anthony, 1994; Ishak *et al*., 1994; Stocker and Conran, 1997). One-third of patients with hepatoblastoma have a congenital anomaly (hemihypertrophy, cleft palate, talipes, cardiac or renal malformation), a syndrome (Beckwith–Widemann and Down) or other childhood tumour (nephroblastoma). There is also an increased incidence in familial adenomatous polyposis. Chromosomal abnormalities are common.

The usual presentation is with failure to thrive, loss of weight and a rapidly enlarging abdominal mass. The serum α-fetoprotein level is almost invariably high. Virilization is seen in a small minority, due to production of chorionic gonadotrophin by the tumour. Increased urinary excretion of cystathionine occurs in about half of cases.

Hepatoblastoma usually forms a single mass and is often large when first detected, up to 25 cm. It is well circumscribed, with a thin capsule and a partly solid, partly cystic, fibrous, gelatinous or haemorrhagic cut surface. The epithelial component consists of embryonal, foetal, occasionally adult-like liver cells or glandular structures and keratinizing squamous cells. The commonest mesenchymal elements are undifferentiated spindle cells and osteoid but cartilage, bone and striated muscle may be seen. Rarely, anaplastic small cells, neuroendocrine differentiation and melanin production are present. Extramedullary haematopoiesis is common. Immunocytochemistry shows a wide range of differentiation pathways.

Most hepatoblastomas fall into the epithelial or mixed epithelial and mesenchymal categories (**Figure 8**) but a more detailed classification defines six categories: epithelial (pure foetal, combined foetal and embryonal, macrotrabecular, small cell undifferentiated), mixed epithelial and mesenchymal and mixed with teratoid features (Stocker and Conran, 1997). These categories have limited prognostic significance.

Although half of patients are inoperable at presentation and one-fifth have pulmonary metastases, pre-operative chemotherapy allows local resection or total hepatectomy followed by transplantation to be carried out in $> 90\%$ of cases. The overall survival rate is 65–70%. A rapid fall of serum α-fetoprotein levels after surgery is a particularly good prognostic sign.

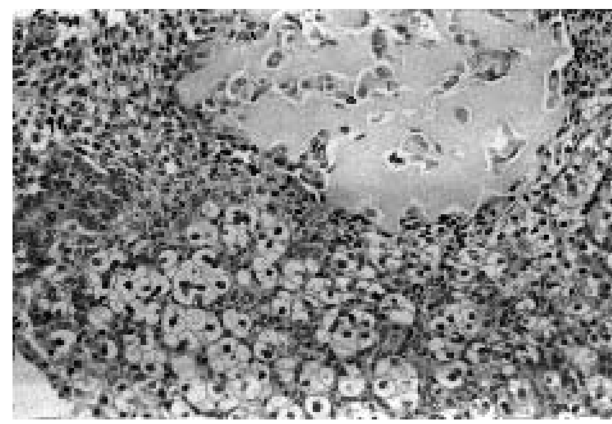


Figure 8 Dark and clear foetal-type cells of hepatoblastoma, mesenchymal spindle cells and structureless osteoid (bone matrix).

CHOLANGIOCARCINOMA (INTRAHEPATIC AND HILAR BILE DUCT CARCINOMA)

Cholangiocarcinoma is a malignant tumour composed of structures resembling bile ducts (Ishak *et al*., 1994). It may be intrahepatic, arising from small bile ducts within the liver or hilar, arising from large bile ducts near the porta hepatis. The clinical presentation is somewhat different according to location. The pathology is simple: all these tumours are mucus secreting adenocarcinomas (Anthony, 1989, 1994; Nakanuma *et al*., 2000).

Epidemiology

Cholangiocarcinoma is much less common than hepatocellular carcinoma and constitutes about 15% of all liver cancers (Parkin *et al*., 1997). It is distributed equally throughout the world except in South-East Asia, particularly Thailand, Laos, Korea, Hong Kong and Canton, where its incidence is increased. This is due to the high prevalence of liver fluke infestation in these areas. Patients are middle aged to elderly, there is no gender difference and the tumour is not associated with cirrhosis.

Aetiology

Liver Fluke Infestation

Infestation with the liver flukes *Opistorchis viverrini* in Thailand and Laos and *Clonorchis sinensis* in

Korea, Hong Kong and Canton is the major cause of cholangiocarcinoma in these high-incidence areas. The life cycle of liver flukes requires poor environmental conditions with infested human waste discharged into stagnant water, take-up by snails as intermediate hosts and the habit of eating raw or undercooked fish which themselves have become infested. A heavy parasite load is acquired over many years and results in cholangitis, liver abscess and cholangiocarcinoma. The presence of liver flukes may not, in itself, be carcinogenic and nitrosamines, derived from diet, may act as cofactors.

Hepatolithiasis

Intrahepatic biliary stones are frequently associated with clonorchiasis but not with opistorchiasis. Cystic lesions of bile ducts may also be complicated by stones.

Chronic Inflammatory Bowel Disease

Cholangiocarcinoma is a significant complication of long-standing ulcerative colitis commonly preceded by sclerosing cholangitis.

Congenital Anomalies of the Biliary Tree

These include cystic dilatation of the biliary tree or Caroli disease, choledochal cysts, biliary microhamartomas and anomalies of the union between the main pancreatic and common bile ducts.

Thorotrast

This once commonly used radiological contrast medium is best known for causing angiosarcoma of the liver but, in recent years, cholangiocarcinomas have also been observed.

Precancerous Changes, Screening and Prevention

The best studied precancerous changes are those associated with liver flukes and biliary stones (**Figure 9**). These consist of adenomatous hyperplasia, dysplasia with multilayering, nuclear enlargement and hyperchromasia and carcinoma-*in-situ* without invasion (Nakanuma *et al*., 2000). Cell kinetic studies have shown increasing proliferative activity along this sequence and the acquisition of genetic abnormalities.

There is no effective screening method for cholangiocarcinoma and prevention consists of reducing the incidence of liver fluke infestation, surveillance of individuals with known high-risk factors and surgical removal of choledochal cysts which carry a particularly high risk of malignant change.

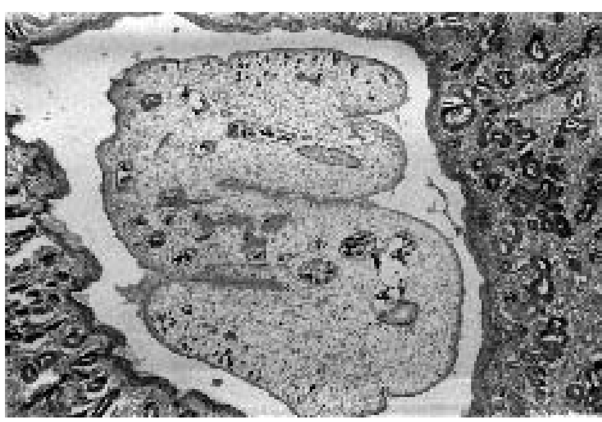


Figure 9 A female *Clonorchis sinensis* fluke lying in a bile duct, the lining of which shows glandular adenomatous hyperplasia. The small, dark, oval objects in the worm's body are eggs.

Macroscopic and Microscopic Pathology

The Macroscopic Pathology of Cholangiocarcinoma

The gross appearance of intrahepatic tumours is of a grey–white, tough, scirrhous type of growth which is often solitary but may be multinodular, or a combination of both. Central necrosis, scarring and calcification may be seen. Finger-like extensions around the main mass represent spread along portal tracts. Metastatic spread is common, to regional lymph nodes, lungs and the peritoneum. Tumours in the hilum present as ill-defined nodules, strictures or, rarely, as an intraductal papillary growth. The flow of bile from the liver is often obstructed and the liver is stained green whilst the gall bladder and common bile ducts are empty.

Microscopic Features: Histological Patterns, Ultrastructure and Special Techniques

Most cholangiocarcinomas are mucus-secreting adenocarcinomas of a tubular pattern and an abundant fibrous stroma is characteristic (**Figure 10**) (Colombari and Tsui, 1995). PAS diastase-resistant or mucicarmine-positive mucus is readily demonstrable and tumour cells express carcinoembryonic antigen in their cytoplasm as well as on their luminal border. The tumour may also grow in solid cords and form papillae. Rarely, it is of signet ring or clear cell type. Large amounts of extracellular mucus are sometimes formed in which tubulopapillary fragments of tumour appear to float freely. Tumours associated with stones, cysts or bile duct anomalies may be adenosquamous or purely squamous. Sarcomatoid cholangiocarcinoma is rare.

Electron microscopy is seldom used in the diagnosis of cholangiocarcinoma. It shows glandular characteristics: lack

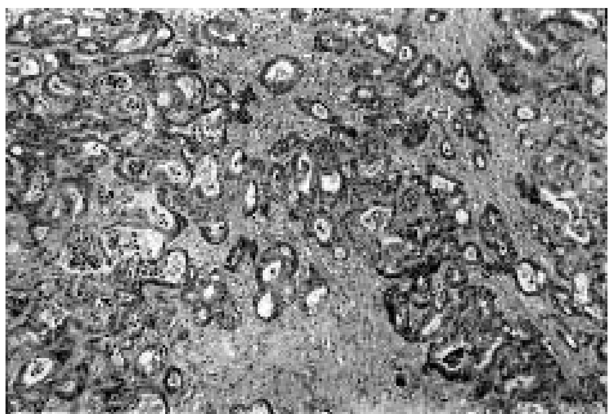


Figure 10 Cholangiocarcinoma made up of abnormally branching bile duct-like structures, separated by a dense fibrous stroma.

of organelles, presence of tonofilaments and a basal lamina.

Immunocytochemistry is used mainly to distinguish cholangiocarcinoma from metastatic adenocarcinoma and, less commonly, from hepatocellular carcinoma. The most useful are demonstration of different cytokeratin patterns for the former and Hep Par 1 for the latter (Leong *et al*., 1998).

Molecular Genetic Findings

Mutations of the *ras* oncogene and the *p53* tumour-suppressor gene are the most common genetic abnormalities in cholangiocarcinoma followed by over-expression of c-*erbB-2* (Nakanuma *et al*., 2000).

Prognostic Factors

Early detection of cholangiocarcinoma is difficult and most patients present with advanced tumours. Lymph node involvement, positive margins and bilobar distribution are associated with high recurrence rates after surgical resection. Patients with unrelieved obstruction of major hepatic ducts and those with cysts and stones may die of complications, e.g. sepsis or liver failure, before the tumour itself becomes evident.

Presentation, Clinical Diagnosis and Management

Malaise, abdominal pain and weight loss are common symptoms of intrahepatic tumours whilst hilar tumours cause unremitting obstructive jaundice. Ultrasound and computer-assisted tomography show the location of the tumour which can then be biopsied for a definitive diagnosis. Only a small minority of patients are suitable for surgical resection and other forms of treatment, e.g. radio- or chemotherapy, are ineffective.

BILIARY CYSTADENOMA AND CYSTADENOCARCINOMA

These are rare tumours. Their pathology is virtually identical with that of similar tumours seen in the ovary and, occasionally, in the pancreas. Most are mucinous and a minority are serous. Patients present with solitary masses which may be fairly large but they are usually amenable to surgical resection and the results are good.

MIXED HEPATOCELLULAR CARCINOMA AND CHOLANGIOCARCINOMA

The presence of both bile secretion and mucus production must be present, supported by appropriate immunocytochemical markers, for such a diagnosis to be made. Effective treatment is seldom possible and the prognosis is poor.

METASTATIC TUMOURS

These are nearly always secondary carcinomas and the common primary sites are the colon, rectum, upper gastrointestinal tract, pancreas, lung and breast. Liver metastases indicate advanced disease and most patients die within a few weeks or at most months. Colonic and, especially, rectal carcinoma metastases may be solitary or few in number and worthwhile results have been achieved by partial hepatectomy.

SARCOMAS OF THE LIVER

These are all rare but angiosarcoma, childhood sarcomas and malignant lymphoma are the most important (Ishak, 1997).

Angiosarcoma

The cause of this tumour is unknown in most cases but a minority are associated with exposure to the once popular radiological contrast agent Thorotrast (thorium dioxide), arsenic and the industrial contaminant vinyl chloride monomer. Thorium is radioactive with a half-life of approximately 400 years and, when injected, most of it is taken up by the liver. In addition to angiosarcoma, hepatocellular carcinoma and cholangiocarcinoma have also been observed. Particles of Thorotrast are readily visualized in histological sections as coarse, pink–brown granules (**Figure 11**).

Angiosarcoma of the liver is always fatal and at autopsy it appears as ill-defined, spongy and haemorrhagic, or greyish–white fibrous nodules which replace the entire

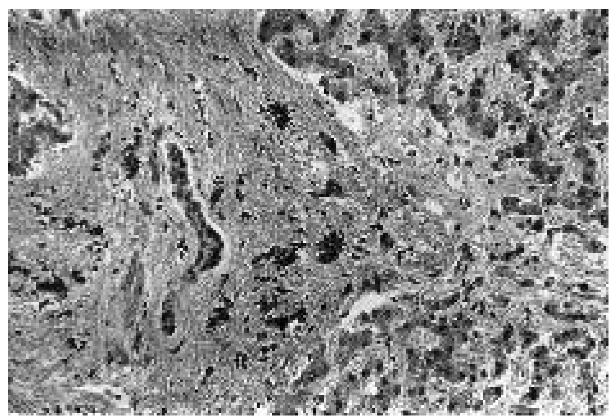


Figure 11 The dark granular material next to a portal bile duct is Thorotrast. There is much fibrosis, separating the liver cell plates around but angiosarcoma has not yet developed.

Figure 12 Reddish, dark and white (fibrous) angiosarcoma replaces the entire liver.

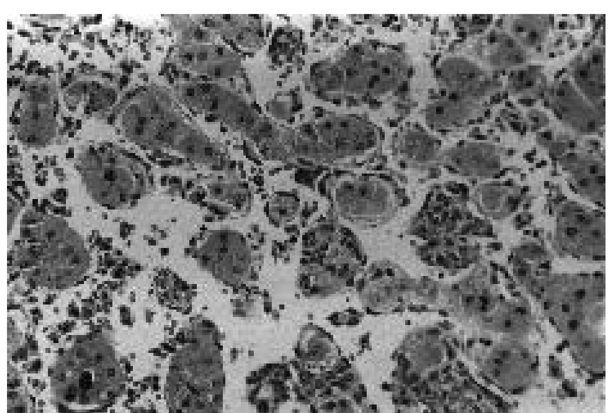


Figure 13 Dark, spindle and polygonal cells of angiosarcoma grow over the surface of liver cell plates in a scaffold-like or tectorial fashion.

liver (**Figure 12**). The histological appearances are variable but the most characteristic is a scaffold-like or tectorial growth of spindle cells on the surface of liver cell plates which eventually atrophy and disappear (**Figure 13**). Immunocytochemistry shows that tumour cells react with endothelial markers such as CD31, CD34 and Factor VIII related antigen or von Willebrand factor.

Another, distinctive form of malignant vascular tumour is epithelioid haemangioma, which occurs in the liver, lung, skin and bone and has a much better prognosis than angiosarcoma.

Childhood Sarcomas

These are embryonal sarcoma and rhabdomyosarcoma. They are much less common than hepatoblastoma. Their prognosis has been hopeless until recently but aggressive treatment combining surgical excision, chemotherapy and radiation has led to 5-year survival rates of around 15%.

Primary Malignant Lymphoma

All types of lymphoma may secondarily involve the liver in their advanced stage and the prognosis is then poor. However, it is increasingly recognized that lymphoma can also be primary in the liver and the outcome, with appropriate treatment, is much more favourable. They are all non-Hodgkin lymphomas of B or T cell lineage. Some have been associated with HBV or HCV infections and AIDS.

BENIGN TUMOURS AND TUMOUR-LIKE LESIONS

The most important benign tumour of the liver is hepatocellular adenoma. The majority of cases occur in young to middle-aged women who have taken oral contraceptive steroids for years or in individuals of either gender on long-term treatment with androgenic/anabolic steroids. Hepatocellular adenoma is often symptomatic, may grow to a large size, and rupture, giving rise to life-threatening intraperitoneal haemorrhage. The relationship of focal nodular hyperplasia to oral contraceptive steroids is much less certain and is often an asymptomatic, incidental finding. Inflammatory myofibroblastic tumour is an intriguing entity, the nature of which is not fully understood, i.e. whether inflammatory or neoplastic, but surgical excision is curative. Many other types of benign tumour have been described but they are all rare or clinically unimportant (Anthony, 1994).

TUMOURS OF THE GALL BLADDER

Nearly all tumours at this site are mucus secreting adenocarcinomas (Albores-Saavedra and Henson, 1986).

The incidence of carcinoma of the gall bladder is variable in different parts of the world and amongst ethnic groups. Overall, it ranks fourth amongst cancers of the digestive tract, after those of the colo-rectum, stomach and pancreas. The highest frequency is seen in South West American Indians and it is also common in Mexico, Chile, Bolivia and Israel.

South West American Indians apart, who seem to have a genetic predisposition to gall bladder carcinoma, the most important predisposing factor in most countries is chronic inflammation of the gall bladder associated with stones (cholelithiasis). Secondary factors are old age, female gender, obesity, abnormalities of bile and lipoprotein metabolism and multiple pregnancies which, themselves, predispose to gall stone formation. There is also a link with chronic inflammatory bowel disease, anomalous pancreaticobiliary duct union and the typhoid carrier state.

Carcinoma of the gall bladder is commonly preceded by epithelial hyperplasia, dysplasia and carcinoma-*in-situ*. Malignant change may also develop in tubular or villous adenomas. Rarely, biliary papillomatosis affects the gall bladder, the intra- and extrahepatic bile ducts and even the pancreatic duct system and carcinomas can arise at any of these sites.

There is no effective screening method for gall bladder carcinoma and the only means of prevention is prophylactic cholecystectomy which has been advocated for high-risk groups such as American Indian females beyond middle age.

Carcinoma of the gall bladder is often a silent disease and the tumour is discovered incidentally at cholecystectomy or else the symptoms are those of cholecystitis or cholelithiasis. These include intolerance of fatty meals, right upper quadrant abdominal pain and intermittent nausea or vomiting. Weight loss and jaundice are indicative of advanced disease. Ultrasonography and computed tomography are used in the assessment of patients with symptoms attributable to gall bladder disease.

Macroscopically, carcinoma of the gall bladder may appear as diffuse thickening of the wall, an ill-defined nodule or a polyp. Microscopically, most tumours are well to moderately differentiated adenocarcinomas with a tubular or tubulo-papillary pattern. Mucus secretion is nearly always demonstrable. Intestinal (with endocrine, Paneth and goblet cells), mucinous (with much extracellular mucin), signet ring, clear and small-cell variants are recognized. Ultrastructural or immunocytochemical studies are not particularly useful for diagnosis: the features are those of an adenocarcinoma occurring at many sites.

The majority of gall bladder carcinomas express mutated *p53* tumour-suppressor gene. *Ras* and other oncogene mutations are late events. Amplification of the c-*erbB-2* gene has also been found.

Surgical excision is the only effective method of treatment. The best results are obtained before the gall bladder wall has been breached. Direct extension into the liver and regional lymph node metastasis are indicators of a poor prognosis. Ultimately, patients die of disseminated disease with direct, lymphatic and blood-borne spread to many possible sites in the body.

TUMOURS OF THE EXTRAHEPATIC BILE DUCTS

As in the gall bladder, the commonest tumour in the extrahepatic bile ducts is an adenocarcinoma. About half arise in the common hepatic, cystic and upper common bile ducts, one quarter in the middle and one-tenth in the lower common bile duct; the rest are diffuse or multiple tumours. In general, the prognosis is worse for carcinomas of the proximal than of the middle or distal segments of the extrahepatic biliary tree. Most patients are elderly and males are more commonly affected than females, in contrast to carcinoma of the gall bladder. Cholelithiasis is not a risk factor. Ulcerative colitis, malunion of the main pancreatic and common bile ducts, congenital malformations such as choledochal cyst, pre-existing adenomas and papillomas are well-known predisposing factors. The onset of invasive malignancy may be preceded by dysplasia and carcinoma-*in-situ*, which are sometimes seen next to it in operative specimens. Most patients present with obstructive jaundice, some with ascending cholangitis and a few with blood in the bile (haemobilia). At laparotomy, extrahepatic bile duct carcinoma appears as a polyp, a stricture or a diffuse thickening. Histologically, most are tubular or tubulo-papillary adenocarcinomas with a fibrous stroma. A good histological grade has been claimed to confer a prognostic advantage but this is not universally accepted. Spread to adjacent structures or lymph node metastasis are indicative of a poor prognosis. The only effective treatment is complete surgical excision.

REFERENCES

Akriviadis, E. A., *et al*. (1998). Hepatocellular carcinoma. *British Journal of Surgery*, **86**, 1319–1331.

Albores-Saavedra, J. and Henson, D. E. (1986). *Tumors of the Gall Bladder and Extrahepatic Bile Ducts, Atlas of Tumour Pathology*, 2nd Series, Fascicle 22. (Armed Forces Institute of Pathology, Washington, DC).

Anthony, P. P. (1989). Epidemiology, aetiology and pathology of bile duct tumours. In: Preece, P. E., *et al*. (eds), *Cancer of the Bile Ducts and Pancreas.* 1–26 (Saunders, Philadelphia).

Anthony, P. P. (1994). Tumours and tumour-like lesions of the liver and biliary tract. In: MacSween, R. N. M., *et al*. (eds), *Pathology of the Liver.* 635–711 (Churchill Livingstone, Edinburgh).

Beasley, R. P. (1988). Hepatitis B virus. The major etiology of hepatocellular carcinoma. *Cancer*, **61**, 1942–1956.

Bosch, F. X., *et al*. (1999). Epidemiology of primary liver cancer. *Seminars in Liver Disease*, **19**, 271–285.

Chen, P.-J. and Chen, D. S. (1999). Hepatitis B virus infection and hepatocellular carcinoma: molecular, genetic and clinical perspectives. *Seminars in Liver Disease*, **19**, 253–262.

Colombari, R. and Tsui, W. M. S. (1995). Biliary tumours of the liver. *Seminars in Liver Disease*, **15**, 402–413.

Colombo, M. (1999). Hepatitis C virus and hepatocellular carcinoma. *Seminars in Liver Disease*, **19**, 263–269.

Craig, J. R., *et al*. (1989). *Tumors of the Liver and Intrahepatic Bile Ducts, Atlas of Tumor Pathology*, 2nd Series, Fascicle 26. (Armed Forces Institute of Pathology, Washington, DC).

Curley, S. A. (ed.) (1998). *Liver Cancer*. (Springer, New York).

European Association for the Study of the Liver. (1999). The liver in systemic diseases. *Journal of Hepatology*, **30** (Supplement 1).

Feitelson, M. A. and Duan, L.-X. (1997). Hepatitis B virus X antigen in the pathogenesis of chronic infections and the development of hepatocellular carcinoma. *American Journal of Pathology*, **150**, 1141–1157.

Geissler, M., *et al*. (1997). Molecular mechanisms of hepatocarcinogenesis. In: Okuda, K. and Tabor, E. (eds), *Liver Cancer.* 59–88 (Churchill Livingstone, New York).

Idilman, R., *et al*. (1998). Pathogenesis of hepatitis B and C-induced hepatocellular carcinoma. *Journal of Viral Hepatology*, **5**, 285–299.

International Agency for Research on Cancer (1972–1999). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans and Supplements to the Monographs*, Vols 1–71 and Supplements 1–8 (International Agency for Research on Cancer, Lyon).

Ishak, K. G. (1997). Malignant mesenchymal tumours of the liver. In: Okuda, K. and Tabor, E. (eds), *Liver Cancer*. 291–314 (Churchill Livingstone, New York).

Ishak, K. G., *et al*. (1994). *Histological Typing of Tumours of the Liver*, 2nd edn. (Springer, Berlin).

Hirohashi, S., *et al*. (2000). Hepatocellular carcinoma. In: Hamilton, S. R. and Aaltonen, L. A. (eds), *Pathology and Genetics of Tumours of the Digestive System. WHO Classification of Tumours.* 159–183. (International Agency for Research on Cancer, Lyon).

Leong, A. S.-Y., *et al*. (1998). Hep Par 1 and selected antibodies in the immunohistological distinction of hepatocellular carcinoma from cholangiocarcinoma, combined tumours and metastatic carcinoma. *Histopathology*, **33**, 318–324.

MacSween, R. N. M. and Scothorne, R. J. (1994). Developmental anatomy and normal structure. In: MacSween, R. N. M., *et al*. (eds), *Pathology of the Liver*. 1–49 (Churchill Livingstone, Edinburgh).

Nakanuma, Y., *et al*. (2000). Intrahepatic cholangiocarcinoma. In: Hamilton, S. R. and Aaltonen, L. A. (eds), *Pathology and Genetics of Tumours of the Digestive System. WHO Classification of Tumours.* 173–180. (International Agency for Research on Cancer, Lyon.)

Okuda, K. (1993). Epidemiology and clinical aspects of hepatocellular carcinoma. *Journal of Gastroenterology and Hepatology*, **8** (Supplement 1), S1–S4.

Parkin, D. M., *et al*. (eds) (1997). *Cancer Incidence in Five Continents*, Vol. VII (International Agency for Research on Cancer, Lyon).

Schafer, D. F. and Sorrell, M. F. (1999). Hepatocellular carcinoma. *Lancet*, **353**, 1253–1257.

Simonetti, R. G., *et al*. (1991). Hepatocellular carcinoma. A worldwide problem and the major risk factors. *Digestive Disease Science*, **36**, 962–972.

Stocker, J. T. and Conran, R. M. (1997). Hepatoblastoma. In: Okuda, K. and Tabor, E. (eds), *Liver Cancer*. 263–278. (Churchill Livingstone, New York).

Wogan, G. N. (1999). Aflatoxin as a human carcinogen. *Hepatology*, **30**, 573–575.

FURTHER READING

Bannasch, P., *et al*. (eds) (1989). *Liver Cell Carcinoma*. Falk Symposium 51. (Kluwer, Dordrecht).

Boyer, J. L. and Ockner, R. K. (eds) (1997). *Progress in Liver Diseases*. (Saunders, Philadelphia).

Clavien, P.-A. (ed.) (1999). *Malignant Liver Tumors*. (Blackwell, Oxford).

Curley, S. A. (ed.) (1998). *Liver Cancer*. (Springer, New York).

Goldin, R. D., *et al*. (eds) (1998). *Pathology of Viral Hepatitis.* (Arnold, London).

MacSween, R. N. M., *et al*. (eds) (1994). *Pathology of the Liver*. (Churchill Livingstone, Edinburgh).

Okuda, K. and Tabor, E. (eds) (1977). *Liver Cancer.* (Churchill Livingstone, New York).

Phillips, M. J., *et al*. (1987). *The Liver: an Atlas and Text of Ultrastructural Pathology*. (Raven Press, New York).

Preece, P. E., *et al*. (eds) (1989). *Cancer of the Bile Ducts and Pancreas.* (Saunders, Philadelphia).

Zakim, D. and Boyer, T. D. (eds) (1996). *Hepatology.* (Saunders, Philadelphia).

Chapter 40

Pancreas

Ralph H. Hruban, Robb E. Wilentz and Michael Goggins
Johns Hopkins Medical Institutions, Baltimore, MD, USA

CONTENTS

- Normal Development and Structure
- Tumour Pathology
- Epidemiology and Aetiology
- Screening and Prevention
- Molecular Genetics
- Prognostic Factors
- Overview of Present Clinical Management
- Conclusions

NORMAL DEVELOPMENT AND STRUCTURE

As is true for many organs, an understanding of the embryology of the pancreas can lead to a better understanding of the diseases that affect the gland. The five parts of the pancreas (head, uncinate, neck, body and tail; **Figure 1a**) develop from two separate outpouchings of the embryological foregut (Solicia *et al*., 1997). The first outpouching, called the 'ventral anlage,' will give rise to most of the head and uncinate process of the pancreas. The second outpouching, called the dorsal anlage, will form the tail, body and inferior (lower) portion of the head of the gland.

These two anlages and their duct systems normally fuse by the eighth week of gestation to form a single organ. The main pancreatic duct, called the 'duct of Wirsung,' is normally formed when the duct in the dorsal anlage fuses with the duct in the ventral anlage at a point close to the duodenum. Because the ventral anlage also forms the common bile duct, the main pancreatic duct usually drains into the duodenum in conjunction with the common bile duct (**Figure 1b**). A remaining portion of the dorsal duct forms the accessory pancreatic duct, called the 'duct of Santorini' and this accessory duct usually drains separately into the duodenum.

The adult pancreas comes to rest in the centre of the abdomen, housed between the duodenum and spleen. The head of the pancreas is closely associated with the duodenum. The uncinate process forms a groove above which important blood vessels, the superior mesenteric artery and vein, pass. The neck, body and tail comprise the parts of the pancreas that are successively closer to the spleen (see **Figure 1a**).

Macroscopically, the pancreas is a single organ. Microscopically, however, the pancreas contains two distinct components, each with an important function (**Figure 2**). The 'exocrine' pancreas secretes digestive enzymes into the duodenum, while the 'endocrine' portion of the pancreas secretes hormones, such as insulin, into the

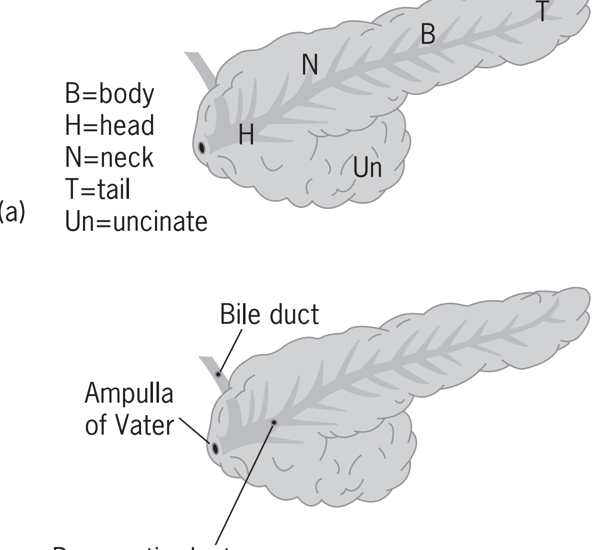


Figure 1 Structure of the adult pancreas. The head is that portion of the pancreas closest to the duodenum, the tail the portion closest to the spleen (a). Note that the main pancreatic duct joins the distal common bile duct at the ampulla of Vater (b). (Adapted from an original medical illustration by Jennifer Parsons, with permission.)

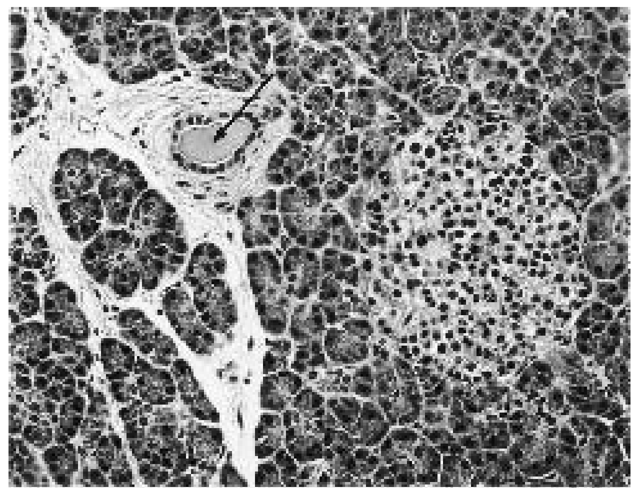


Figure 2 Microscopic section of normal pancreas. The acini comprise the bulk of the exocrine pancreas. The acini produce digestive enzymes which are released into small ductules (arrow) and from there the enzymes travel into the duodenum. The islets of Langerhans (nodule on right) form the endocrine pancreas. They release hormones such as insulin directly into the bloodstream.

bloodstream. The pancreas has cells specialized for each of these tasks, as well as the cells needed to support and nourish the organ (connective tissue, blood vessels).

The exocrine structures of the pancreas include the acini and ducts (**Figure 2**). Acini, which comprise over 80% of the pancreatic mass, are lobular units that secrete fluids rich in bicarbonate and enzymes into small ductules. These small ductules merge to drain their contents into larger ducts, which in turn merge into even larger ducts. This arborizing pattern continues to the level of the main and accessory pancreatic ducts, and the digestive enzymes of the exocrine pancreas are thereby released into the duodenum.

The endocrine portion of the pancreas consists of the islets of Langerhans (**Figure 2**). Although a normal pancreas contains over one million of these islets, they comprise only 1–2% of the organ mass (Solcia *et al.*, 1997). Approximately 70% of the cells in normal islets are α cells, 20% are β cells and 10% are δ cells. The α cells secrete the hormone glucagon, the β cells secrete the hormone insulin and the δ cells secrete the hormone somatostatin.

The various neoplasms of the pancreas can be best understood if one keeps these structures of the normal pancreas in mind.

TUMOUR PATHOLOGY

Cancer of the pancreas is not one disease. Instead, it is a number of different diseases broadly classified under one umbrella term (Klöppel *et al.*, 1996; Solcia *et al.*, 1997). For example, primary, metastatic and systemic tumours can all involve the gland. Primary cancers arise in the pancreas. Metastatic cancers originate in other organs and spread to the pancreas secondarily. Systemic malignancies derive from the blood or lymph nodes and, by definition, simultaneously involve multiple sites, one of which may be the pancreas.

This section provides an overview of the many types of neoplasms that can arise in or spread to the pancreas. It focuses on primary pancreatic neoplasms, which can be benign (tumours which usually do not spread and which usually will not lead to a patient's death), borderline (tumours for which behaviour is difficult to predict) or malignant (tumours which, if untreated, will spread beyond the gland and lead to the patient's death). Not surprisingly, if we remember the dual composition of the normal pancreas, primary pancreatic neoplasms can show either endocrine or exocrine (non-endocrine) differentiation. The non-endocrine neoplasms can be further subclassified into solid or cystic. For example, ductal adenocarcinoma, the most common primary pancreatic malignancy, is a non-endocrine, solid neoplasm. Needless to say, each of the various pancreatic neoplasms is grossly, microscopically and clinically distinct. Therefore, understanding the pathology of pancreatic cancer forms the cornerstone for rational patient diagnosis, treatment and prognostication.

The following sections will describe the most common types of tumours of the pancreas. Wherever possible, the classification systems published by the Armed Forces Institute of Pathology (AFIP) and by the World Health Organisation (WHO) will be followed (see **Table 1**) (Klöppel *et al.*, 1996; Solcia *et al.*, 1997).

Solid Non-endocrine Neoplasms

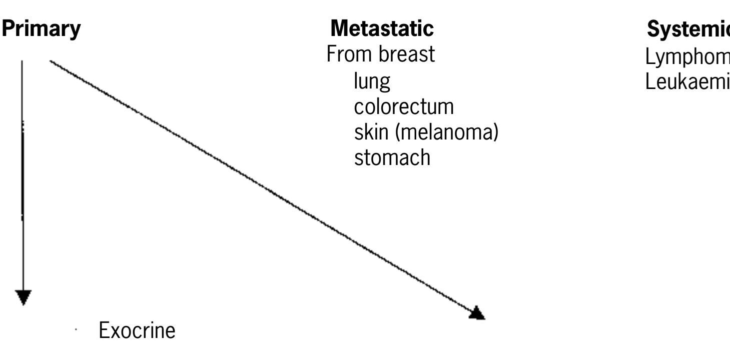
Ductal Adenocarcinoma

Gross and Microscopic Features

Ductal adenocarcinoma is the most common malignancy of the pancreas, accounting for almost three-quarters of all primary cancers (Solcia *et al.*, 1997). Most, but not all, ductal adenocarcinomas arise in the head of the pancreas (Solcia *et al.*, 1997; Hruban *et al.*, 2000). These tumours are infiltrative (**Figure 3; see colour plate section**), firm masses. Those that arise in the head of the pancreas often obstruct and dilate the distal common bile and pancreatic ducts. As a result, many patients with pancreatic cancer develop jaundice, a yellowish discoloration of the skin and eyes caused by obstruction of the flow of bile. Microscopically, ductal adenocarcinomas are composed of infiltrating glands of various shapes and sizes surrounded by reactive connective tissue (**Figure 4a**). The cancer cells may contain abnormal mitotic figures and the nuclei of these cells can show marked pleomorphism (variation in size and shape) and hyperchromasia (increased nuclear staining).

Most ductal adenocarcinomas grow into nerves (**Figure 4b**) and blood vessels, and it is therefore

Table 1 Histological classification of tumours of the pancreas. (Adapted from Solcia *et al.*, 1997.)



Exocrine		Endocrine
	Benign	
1. Serous cystadenoma 2. Mucinous cystadenoma 3. Intraductal papillary–mucinous adenoma		1. Well- or moderately-differentiated islet cell tumour
	Borderline	
1. Borderline mucinous cystic tumour 2. Borderline intraductal papillary–mucinous tumour 3. Solid pseudopapillary tumour		1. Well- or moderately differentiated endocrine neoplasm of uncertain malignant potential
	Malignant	
1. Ductal adenocarcinoma 2. Osteoclast-like giant-cell tumour 3. Mucinous cystadenocarcinoma 4. Serous cystadenocarcinoma 5. Intraductal papillary–mucinous carcinoma 6. Acinar cell carcinoma 7. Pancreatoblastoma 8. Medullary carcinoma		1. Well- or moderately differentiated endocrine carcinoma 2. Poorly differentiated carcinoma (small-cell carcinoma, high-grade neuroendocrine carcinoma)

not surprising that many patients with pancreatic cancer develop severe back pain. In addition, most ductal adenocarcinomas spread to lymph nodes (Solcia *et al.*, 1997). Ductal adenocarcinomas also frequently spread to other organs ('metastasize') (Solcia *et al.*, 1997; Hruban *et al.*, 2000). This spread to lymph nodes and other organs greatly reduces the effectiveness of surgery in the treatment of pancreatic cancer and, not unexpectedly, patients with metastases do significantly worse than do patients without them (Yeo *et al.*, 1995; Solcia *et al.*, 1997; Hruban *et al.*, 2000). Unfortunately, most patients with pancreas cancer do not come to clinical attention until after their cancers have spread. Overall survival from ductal adenocarcinoma is therefore extremely poor, with average survival time of about 6 months.

Special Features: Histological Precursors

One of the more important findings in the pathology of ductal adenocarcinoma of the pancreas has been the identification of the morphological precursors to invasive cancer. Called 'pancreatic intraepithelial neoplasias' (or PanINs for short), these precursors arise in the small pancreatic ducts and ductules, and they are composed of mucin-producing cells with varying degrees of nuclear and architectural atypia. A complete description of the various PanINs as well as numerous examples can be found on the Web (http://pathology.jhu.edu/pancreas_panin). As shown in **Figure 5**, PanINs progress from flat lesions to papillary lesions to atypical papillary lesions to carcinomas *in situ* (non-invasive carcinoma) and finally to invasive cancers (Cubilla and Fitzgerald, 1976; DiGiuseppe *et al.*, 1996; Hruban *et al.*, 2000).

Several lines of evidence suggest that PanINs are the precursors of infiltrating pancreatic ductal adenocarcinoma, just like adenomas are the precursors of infiltrating cancer in the colon. First, PanINs are frequently found in pancreata adjacent to infiltrating cancers (Cubilla and Fitzgerald, 1976). Second, isolated clinical case reports have suggested that PanINs can progress to infiltrating cancer over time (Brat *et al.*, 1998). For example, Brat *et al.* reported three patients who developed infiltrating ductal adenocarcinomas months to years after high-grade PanINs

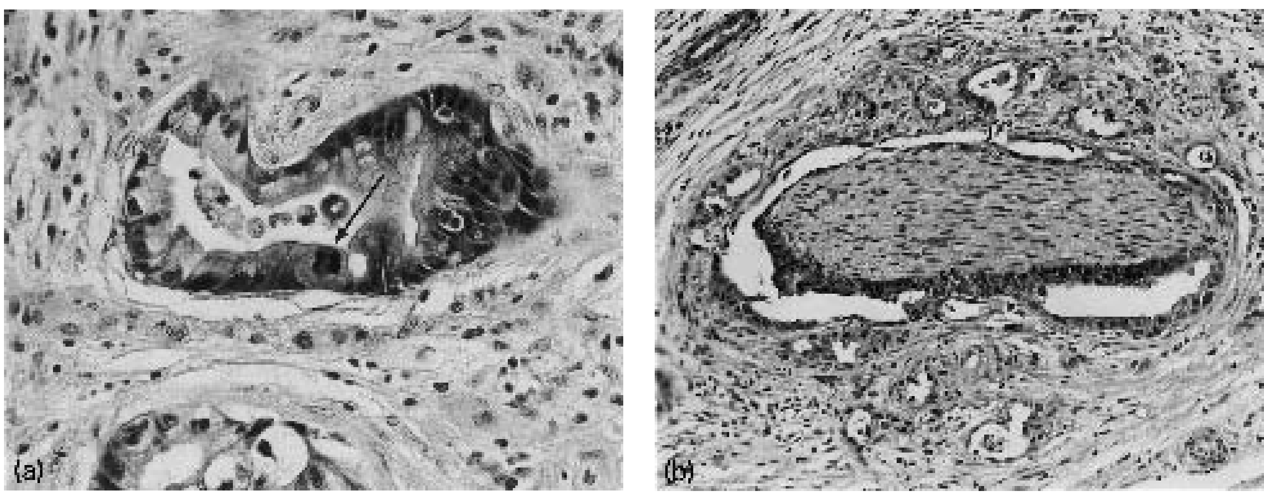


Figure 4 Microscopic section of infiltrating ductal adenocarcinoma. Note how the cancer cells form irregular glands and note the presence of an atypical mitotic figure (arrow) (a). Ductal adenocarcinomas often invade around nerves (b).

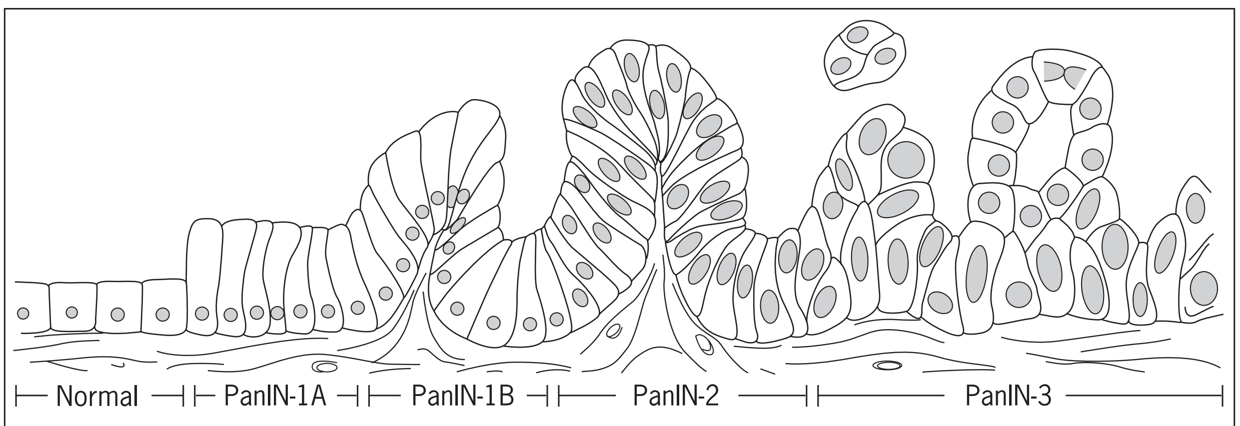


Figure 5 Illustration depicting the progression from normal pancreatic duct (left), to flat duct lesion without atypia (PanIN-1A), to papillary duct lesion without atypia (PanIN-1B), to papillary duct lesion with atypia (PanIN-2), to carcinoma-*in-situ* (PanIN-3). (Adapted from an original medical illustration by Jennifer Parsons, with permission.)

were identified in their pancreata. Third, PanINs display many of the same fundamental genetic changes (mutations in the K-*RAS*, *HER-2/neu*, *BRCA2*, *p16*, *p53* and *DPC4* genes) as do infiltrating adenocarcinomas (Caldas *et al*., 1994; DiGiuseppe *et al*., 1994a, b; Day *et al*., 1996; Moskaluk *et al*., 1997; Willentz *et al*., 1998) (see Molecular Genetics).

An understanding of PanINs and the development of this progression model is important because it suggests that the detection of precursor lesions and curable early pancreatic cancers should one day be possible (Caldas *et al*., 1994). Early detection is especially important in the case of ductal adenocarcinoma of the pancreas because, as mentioned earlier, most pancreatic cancers have already spread beyond the pancreas when they come to clinical attention.

Medullary Carcinoma

Gross and Microscopic Findings

Although historically grouped with ductal adenocarcinomas, medullary carcinomas are a newly recognized subtype of pancreatic cancer with distinct gross and microscopic appearances. Grossly, these neoplasms may be relatively well circumscribed. Microscopically, they have poorly defined cellular boundaries ('syncytial growth pattern') and expanding, rather than infiltrating, tumour borders (Goggins *et al*., 1998). Most also have extensive necrosis.

Special Features: Genetics

Medullary carcinomas are important to recognize because they are genetically distinct tumours, in that they frequently have 'microsatellite instability' (MSI) and wild-type K-*RAS*

genes (Goggins *et al*., 1998). (see Molecular Genetics). These data are atypical for the usual ductal adenocarcinomas, which nearly universally harbour K-*RAS* gene mutations and seldom if ever have MSI (Hruban *et al*., 1993). In addition, the authors have recently shown that patients with medullary carcinomas often have a familial history of cancer and that medullary carcinoma may be a sign of an inherited propensity to develop cancer. Medullary carcinomas also may be associated with an outcome better than that for ductal adenocarcinomas, but more study is necessary (Goggins *et al*., 1998).

Acinar Cell Carcinoma

Gross and Microscopic Findings

Acinar cell carcinomas are usually large and most arise in the head of the pancreas (Klimstra *et al*., 1992; Solcia *et al*., 1997; Hruban *et al*., 2000). In contrast to the infiltrative appearance of ductal adenocarcinomas, acinar cell carcinomas typically have smooth borders. Microscopically, the neoplastic cells in acinar cell carcinomas form small glands, called acini. The cells are pink and granular. These carcinomas usually express the digestive enzymes trypsin, lipase, chymotrypsin and/or amylase, and staining for these substances may be helpful in distinguishing acinar cell carcinomas from other neoplasms that arise in the pancreas. Electron microscopy will reveal the presence of zymogen granules in the neoplastic cells (Klimstra *et al*., 1995). Zymogen granules are the granules which hold or package the digestive enzymes in normal acinar cells.

Special Features: Clinical Presentation

While most patients with acinar cell carcinomas present with non-specific symptoms, as many as 20% develop the clinical syndrome of subcutaneous fat necrosis. This dramatic syndrome is characterized by a skin rash (erythema nodosum-like), peripheral eosinophilia (increased numbers of eosinophils in the blood) and/or polyarthralgias (joint pain involving multiple joints) (Solcia *et al*., 1997). These latter signs and symptoms are caused by the release of massive amounts of the digestive enzyme lipase into the bloodstream by the neoplasm (Klimstra *et al*., 1992). The mean survival for patients with acinar cell carcinoma of the pancreas is only 18 months (Klimstra *et al*., 1992; Solcia *et al*., 1997).

Osteoclast-like Giant-cell Tumour (OCGT)

Gross and Microscopic Findings

OCGTs are typically well circumscribed, yellow–pink and fleshy. By light microscopy these tumours are composed of striking giant cells containing multiple nuclei (multinucleated giant cells) dispersed among bland tumour cells with only one nucleus (mononuclear cells) (Solcia *et al*., 1997; Hruban *et al*., 2000). The multinucleated giant cells in OCGTs closely resemble osteoclasts, a type of cell found in resorbing bone. The multinucleated giant cells in OCGTs express the macrophage markers KP-1 and Mac-387, while the bland mononuclear tumour cells variably express the epithelial marker cytokeratin.

Special Features: Genetics

The striking resemblance of the multinucleated giant cells in OCGTs to the osteoclasts normally found in bone confused pathologists for years. Are these distinctive tumours ductal adenocarcinomas which had elicted an unusual reaction, or are OCGTs more closely related to bone tumours? The answer has come from genetic analysis. OCGTs frequently harbour activating point mutations in codon 12 of the K-*RAS* gene and, as discussed in the Molecular Genetics section, K-*RAS* gene mutations are one of the most common genetic alterations in ductal adenocarcinomas (Hruban *et al*., 1993; Westra *et al*., 1998). These data help establish that OCGTs are really carcinomas which elicit a non-neoplastic, giant-cell response (Westra *et al*., 1998). Osteoclast-like giant cell tumours of the pancreas are unrelated to bone tumours. The survival rate from OCGTs is somewhat better than that for ductal adenocarcinoma (Solcia *et al*., 1997).

Pancreatoblastoma

Gross and Microscopic Features

Pancreatoblastomas have striking gross and microscopic appearances. They are large, necrotic and white–grey (Klimstra *et al*., 1995; Solcia *et al*., 1997; Hruban *et al*., 2000). Microscopically, pancreatoblastomas contain back-to-back, small cells with round nuclei (Klimstra *et al*., 1995). In this sea of small cells are embedded nests of swirled cells, called 'squamoid corpuscles' (**Figure 6**).

Special Features: Clinical Presentation

This rare neoplasm occurs primarily in children, with an age range of 1–15 years (Klimstra *et al*., 1995). It has therefore been referred to as 'pancreatic carcinoma of infancy.' The survival rate for patients with pancreatoblastomas is better than it is for patients with infiltrating ductal adenocarcinomas (Klimstra *et al*., 1995). Unfortunately, very little is known about what causes these tumours to develop in children.

Cystic Non-endocrine Neoplasms

Serous Cystadenoma and Serous Cystadenocarcinoma

Gross and Microscopic Findings

Grossly, serous cystadenomas are large, spongy tumours filled with a watery fluid. They often contain a central, calcified scar (Compagno and Oertel, 1978a). By light microscopy, a layer of simple cuboidal (square-shaped) cells with uniform nuclei can be seen lining the cysts (**Figure 7**) (Solcia *et al*., 1997; Hruban *et al*., 2000).

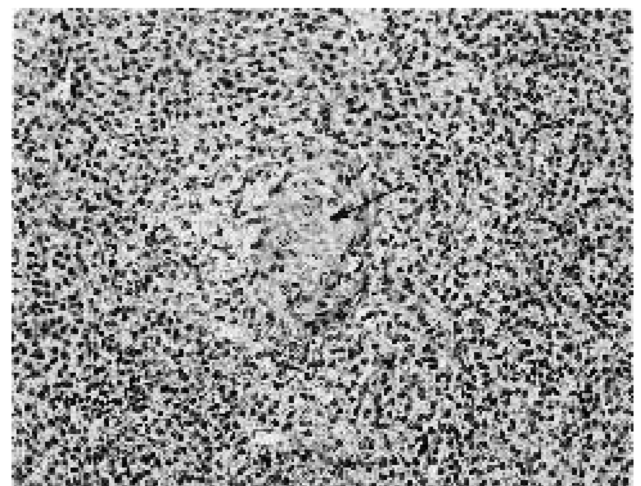


Figure 6 Microscopic section of a pancreatoblastoma. Note the prominent squamous corpuscle (arrow).

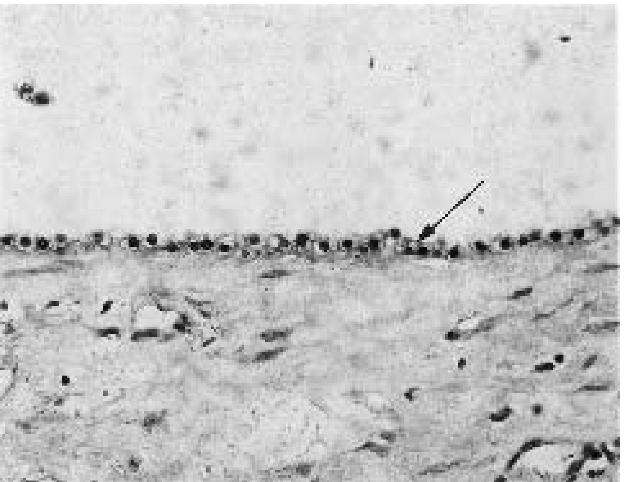


Figure 7 Microscopic section of a serous cystadenoma. The cysts (empty space at top of photograph) in serous cystadenomas are relatively small and are lined by cuboidal, cleared-out cells (arrow).

Because the cells contain large amounts of glycogen, they stain strongly with the periodic acid–Schiff (PAS) stain.

Special Features: Clinical Presentation and Prognosis

These neoplasms are more common in women than in men, and the average age at diagnosis is in the seventh decade. Patients with von Hippel–Lindau syndrome develop serous cystadenomas more frequently than the general population (Solcia *et al*., 1997). The vast majority of serous cystic neoplasms are benign and patients with even very large (football-sized) serous cystadenomas can be cured of their disease if their tumours are surgically removed (Compagno and Oertel, 1978a; Solcia *et al*., 1997). Recently, however, there have been a few isolated case reports of aggressive behaviour in serous cystic neoplasms ('serous cystadenocarcinomas').

Mucinous Cystic Neoplasms

Gross and Microscopic Findings

In contrast to the relatively homogeneous appearance and behaviour of serous cystadenomas, mucinous cystic neoplasms of the pancreas are morphologically and clinically heterogeneous (Albores-Saavedra *et al*., 1987; Solcia *et al*., 1997; Hruban *et al*., 2000). Grossly mucinous cystic neoplasms are large tumours composed of cysts filled with tenacious fluid (mucin) (Compagno and Oertel, 1978b). The cysts are lined by tall, mucin-producing cells (**Figure 8**); expectedly, stains for mucin are positive. In some women a dense layer of spindle-shaped cells (stroma) resembling ovarian stroma surrounds the epithelial cells. The clinical significance of this finding is not clear, but it has been used to suggest a common origin for some neoplasms of the ovary and pancreas.

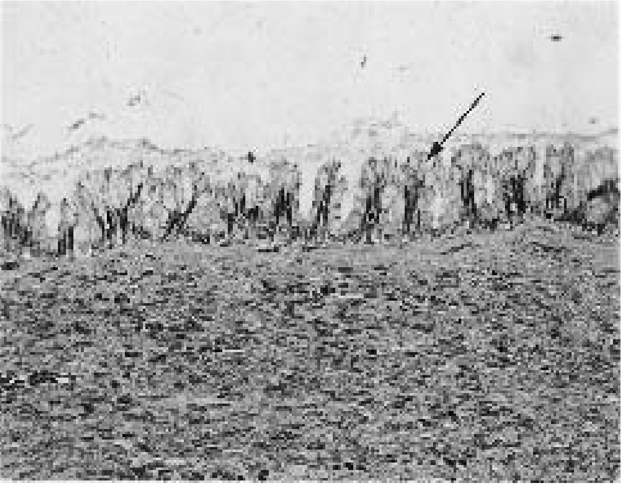


Figure 8 Microscopic section of a mucinous cystic neoplasm. The cysts (empty space at top of photograph) in a mucinous cystadenoma are larger than the cysts found in serous cystadenomas and the cysts in mucinous cystic neoplasms are lined by tall cells producing mucin (arrow). These cells sit on a dense stromal layer that resembles ovarian stroma (bottom portion of the figure).

Mucinous cystic neoplasms can be divided into three groups, mucinous cystadenomas, borderline mucinous cystic neoplasms and mucinous cystadenocarcinomas (Klöppel *et al*., 1996; Solcia *et al*., 1997). Mucinous cystadenomas contain a single layer of cells lacking significant atypia. In borderline mucinous cystic neoplasms, the cells may form finger-like projections (papillae) and complex architectural patterns. The cells in these tumours show significant atypia (e.g. loss of nuclear polarity and

pleomorphism), but no carcinoma is seen. When an *in situ* carcinoma or an invasive carcinoma is present, the diagnosis of a 'mucinous cystadenocarcinoma' should be made. Importantly, otherwise benign-appearing mucinous cystic neoplasms may harbour small foci of invasive carcinoma (Compagno and Oertel, 1978b). Therefore, when possible, mucinous cystic neoplasms should be completely resected surgically, and the surgical pathologist should entirely submit and carefully examine the neoplasm. Failure to do so may explain the occasional reports of metastasizing 'mucinous cystadenomas' (Compagno and Oertel, 1978b; Wilentz *et al*., 1999).

Special Features: Clinical Presentation and Prognosis

Mucinous cystic neoplasms are more common in women than they are in men and the mean age at diagnosis is in the late fifth decade (Compagno and Oertel, 1978b; Solcia *et al*., 1997; Hruban *et al*., 2000). The prognosis for patients with mucinous cystic neoplasms depends on the presence of invasive carcinoma. The authors recently showed that all patients with non-invasive mucinous cystic neoplasms are cured if their tumours are completely resected (Wilentz *et al*., 1999). In addition, while invasive mucinous cystadenocarcinomas are fully malignant tumours, patients with these neoplasms usually live longer than do patients with typical solid infiltrating ductal adenocarcinomas (Wilentz *et al*., 1999). In fact, approximately 50% of patients who have had an invasive mucinous cystadenocarcinoma completely resected will live at least 5 years. This survival rate is much better than the survival rate for infiltrating ductal adenocarcinomas of the pancreas, underscoring the importance of correct pathological classification in patient prognostication and treatment.

Intraductal Papillary Mucinous Neoplasm (IPMN)

Gross and Microscopic Findings

IPMNs are frequently papillary (finger-like) neoplasms that arise in the main pancreatic duct system. By light microscopy, dilated pancreatic ducts are lined by tall, mucin-secreting cells that form papillae (**Figure 9**). Approximately 25% of these tumours are associated with an invasive adenocarcinoma. These invasive cancers often show abundant extracellular mucin production and are called 'colloid' or 'mucinous' adenocarcinomas.

The WHO and the AFIP grading schemes include a three-tiered classification for IPMNs, similar to that for mucinous cystic neoplasms (Klöppel *et al*., 1996; Solcia *et al*., 1997). 'Intraductal papillary mucinous adenomas' are IPMNs without significant cytological or architectural atypia. 'Borderline IPMNs' show a moderate amount of atypia. Finally, 'papillary mucinous carcinoma' is the designation given to those tumours in which the intraductal lesion displays significant cytological and architectural

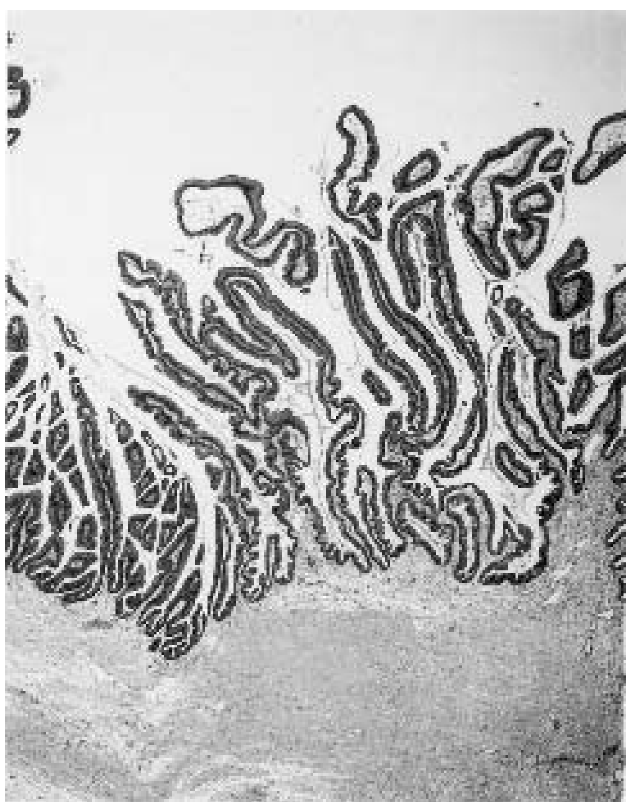


Figure 9 Microscopic section of an intraductal papillary mucinous neoplasm (IPMN). IPMNs are characterized by prominent papillary projections into the dilated pancreatic ducts.

atypia (carcinoma-*in-situ*) or in which an invasive cancer is identified (Klöppel *et al*., 1996; Solcia *et al*., 1997).

Special Features: Clinical Presentation and Prognosis

IPMNs occur with approximately equal frequency in both genders. Their origin in the main pancreatic duct or one of its branches helps distinguish IPMNs from mucinous cystic neoplasms, and it also helps explain why patients with IPMNs are often found to have mucin oozing from the ampulla of Vater, if they are examined endoscopically. The prognosis for patients with IPMNs is probably very similar to that of mucinous cystic neoplasms, but more study is needed (Solcia *et al*., 1997; Hruban *et al*., 2000).

Solid-pseudopapillary Neoplasm

Gross and Microscopic Findings

Solid pseudopapillary neoplasms form large, well-demarcated masses that are cystic, haemorrhagic and necrotic (Solcia *et al*., 1997; Hruban *et al*., 2000). Microscopically the tumour shows solid, cystic and papillary components. The solid areas are composed of nests of small, pink cells

with bland nuclei. The cysts are formed by pools of blood and the papillae usually have vascular cores.

Special Features: Clinical Presentation and Prognosis

Remarkably, almost all solid-pseudopapillary neoplasms of the pancreas occur in women in their 20s (Solcia *et al*., 1997; Hruban *et al*., 2000). Fortunately, most patients with solid pseudopapillary neoplasms survive for many years after surgical resection; however, metastases do occur, and surgeons should try to remove these neoplasms completely.

Endocrine Tumours

Endocrine tumours, also known as islet cell tumours, account for only 1% of all pancreatic neoplasms (Solcia *et al*., 1997; Hruban *et al*., 2000). Unlike their exocrine counterparts, the majority of endocrine neoplasms of the pancreas are not easily divisible into distinct subtypes highly predictive of behaviour (Solcia *et al*., 1997). Sometimes even a combination of gross, microscopic, immunohistochemical and clinical findings cannot definitively predict the overall malignant potential of one of these lesions. The exception is poorly differentiated endocrine carcinoma (high-grade neuroendocrine carcinoma, small-cell carcinoma), which is unequivocally malignant.

Well- and Moderately Differentiated Endocrine Neoplasms (Islet Cell Tumours)

Gross and Microscopic Findings

Well- and moderately differentiated endocrine neoplasms can be benign ('adenoma'), borderline ('neoplasm of uncertain malignant potential') or malignant ('carcinoma'). Grossly all three types of well- and moderately differentiated endocrine neoplasms are usually solid and well circumscribed. They contain uniform cells with granular nuclei (**Figure 10**). The cells can form ribbons, cords, tubules, sheets or nests. The tumours usually stain for the endocrine markers chromogranin, synaptophysin and neuron-specific enolase (NSE). Electron microscopy reveals 100–400-nm neurosecretory granules within the neoplastic cells. These granules are similar to the granules found in normal endocrine cells and they are easily distinguished from the much larger zymogen granules found in acinar cell carcinoma of the pancreas (Solcia *et al*., 1997; Hruban *et al*., 2000).

The best way to determine if a well- or moderately differentiated endocrine neoplasm is benign, borderline or malignant is to look at its behaviour. Microscopic findings such as cytological and architectural atypia are not as helpful. Therefore, well- and moderately differentiated endocrine tumours that show gross extension to other organs, that invade into large blood vessels or that metastasize are prima facie low-grade malignancies. Unfortunately, some endocrine neoplasms that do not show these characteristics later behave in a malignant fashion. There has therefore been a great deal of interest in developing *indirect* markers of malignancy (Solcia *et al*., 1997). When gross local extension, large-vessel invasion or metastasis is not seen, pathologists can apply these indirect markers (size of the tumour, mitotic rate, microscopic invasion of blood vessels and nerves) to assess whether a lesion is benign, borderline or malignant. These markers should be applied with the realization that indirect markers are imperfect markers of malignancy (Solcia *et al*., 1997; Hruban *et al*., 2000).

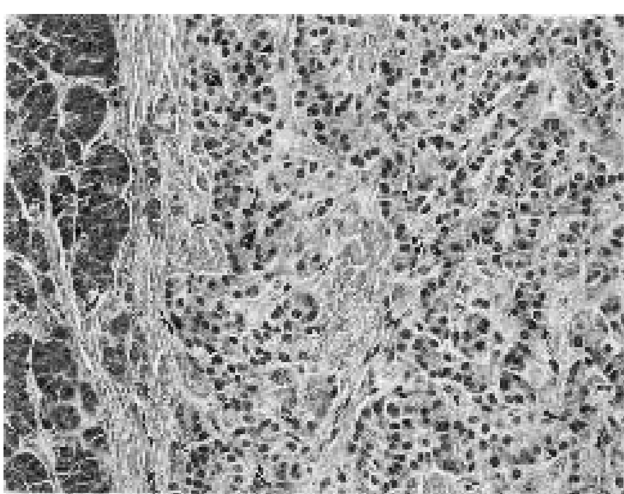


Figure 10 Well-differentiated neuroendocrine neoplasm. In this microscopic section the normal pancreas is to the left and the tumour to the right.

Special Features: Hormone Production

Remarkably, and often dramatically, endocrine neoplasms of the pancreas can produce excessive quantities of many of the same hormones normally produced in small quantities by the islets of Langerhans. Some patients with endocrine neoplasms therefore develop striking symptoms. For example, patients with insulin-producing tumours (insulinomas) can present with marked hypoglycaemia (low blood sugar), headaches, weakness, dizziness and/or seizures. Patients with tumours which produce the hormone gastrin (gastrinomas) can develop the Zollinger–Ellison (ZE) syndrome, characterized by gastric hyperacidity, multiple, recurrent peptic ulcers, gastro-oesophageal reflux and diarrhoea. Tumours that produce VIP (VIPomas) can lead to the Verner–Morrison or WDHA syndrome, which primarily results in watery diarrhoea, hypokalaemia (low potassium levels) and achlorhydria (or hypochlorhydria). Patients with glucagon-producing tumours (glucagonomas) experience a striking, symmetrical skin rash ('necrolytic migratory erythema') on their buttocks, groin, perineum, thighs and distal extremities. Patients with somatostatin-producing tumours (somatostatinomas) present with diabetes mellitus, cholelithiasis (gallstones), diarrhoea, hypochlorhydria, weight loss and anaemia.

Clinical presentation is very important in evaluating a well- or moderately differentiated endocrine neoplasm. Although 10% of glucagon-cell, somatostatin-cell, gastrin-cell and VIP-cell tumours discovered incidentally while the patient was being evaluated for some other problem are malignant, the majority of the *same* tumours are malignant if hormone production by the tumour produces a clinically recognizable syndrome (Solcia *et al*., 1997). The exception is an insulin-producing tumour: only 10% of all insulin-producing neoplasms are malignant, regardless of clinical presentation.

Multiple Endocrine Neoplasia Syndrome, Type 1 (MEN1 Syndrome)

The multiple endocrine neoplasia syndrome, type 1 (MEN1), or Werner's syndrome, deserves special mention. This syndrome is characterized by concomitant multiple hyperplasias and neoplasias of the parathyroid gland (up to 97% of patients), pancreas (80%) and anterior pituitary (50%) (Solcia *et al*., 1997). The MEN1 syndrome can be inherited in an autosomal dominant pattern or it can occur sporadically. The syndrome results from a germ-line mutation or deletion in the *MEN1* gene on the long arm of chromosome 11 (11q13) (Lubensky *et al*., 1996).

As many as 80% of patients with MEN1 develop pancreatic tumours, most frequently gastrinomas, and it therefore should not be surprising that the Zollinger–Ellison syndrome occurs in at least one-third of MEN1 patients (Solcia *et al*., 1997). Insulinomas (approximately 25% of MEN1 patients), VIPomas, glucagonomas and growth hormone-releasing tumours are less common in patients with this syndrome. (See also the chapter *Inherited Predispositions to Cancer*.)

Poorly Differentiated Endocrine Neoplasms (High-grade Neuroendocrine Carcinomas, Small-cell Carcinomas)

Gross and Microscopic Features

Poorly differentiated endocrine neoplasms account for only 2% of all pancreatic endocrine tumours (Solcia *et al*., 1997; Hruban *et al*., 2000). Grossly these tumours are infiltrative, haemorrhagic, necrotic and grey–white. The cells have extremely high nuclear-to-cytoplasmic ratios. Nuclear moulding, where nuclei wrap around one another, is prominent. The mitotic rate is extremely high (Solcia *et al*., 1997).

Special Features: Prognosis

These tumours are unquestionably malignant, and patients with them have extremely poor survival rates.

Primary Mesenchymal Tumours

Benign and malignant mesenchymal (connective tissue) tumours of the pancreas are extremely rare. However, schwannomas (benign tumours of the nerve sheath), leiomyosarcomas (malignant tumours showing smooth muscle differentiation), liposarcomas (malignant tumours showing fat differentiation) and malignant fibrous histiocytomas (malignant tumours showing fibrous and histiocytic differentiation) of the pancreas have been reported. The diagnostic criteria for these neoplasms are the same as they are for mesenchymal tumours arising in other sites. Generally, survival rates for patients with sarcomas are low, as they are for sarcomas primary to other locations. (See also the chapter *Soft Tissues*.)

Metastatic Malignancies

Masses in the pancreas need not have arisen in the pancreas. They may arise in other organs and only later spread (metastasize) to the pancreas. The most common metastases to the pancreas originate in the breast (20%), lung (19%), colorectum (10%), skin (melanoma, 9%) and stomach (7%) (Solcia *et al*., 1997).

Systemic Malignancies

Leukaemia and lymphoma simultaneously involve more than one site, one of which may be the pancreas. In fact, the pancreas is sometimes the presenting site for these malignancies. Of the 67 patients with systemic malignancies involving the pancreas reviewed by Cubilla and Fitzgerald (1976), three-quarters had lymphomas and one-quarter had leukaemias.

EPIDEMIOLOGY AND AETIOLOGY

Because ductal adenocarcinomas are the most common malignancy in the pancreas, and the most is known about ductal adenocarcinomas, the remainder of discussion in this chapter will centre around this type of cancer unless specified otherwise. As outlined in **Table 2**, a number of risk factors have been identified for the development of pancreatic cancer (Gold and Goldin, 1998). These include older age, cigarette smoking, family history of pancreas cancer, previous gastric surgery, chronic pancreatitis, diabetes mellitus, occupational exposure to certain chemicals, radiation exposure and a diet high in fat and low in fruits and vegetables (Gold and Goldin, 1998).

More than 80% of pancreatic cancers occur between the ages of 60 and 80, and cases before the age of 40 are rare (Solcia *et al*., 1997; Gold and Goldin, 1998). Those that do occur at a younger age are usually special variants of pancreatic cancer, such as the pancreaticoblastoma, which occurs primarily in children (Klimstra *et al*., 1995). Pancreatic cancer is more common in men than in women and in blacks than in whites (Solcia *et al*., 1997). Of interest, pancreatic cancer may occur more frequently in individuals of Jewish descent (Gold and Goldin, 1998). As will

Table 2 Risk factors for pancreatic cancer

Patient characteristics
Older age
Male gender
Black race
Jewish ancestry
Family history of pancreatic cancer
Diabetes mellitus
Chronic pancreatitis
Patient behaviour
Cigarette smoking
Diet low in fruits and vegetables
Diet high in fat
Occupational exposure to certain chemicals
Radiation exposure

be discussed in greater detail in the Molecular Genetics section, this may be because of the high prevalence of inherited mutations in the second breast cancer gene (*BRCA2*) in Ashkenazi Jews (Ozcelik *et al*., 1997).

Cigarette smoking has been clearly established as a risk factor for pancreatic cancer. The increased risk of developing pancreatic cancer in smokers is 2–5-fold over non-smokers, and the risk increases with the number of cigarettes smoked. Importantly, those who stop smoking can quickly reduce their risk for developing pancreatic cancer. For example, Ghadirian *et al*. (1991a) conducted a population-based case-control study of pancreatic cancer in Montreal, Canada, and found that, depending on the number of cigarettes already smoked in a lifetime, patients can reduce their risk of pancreatic cancer as much as threefold by stopping smoking. Remarkably, Mulder *et al*., (1999) have estimated that a moderate reduction in smoking in Europe could save as many as 68 000 lives that will otherwise be lost to pancreatic cancer between 1999 and 2020. Clearly, smoking plays a major role in the development of pancreatic cancer.

While smoking is a habit we can control, family cancer history is something we cannot. Researchers at The Johns Hopkins Hospital and others have found that familial inheritance plays a significant role in the development of some pancreatic cancers (Hruban *et al*., 1998). The evidence for this comes from three areas. First, there have been a number of isolated case reports of pancreas cancer aggregating in certain families (Hruban *et al*., 1998). For example, one of former President Carter's parents, his brother and two of his sisters all died from pancreatic cancer. While these reports suggest that there is a genetic (inherited) basis for the development of pancreatic cancer, the occurrence of multiple cancers in a family could be just bad luck. A number of investigators have therefore conducted case-control studies of pancreatic cancer. For example, Ghadirian *et al*. (1991b) conducted a population-based case-control study of pancreatic cancer in the Francophone community of Montreal, Canada, and found that 7.8% of the patients with pancreatic cancer reported a positive family history of pancreatic cancer, compared with only 0.6% of the controls. This 13-fold difference between cases and controls did not appear to be due to environmental factors such as smoking, strongly suggesting the genetic transmission of an increased risk of developing pancreatic cancer. The National Familial Pancreas Tumour Registry (NFPTR) at Johns Hopkins, and other registries like it, were therefore established to track families with multiple pancreatic cancers (Hruban *et al*., 1998). The NFPTR is perhaps the largest of these registries and it is truly an international registry with 719 kindred enrolled as of July 1, 2001. These kindred come from the United States, Europe and Australia, and include 284 families in which two or more first-degree relatives have been diagnosed with pancreatic cancer. This registry provides an invaluable resource to scientists studying the genetics of familial pancreatic cancer (see the Molecular Genetics section). Furthermore, a prospective study of the families enrolled in the NFPTR has demonstrated that the first-degree relatives of patients with familial pancreatic cancer have an increased risk of developing pancreatic cancer themselves. Tersmette *et al*. (2001) studied the families enrolled in the NFPTR. They found that if two family members had been diagnosed with pancreatic cancer at the time of enrollment into the NFPTR, then the risk of other previously healthy family members developing a new pancreatic cancer was 18-fold greater than expected. This risk increased to 56-fold when three or more family members had pancreatic cancer at the time the family enrolled in the registry. The *prospective* development of pancreatic cancer in these families clearly establishes that inherited susceptibility plays a significant role in the development of pancreatic cancer in some families.

Those wishing to learn more about the NFPTR and those wishing to register may contact the National Familial Pancreas Tumour Registry, c/o Dr Ralph Hruban, The Johns Hopkins Hospital, 401 N. Broadway, Baltimore, MD 21231, USA. E-mail: kbrune@jhmi.edu.

The final line of evidence establishing the genetic transmission of an increased risk of developing pancreatic cancer is the identification of some of the genes responsible for familial pancreatic cancer. These genes include *BRCA2*, *p16*, *STK11/LKB1* and *cationic trypsinogen* (Goggins *et al*., 1996; Whitcomb *et al*., 1996; Ozcelik *et al*., 1997; Su *et al*., 1999). These genes will be discussed in greater detail in the Molecular Genetics of Pancreatic Cancer section.

Previous gastric surgery has also been identified as a risk for pancreatic cancer. This association has been extensively studied by Offerhaus *et al*. (1998). They studied several groups of patients, both from the United States and from Europe, and found that patients who had peptic ulcer surgery have a 2–5-fold increased risk of developing pancreatic cancer, especially if the surgery was more than

20 years earlier. The reason for this increased risk is not clear, but it has been hypothesized that increased levels of cholecystokinin and the increased production of nitrosamines in the operated stomach may play a role.

Diet may also contribute to the development of pancreatic cancer (Gold and Goldin, 1998). Diets high in fruits and vegetables have shown to reduce the risk of pancreatic cancer, whereas diets high in fat increase the risk. The mechanism for this has not been established, however, Stolzenberg-Solomon *et al*. (1999) have shown that vitamins such as folate may play a role. They conducted a nested case-control study of a large cohort of male Finnish smokers and found that serum folate and pyridoxal-5′-phosphate concentrations have an inverse dose–response relationship with pancreatic cancer. Additional dietary factors that have been studied include alcohol and coffee consumption; however, there is insufficient evidence to support either as a causal factor in the development of pancreatic cancer (Stolzenberg-Solomon *et al*., 1999).

Finally, two medical conditions, diabetes mellitus and chronic pancreatitis, have also been implicated in the development of pancreatic cancer (Stolzenberg-Solomon *et al*., 1999). Both of these are complex factors to study. Not only have these factors been implicated in the development of pancreatic cancer, but cancer of the pancreas can also destroy normal pancreatic tissues and thus cause both diabetes and pancreatitis. The mechanism by which pancreatic cancer produces diabetes is not well defined; however, some have suggested that amylin production by pancreatic cancer may contribute to the development of diabetes. The increased risk of pancreatic cancer is, however, clear in familial pancreatitis (Whitcomb *et al*., 1996). Familial pancreatitis is caused by inherited mutations in the cationic trysinogen gene and affected family members develop severe recurrent bouts of pancreatitis at a young age. These patients have been shown to have a 40% lifetime risk of developing pancreatic cancer.

SCREENING AND PREVENTION

Screening

Population-based screening programmes have been shown to improve survival of breast, colon and cervical cancer. In contrast, the pancreas is a relatively inaccessible organ and current screening tests for pancreatic cancer are not effective. Nonetheless, there is an enormous need for such a test. Most patients with pancreatic cancer have a dismal prognosis because they do not come to clinical attention until after the disease has spread beyond the pancreas. This need for an effective screening test is perhaps felt most by those with an increased risk for developing pancreatic cancer, such as individuals with an inherited predisposition to develop the disease. These would include individuals with inherited genetic abnormalities in cancer-causing genes, such as *BRCA2, p16* and *STK11/LKB1* and individuals with familial pancreatic cancer of unknown cause (Goggins *et al*., 1996; Hruban *et al*., 1998; Su *et al*., 1999). The importance of screening for pancreatic cancer can be seen in the survival statistics for pancreatic cancer surgery. Patients who have their cancer surgically resected and who are found to have small tumours confined to the gland ('node-negative disease, negative margins and tumours <2 cm') have a 5-year survival of ~25% (Yeo *et al*., 1995). By contrast, most (~85%) patients with pancreatic cancer do not present to their doctor until after their cancers have grown so large that they are inoperable at the time of diagnosis, and patients with inoperable cancer have an average survival of only 6 months.

A great deal of effort is therefore being focused on research efforts for the early detection of pancreatic cancer in asymptomatic high-risk individuals. It is logical that with advances in the molecular genetics of pancreatic cancer and technological advances in endoscopy and radiology, more accurate screening tests for pancreatic cancer will soon become available.

Three groups of diagnostic tests could be applied to screening for pancreatic cancer: (1) radiological methods, (2) tumour markers and (3) endoscopic methods.

Radiological Methods

Radiological methods are the most commonly used methods to diagnose pancreatic cancer. The accuracy of radiology has improved in recent years. Computerized tomography (CT) scanning is commonly the first test used to image the pancreas (Bluemke *et al*., 1995). One recent advance in CT imaging has been the development of spiral or helical CT. This new imaging technique allows the radiologist to obtain higher resolution images in three dimensions. Currently, helical CT and magnetic resonance imaging (MRI) diagnose pancreatic cancer with a sensitivity of ~90%. The ~10% of cancers missed using CT scanning are often those tumours that cause mild diffuse enlargement of the pancreas rather than a discrete mass. In addition, the very small cancers, those less than 1–2 cm in diameter, are often not visible using CT or MRI.

Many pancreatic cancers that are not visualized on CT can be seen using endoscopic ultrasound (EUS). During EUS the endosonographer introduces an endoscope through the patients' mouth, through the stomach and into the duodenum. The tip of this endoscope contains an ultrasound transducer. The pancreas lies adjacent to the stomach and duodenum; this approach therefore allows the endosonographer to obtain close-up images of the pancreas. The main variable determining the quality of the EUS image obtained is the skill and experience of the endosonographer. The sensitivity of endoscopic ultrasound is at least as good as that of helical CT or MRI for

visualizing pancreatic lesions. EUS has an advantage over CT in that the pancreas can also be biopsied using fine needle aspiration (FNA) techniques through the same endoscope, enabling multiple samples to be taken painlessly from any suspicious lesions encountered.

Endoscopic retrograde cholangiopancreatography (ERCP) is a valuable diagnostic and therapeutic tool for managing pancreatic diseases. During ERCP the gastroenterologist again passes an endoscope through the patient's mouth, through the stomach and into the duodenum. Once the tip of the scope is in the duodenum, the endoscopist can visualize the ampulla of Vater, the site where the pancreatic and bile ducts enter the duodenum. A small catheter is then placed through the endoscope and through the ampulla of Vater into the biliary and pancreatic ducts. By injecting radio-opaque dye into the catheter, abnormalities are accurately identified with ERCP. Unfortunately, small lesions in the parenchyma of the pancreas, especially those that have minimal effects on the duct system, can be missed at ERCP, and ERCP is not without risks. ERCP can be complicated by acute pancreatitis (~1/20 procedures), bleeding, complications of sedation, perforation and occasionally even death (< 1/500). The risks associated with undergoing an ERCP therefore preclude its use as a general screening test.

Other imaging modalities under investigation as tests for pancreatic cancer diagnosis include positron emission tomography (PET) scanning and intraductal ultrasound. While the availability of PET is limited, intraductal ultrasound is used to help interpret suspicious findings on ERCP. With the latter procedure, a catheter with an ultrasound probe at its tip is placed into the pancreatic or biliary duct during ERCP. Ultimately these probes may be helpful in identifying very early carcinomas.

In general, radiological imaging of the pancreas is useful in diagnosing pancreatic cancer in a patient for whom there is a high degree of suspicion. The methods are, however, either too expensive or too invasive to be used in the general population as screening tests. (See the section on *Diagnostic Imaging and Image-Guided Intervention*.)

Tumour Markers

Given the limitations of current radiological tools, much effort has been put into identifying molecular markers that have the potential to be sensitive and specific for pancreatic cancer.

An ideal marker would be both highly sensitive (it would correctly identify almost everyone who has pancreatic cancer) and specific (positive tests are only seen in patients with cancer) for pancreatic cancer and it could be applied to samples obtained relatively non-invasively. Although a large number of potential tumour markers have been evaluated, none of them yet have been shown to be sufficiently sensitive or specific for use in a screening setting for pancreatic cancer.

The tumour markers most studied are the carbohydrate antigen 19-9 (CA19-9), K-RAS and telomerase.

CA 19-9

CA19-9 is a valuable tumour marker for following the therapeutic response in patients who are being treated for pancreatic cancer (Ritts and Pitt, 1998). In this setting, CA19-9 levels correlate well with tumour volume and response to therapy. However, CA19-9 is not useful as a screening test for early pancreatic cancer. First, only patients with certain blood types are capable of making CA19-9. Approximately 10–15% of individuals do not secrete CA19-9 because of their Lewis antigen blood type. In addition, CA19-9 levels may be within the normal range while the cancer is still at a small and asymptomatic stage and conversely CA19-9 levels may be elevated in benign biliary or pancreatic conditions. These limitations also apply to related carbohydrate antigens such as CA-125, KAM17.1, CA2.2, CA-50 and CA-242.

K-RAS

K-*RAS* gene mutations are present in ~90% of pancreatic cancers (see Molecular Genetics below) and these mutations can be detected in specimens obtained distant from the cancer such as pancreatic juice, duodenal fluid, stool and blood using sensitive mutation assays (Hruban *et al.*, 1993; Caldas *et al.*, 1994). Unfortunately, a number of limitations preclude the use of K-*RAS* as a screening marker. First, K-*RAS* gene mutations are not specific for pancreatic cancer and also occur in the small non-invasive pancreatic duct lesions (PanINs) that are prevalent in individuals of increasing age. In autopsy series, PanINs can be found in as many as 10–30% of individuals, especially in smokers. Similarly, K-*RAS* gene mutations have also been reported in chronic pancreatitis (Caldas *et al.*, 1994). Clearly, mutant K-*RAS* is not a specific marker of pancreatic cancer. By contrast, circulating mutant K-*RAS* genes in the blood is much more specific for pancreatic cancer and is detectable in blood samples from 25 to 65% of patients with pancreatic cancer. However, the presence of mutant K-*RAS* in the blood may be a late event as it correlates with inoperable pancreatic cancer, with a poor prognosis and relapse after surgical resection.

Telomerase

Telomerase is an exciting new potential marker for pancreatic cancer. Telomerase is an enzyme which helps maintain the ends (telomeres) of chromosomes. Telomerase activity is lost in most normal cells after embryonic development while as many as 90% of cancers and some inflammatory cells express telomerase. Since telomerase is expressed in inflammatory cells, it may not be sufficiently specific for use as a cancer-screening marker. Nonetheless, several groups have reported that as many as 90% of patients with pancreatic cancer have measurable telomerase activity in their pancreatic juice (Suehara *et al.*, 1997),

and telomerase may yet prove to be a relatively sensitive and specific marker for differentiating benign from malignant lesions of the pancreas.

Identifying New Markers

Several novel approaches have been used to identify new markers that might be specific for pancreatic cancer. One of these approaches that is particularly exciting is serial analysis of gene expression (SAGE). SAGE generates a quantitative list of the genes that are expressed by a tissue (Zhang *et al*., 1997). Using SAGE, comparisons can be made between the expression of genes in cancer and normal tissue and a list of genes can then be generated that are highly over-expressed in the cancer relative to the normal (Zhang *et al*., 1997). One marker identified using this approach is tissue inhibitor metalloproteinase 1 (TIMP-1). When used in combination with CA19-9, measurements of TIMP-1 levels in the blood can distinguish patients with pancreatic cancer from controls with greater sensitivity and specificity than CA19-9 alone. Recently, SAGE has been used to discover other additional exciting markers of pancreatic cancer, including prostate stem cell antigen (Argani *et al.*, 2001).

Another powerful technology likely to enhance the prospects for finding cancer-specific markers is gene expression arrays. Gene expression arrays contain arrays of many thousands of genes gridded on to small templates such as a glass slide. Therefore, a tissue or sample can be probed to see if it contains any one of thousands of genes simultaneously using just one slide. It is therefore now easier to obtain gene expression profiles of cells from cancer and these profiles can be compared with the gene expression patterns of non-cancerous cells. For both SAGE and microarrays, complex analytical and statistical software programs are required to interpret complex gene expression data and results require confirmation using other experimental approaches. Nonetheless, both are exciting techniques which my help in the development of novel screening tests for early pancreatic cancer.

DNA Methylation

DNA is frequently methylated in mammalian DNA. Methylation refers to the addition of single carbon groups and methylation occurs at specific sites in DNA called CpG islands. Regions of DNA rich in CpGs are frequently found in the portion of genes which control the expression of the gene (the promoter). When CpG islands are methylated in a promoter of a gene, it can inhibit transcription of that gene by preventing RNA polymerase and the RNA transcription machinery from producing messenger RNA. Hence DNA methylation is a common mechanism for regulating gene expression. Both selective hyper- and hypomethylation of DNA are known to occur in cancer.

Several techniques have been used to screen cancers for methylation abnormalities. DNA methylation changes in cancer can be detected even when they are admixed with many more copies of normal DNA. DNA methylation is therefore being studied as a possible screening tool for the early detection of cancer. For example, DNA methylation of the *p16* gene has been found in the sputum of patients with early lung cancer (Belinsky *et al*., 1998).

Several genes (mostly tumour-suppressor genes) have recently been shown to be selectively hypermethylated in a subset of pancreatic cancers. These genes include *p16* and *hMLH1*.

Prevention

Unfortunately, there are no good published clinical trials for the prevention of pancreatic cancer. As the genetic and environmental factors responsible for pancreatic cancer have become more defined, so to has the need for preventive strategies become apparent.

Several approaches to prevention can be considered. First, general health measures are prudent such as avoiding smoking and alcohol consumption and maintaining a balanced diet. Balanced diets should be low in fat and high in fruits and vegetables. Second, groups with a high risk of developing pancreatic cancer can be enrolled into screening programmes as these programmes become available. Third, chemopreventive strategies should be tested on high-risk populations. For example, nonsteroidal anti-inflammatory drugs (NSAIDs) and COX-2 inhibitors have chemopreventive activity in a variety of animal and clinical studies, and epidemiological studies suggest that NSAIDs protect against colorectal, oesophageal and gastric cancer. Many cancers, including pancreatic cancer, overexpress the enzyme COX-2, the likely target of NSAIDs, and COX-2 inhibitors have been shown to decrease the growth of pancreatic cancers in animal models (Molina *et al*., 1999). Finally, prophylactic pancreatic resection may be appropriate for a very few individuals at very high risk of developing pancreatic cancer (Brentnall *et al*., 1999). Two groups of individuals with the highest risk of developing pancreatic cancer are patients with hereditary pancreatitis (lifetime risk of pancreatic cancer 30–40%) and patients with idiopathic familial pancreatic cancer (individuals with three or more first-degree relatives with pancreatic cancer have a lifetime risk of pancreatic cancer of ~20%). To ensure protection against pancreatic cancer in this setting, prophylactic total pancreatectomy may be indicated, but this is a very high-risk procedure. Total pancreatectomy is associated with significant short-term and long-term morbidity and mortality, including brittle diabetes, and for this reason it is rarely performed.

MOLECULAR GENETICS

The last 10 years have seen a revolution in our understanding of the molecular genetics of pancreatic cancer. In

Table 3 Genes involved in the development of apparently sporadic pancreatic cancer

Gene	% of Cancers	Chromosome	Mechanism of alteration[a]
Oncogenes			
K-*RAS*	90	12p	Point mutation
AKT2	10–20	19q	Amplification
AIB1	65	20q	Amplification
HER/2-neu	70	17q	Overexpression
Tumour-suppressor genes			
p16	95	9p	HD, LOH and IM, PM
p53	50–70	17p	LOH and IM
DPC4	55	18q	HD, LOH and IM
BRCA2	5–10	13q	Germline with LOH
MKK4	4	17p	HD, LOH and IM
LKB1/STK11	5–6	19p	LOH and IM, HD
TGFβR1 and *TGFβR2*	4	9q,3p	HD
DNA mismatch repair genes			
MSH2	<4	2p	Unknown
MLH1	<4	3p	Unknown

[a] HD = homozygous deletion; LOH = loss of heterozygosity; IM = intragenic mutation; PM = promoter hypermethylation.

a few short years, pancreatic cancer has gone from one of the most poorly understood diseases to one of the best. It is now clear that pancreatic cancer is a genetic disease. The genetic alterations which lead to the development of pancreatic cancer can be inherited (see Epidemiology and Aetiology section) or acquired and the genes affected can be classified into three broad groups: oncogenes, tumour-suppressor genes and DNA mismatch repair genes (see **Table 3**).

Oncogenes

Oncogenes are genes which, when *activated* by mutation or overexpression, possess transforming (cancer-causing) properties. The oncogenes which have been shown to play a role in the development of pancreatic cancer include the K-*RAS*, *HER2-neu*, *AKT2*, *AIB1* and *MYB* genes (Hruban *et al*., 1993; 1998; Day *et al*., 1996). The K-*RAS* gene resides on chromosome 12p and it is activated by point mutation in ~90% of the cancers (Hruban *et al*., 1993). The *HER2-neu* gene on chromosome 17q is overexpressed in ~70% of pancreatic cancers and amplification of *AKT2* on chromosome 19q, *AIB1* on 20q and *MYB* on chromosome 6q has been reported in a smaller percentage of the tumours (Day *et al*., 1996). The demonstration that these oncogenes are activated in pancreatic cancer is important for a number of reasons. First, these genes are potential targets for novel therapies. For example, in order for the K-*RAS* gene product to be functional, it must first be activated by the enzyme farnesyl transferase. Several groups have already developed farnesyl transferase inhibitors, some of which may be effective in treating pancreatic cancer. Similarly, mutated K-Ras peptides have been used as vaccines to treat pancreatic cancers. Second, as discussed in the section on tumour markers, activated oncogenes are potential targets for gene-based screening tests for pancreatic cancer. For example, mutant K-*RAS* genes shed from a pancreatic cancer have been detected in pancreatic and duodenal fluids and in the stool of patients with pancreatic cancer (Caldas *et al*., 1994; Brentnall *et al*., 1999). Such gene-based screening tests are exciting because they could potentially detect as few as one mutant copy of a gene admixed with 10 000 normal copies of that same gene (see Screening and Prevention). Third, the patterns of alterations in oncogenes can provide a clue as to what caused the alterations and therefore what caused the cancer. For example, we and others have shown that activating point mutations in K-*RAS* are slightly more common in cancers obtained from smokers then they are in cancers obtained from non-smokers (Hruban *et al*., 1993). This finding can be likened to finding the fingerprint of cigarette smoke in a cancer. It helps establish who the criminal is.

Tumour-suppressor Genes

The second class of genes which are altered in pancreatic cancer are the tumour-suppressor genes. Tumour-suppressor genes are genes which encode for proteins which normally function to restrain cell proliferation, so the *loss* of their activity may lead to unrestrained cell growth. The tumour-suppressor which have been shown to be inactivated in pancreatic cancer include *p16* (in 95% of the cancers), *p53* (in 75%), *DPC4* (in 55%), *BRCA2* (in 10%), *MKK4* (in 4%), *RB1* (in $<5\%$), *LKB1/STK11* (in 4%) and the transforming growth factor β receptor genes I and II (**Table 3**) (Schutte *et al*., 1995, 1997; Goggins *et al*., 1996; Hahn *et al*., 1996; Ozcelik *et al*., 1997; Rozenblum *et al*., 1997; Hruban *et al*., 1998; Wilentz *et al*., 1998; Su *et al*., 1999). Tumour-suppressor genes, like other autosomal

Table 4 Known causes of familial pancreatic cancer

Syndrome[a]	Gene	Chromosome	Familial characteristics
Breast cancer 2	*BRCA2*	13q	Breast cancer and pancreatic cancer
Peutz-Jeghers	*STK11/LKB1*	19p	Pigmented spots on lips
FAMMM	*p16*	9p	Multiple moles, melanoma and pancreatic cancer
HNPCC	Multiple	Multiple	Non-polyposis colon cancer, other cancers
Familial pancreatitis	Cationic trypsinogen	7q	Recurrent episodes of pancreatitis starting at a young age

[a] FAMMM = familial atypical multiple mole melanoma; HNPCC = hereditary non-polyposis colorectal cancer.

genes, are inherited in pairs; a maternal copy and a paternal copy. Both of these copies (called 'alleles') of a tumour-suppressor gene must be inactivated for there to be loss of function of the gene product, i.e. tumour-suppressor genes act as recessive genes. In pancreatic cancer, the inactivation of both alleles of a tumour-suppressor gene occurs by one of three mechanisms: (1) loss of one allele ('loss of heterozygosity') coupled with a mutation within (intragenic) the second allele; (2) loss of both copies of the gene ('homozygous deletions'); and (3) loss of one copy (LOH) coupled with epigenetic inactivation of the second copy (hypermethylation of the gene's promoter) (Schutte *et al*., 1997).

The identification of the tumour-suppressor genes involved in the development of pancreatic cancer is important for a number of reasons. First, the identification of the pathways inactivated when a tumour-suppressor gene is mutated may provide new targets for treatment. For example, inactivation of the *DPC4* tumour-suppressor gene is relatively specific for pancreatic cancer and several potential targets within this pathway have been identified which could be used to develop new chemotherapeutic agents (Hahn *et al*., 1996). Second, the demonstration that specific tumour-suppressor genes are inactivated in pancreatic cancers has proved a critical advance in our understanding of the causes of familial pancreatic cancers (Hruban *et al*., 1998). To understand why this is true, one most go back over 20 years to Alfred Knudson's seminal research on retinoblastoma. He studied childhood eye cancers (retinoblastomas) and hypothesized that the gene that caused sporadic (non-familial) retinoblastoma also caused familial retinoblastoma. In the familial form of the cancer, affected family members inherit one defective copy of the gene. As a result, they have only one good copy of the gene remaining. If this second copy is inactivated (mutated) later in life, then the cancer develops. Cancers are therefore common in these families. This is analogous to going up in the space shuttle with one functioning computer and one broken computer. If something goes wrong with the good one, there is no backup. By contrast, in non-familial (sporadic) forms of the cancers, the patients inherit two good copies of the gene. Only the rare cases in which both copies are inactivated later in life does the cancer develop.

Knudson's hypothesis appears to operate for pancreatic cancer. Kern and colleagues at Johns Hopkins have demonstrated that some forms of familial pancreatic cancer are caused by germ-line (inherited) mutations in tumour-suppressor genes. As summarized in **Table 4**, the genes which are targeted in familial pancreatic cancer include *BRCA2*, *STK11/LKB1* and *p16* (Goggins *et al*., 1996; Su *et al*., 1999). Patients who inherit a defective copy of one of these genes are more likely to develop a cancer later in life because they begin life with only one, instead of the usual two, functional copies of the gene. Should that good copy be inactivated, gene function would be lost. For example, Goggins *et al*. (1996) studied a large series of patients with pancreatic cancer and found that 7% had a germ-line (inherited) mutation in the second breast cancer gene (*BRCA2*). These patients presumably developed their pancreatic cancer when a cell in their pancreas lost the second, only remaining, good copy of the *BRCA2* gene (Goggins *et al*., 1996; Ozcelik *et al*., 1997). Similarly, germ-line mutations in the *p16* gene predispose to both pancreatic cancer and melanoma (Hruban *et al*., 1998); and germ-line mutations in the *STK11/LKB1* gene to pancreatic cancer and a rare syndrome called the 'Peutz–Jeghers' syndrome (Su *et al*., 1999).

The discovery of the genes responsible for some forms of familial pancreatic cancer represents a critical advance, because it means that members of families in which there has been an aggregation of cancer can now be genetically tested. Those found to carry a mutation in one of these genes can be more carefully screened for cancer or may even choose prophylactic surgery, while those found not to carry of a mutation can be relieved of their anxiety.

Mismatch Repair Genes

The final class of genes which play a role in the development of pancreatic cancer are the DNA mismatch repair genes. The products of DNA mismatch repair genes function to ensure the fidelity of DNA replication. Inactivation of DNA mismatch repair genes can therefore be thought of as analogous to having a drunk mechanic inspect your car. The failure of this mechanic to fix problems can have a devastating long-term effect. Remember that every time a cell divides it must copy all three billion DNA base pairs. Errors are made, and if the enzymes which help repair these errors are inactivated, these errors will not be repaired. Over time, this will lead to the accumulation of mutations

in both oncogenes and tumour-suppressor genes and therefore in the development of cancer. The inactivation of a DNA mismatch repair gene in a cancer produces a characteristic change in DNA called 'microsatellite instability', (MSI) and Goggins *et al*. (1998) have recently demonstrated MSI in ~4% of pancreatic cancers. These cancers are remarkable because as discussed in the section on tumour pathology, they appear to have a distinct microscopic appearance ('medullary phenotype') and because these cancers may have a different response to certain chemotherapeutic agents (Goggins *et al*., 1998).

Thus, there has been a revolution in our understanding of the molecular genetics of pancreatic cancer. This understanding is already being applied to the development of new screening tests for pancreatic cancer and new treatments for the disease, and it has led to a better understanding of why pancreatic cancer aggregates in some families.

PROGNOSTIC FACTORS

Overall, pancreatic cancer has the worst prognosis of all the common forms of cancer. Affected patients have a median survival of only ~6 months and less than 5% of patients live to 5 years. Given such dismal statistics, it is not surprising that few markers are available which can identify patients with a good prognosis. Among the subgroup of patients who undergo a surgical resection of the head of the pancreas (Whipple procedure, or pancreaticoduodenectomy), several prognostic factors influence outcome. The size of the carcinoma, the presence or absence of positive margins (cancer extending to where the surgeon cut), the histological grade (how closely it resembles normal tissue under the microscope) and the presence of lymph node metastases predict survival (Yeo *et al*., 1995). In addition, Allison *et al*. (1998) have shown that tumour DNA content (ploidy) also has prognostic significance and, as discussed in the Pathology section, histological classification can be an important predictor of prognosis. Finally, because of their late presentation, pancreatic carcinomas involving the tail and body of the pancreas have a poorer prognosis than cancers of the head of the pancreas.

OVERVIEW OF PRESENT CLINICAL MANAGEMENT

The management of pancreatic cancer depends on several factors, including the patient's symptoms, the performance status of the patient, the histological classification of the patient's tumour, the stage of the disease (whether or not it has spread beyond the gland) and the presence of complications. The diagnosis of pancreatic cancer is usually suspected from complaints of progressive obstructive jaundice (a yellow discoloration of the skin), profound weight loss and pain in the abdomen or mid-back. Less often patients can present with diabetes mellitus, thrombophlebitis migrans (blood clots developing at multiple sites), depression or evidence of metastatic disease. Generally, the diagnosis is established using CT (Bluemke *et al*., 1995), EUS or ERCP with histological (or cytological) confirmation. If curative resection is considered, staging investigations using EUS or angiography are performed, searching for evidence of spread to lymph nodes, to the peritoneum (the lining of the abdomen) or to the liver and for signs of large blood vessel involvement (invasion of the splenic or portal vein). Helical CT with contrast usually provides good assessment of the status of the blood vessels around the pancreas.

Pancreatic Adenocarcinoma

Surgery (Whipple resection) remains the only realistic curative modality for pancreatic cancer (Yeo *et al*., 1995). The Whipple procedure involves the resection of the head of the pancreas, the duodenum, lower common bile duct, local lymph nodes and peripancreatic tissue. This is not an easy operation. Operative mortality rates vary considerably with the experience of the surgeon, but mortality rates in expert centres are an acceptable 2–3%. Unfortunately, even this radical surgery is not curative in most cases. Most individuals who undergo Whipple operation will ultimately die of their disease with a median survival after surgery of ~18 months. Therefore, many patients who undergo a curative resection for pancreatic adenocarcinoma also receive adjuvant (post-operative) chemoradiotherapy (Yeo *et al*., 1995). Some centres are also investigating the potential benefit of neoadjuvant (pre-operative) chemoradiotherapy.

There are few chemotherapeutic agents that are active against pancreatic cancer. Agents such as gemcitabine, taxotere, 5-fluorouracil and others are effective in only 10–20% of patients with the disease. A number of experimental approaches are therefore being tried, including herceptin to target overexpression of the ErbB2 receptor (Day *et al*., 1996), angiogenesis inhibitors and gene therapies which can deliver immunomodulators, prodrugs and tumour-suppressor genes. One of the more novel approaches is a vaccine approach developed by Jaffee *et al.* (2001) at Johns Hopkins, who have developed a pancreatic cancer vaccine that recruits the patients' own immune system to fight the cancer.

It is also important to consider the quality of life for patients with this disease. Pain control is important and can be achieved with the use of opiate analgesia, which can be given in the form of infusion pump. In addition, 'nerve blocks' may be effective in some patients. This procedure involves the destruction of the nerves around the pancreas ('coeliac axis nerve block') and it is achieved by injecting 100% alcohol percutaneously, intraoperatively or during EUS into the nerve bed.

Weight loss is also a common problem for patients with pancreatic cancer. Many patients lose their appetite and the normal taste of food as a result of byproducts (anorectic factors) released from the cancer. In addition, food may not be adequately digested if the pancreas fails to release sufficient pancreatic enzymes owing to pancreatic duct obstruction. Such patients may benefit from taking pancreatic enzyme supplements. However, most patients with pancreatic cancer will lose weight even if they are eating and digesting their food sufficiently. This may be because pancreatic cancers often release cachectic factors (tumour lipid mobilizing factors, proteolysis-inducing factors) which cause muscle and fat breakdown. Such profound weight loss causes weakness and can shorten survival. This cancer-related weight loss can be refractory to treatment, but appetite stimulants may help the anorexia and fish oil supplements appear promising for reversing the cachexia. In the laboratory, fish oils appear to block the effects of cancer on muscle and fat wasting.

Another frequent complication of pancreatic cancer is common bile duct obstruction. Biliary drainage can relieve symptoms of obstruction and it can be achieved with biliary stents introduced during ERCP or percutaneously (percutaneous transhepatic cholangiography (PTC)). Biliary stents can be placed as an outpatient procedure with minimal patient discomfort and they can provide good short-term palliation of symptoms. Unfortunately, such stents frequently block off due to progressive tumour growth.

Pancreatic Neuroendocrine Carcinoma

Patients with islet cell (or neuroendocrine) carcinomas usually present either with symptoms due to hormone hypersecretion or in the context of multiple endocrine neoplasia (MEN) type I or MEN type II. Common presentations are refractory peptic ulcer disease, hypoglycaemia (low blood sugar), carcinoid syndrome (flushing, diarrhoea and asthma), secretory diarrhoea, hypercalcoemia (elevated serum calcium levels) and bone pain and a necrolytic skin rash. These symptoms arise from oversecretion of gastrin, insulin, serotonin (and histamine and other peptides), vasoactive intestinal peptide (VIP), parathyroid hormone or glucagon, respectively. Rarely, other peptides are released, such as somatostatin, pancreatic polypeptide, CRF, GRF or neurotensin. With the appropriate clinical suspicion, diagnosis of the presence of a neuroendocrine carcinoma can be achieved by measuring the levels of the islet cell hormones in the blood. Because they can be small (1 cm or less), multiple and they may involve the duodenum, islet cell tumours can be hard to identify radiologically. EUS is the best test for localizing these tumours. When possible, surgical resection is performed to treat these tumours, but surgery may not be possible if the patient has multiple foci of metastatic disease. Somatostatin analogues are very effective in combating the release of secretory peptides usually observed with these tumours. Patients with gastrinomas usually undergo total gastrectomy, although proton pump inhibitor therapy has reduced the need for this operation. In addition to somatostatin analogues, hyperglycaemic agents such as diazoxide are used to limit hypoglycaemia in patients with inoperable insulinomas.

Finally, chemotherapeutic agents such as streptozocin, 5-fluorouracil and interferon are often used for patients with inoperable islet cell tumours.

CONCLUSIONS

Pancreatic cancer is one of the deadliest of all cancers. The average life expectancy for patients with pancreatic cancer is only 6 months. While there are no effective therapies or screening tests currently available, we believe that the revolution which has occurred in our understanding of the genetics of pancreatic cancer will soon be translated into new effective screening tests, novel treatments and a better understanding of why pancreatic cancer aggregates in some families.

REFERENCES

Albores-Saavedra, J., *et al*. (1987). Mucinous cystadenocarcinoma of the pancreas. Morphologic and immunocytochemical observations. *American Journal of Surgery and Pathology*, **11**, 11–20.

Allison, D. C., *et al*. (1998). DNA content and other factors associated with ten-year survival after resection of pancreatic carcinoma. *Journal of Surgical Oncology*, **67**, 151–159.

Argani, P., *et al*. (2001). Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. *Cancer Research*, **61**, 4320–4324.

Belinsky, S. A., *et al*. (1998). Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proceedings of the National Academy of Sciences of the USA*, **95**, 11891–11896.

Bluemke, D. A., *et al*. (1995). Potentially resectable pancreatic adenocarcinoma: spiral CT assessment with surgical and pathologic correlation. *Radiology*, **197**, 381–385.

Brat, D. J., *et al*. (1998). Progression of pancreatic intraductal neoplasias (high-grade PanIN) to infiltrating adenocarcinoma of the pancreas. *American Journal of Surgery and Pathology*, **22**, 163–169.

Brentnall, T. A., *et al*. (1999). Early diagnosis and treatment of pancreatic dysplasia in patients with a family history of pancreatic cancer. *Annals of Internal Medicine*, **131**, 247–255.

Caldas, C., *et al*. (1994). Detection of K-*ras* mutations in the stool of patients with pancreatic adenocarcinoma and pancreatic ductal hyperplasia. *Cancer Research*, **54**, 3568–3573.

Compagno, J. and Oertel, J. E. (1978a). Microcystic adenomas of the pancreas (glycogen-rich cystadenomas). A clinicopathologic study of 34 cases. *American Journal of Clinical Pathology*, **69**, 289–298.

Compagno, J. and Oertel, J. E. (1978b). Mucinous cystic neoplasms of the pancreas with overt and latent malignancy (cystadenocarcinoma and cystadenoma). A clinicopathologic study of 41 cases. *American Journal of Clinical Pathology*, **69**, 573–580.

Cubilla, A. L. and Fitzgerald, P. J. (1976). Morphological lesions associated with human primary invasive nonendocrine pancreas cancer. *Cancer Research*, **36**, 2690–2698.

Day, J. D., *et al*. (1996). Immunohistochemical evaluation of HER-2/neu oncogene expression in pancreatic adenocarcinoma and pancreatic intraepithelial neoplasms. *Human Pathology*, **27**, 119–124.

DiGiuseppe, J. A., *et al*. (1994a). Detection of K-*ras* mutations in mucinous pancreatic duct hyperplasia from a patient with a family history of pancreatic carcinoma. *American Journal of Pathology*, **144**, 889–895.

DiGiuseppe, J. A., *et al*. (1994b). Overexpression of p53 protein in adenocarcinoma of the pancreas. *American Journal of Clinical Pathology*, **101**, 684–688.

DiGiuseppe, J. A., *et al*. (1996). Molecular biology and the diagnosis and treatment of adenocarcinoma of the pancreas. *Advances in Anatomical Pathology*, **3**, 139–155.

Ghadirian, P., *et al*. (1991a). Tobacco, alcohol, and coffee and cancer of the pancreas. A population-based, case-control study in Quebec, Canada. *Cancer*, **67**, 2664–2670.

Ghadirian, P., *et al*. (1991b). Reported family aggregation of pancreatic cancer within a population-based case-control study in the Francophone community in Montreal, Canada. *International Journal of Pancreatology*, **10**, 183–196.

Goggins, M., *et al*. (1996). Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas. *Cancer Research*, **56**, 5360–5364.

Goggins, M., *et al*. (1998). Pancreatic adenocarcinomas with DNA replication errors (RER+) are associated with wild-type K-ras and characteristic histopathology: poor differentiation, a syncytial growth pattern, and pushing borders suggest RER$^+$. *American Journal of Pathology*, **152**, 1501–1507.

Gold, E. B. and Goldin, S. B. (1998). Epidemiology of and risk factors for pancreatic cancer. *Surgical Oncology Clinics of North America*, **7**, 67–91.

Hahn, S. A., *et al*. (1996). *DPC4*, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*, **271**, 350–353.

Hruban, R. H., *et al*. (1993). K-*ras* oncogene activation in adenocarcinoma of the human pancreas. A study of 82 carcinomas using a combination of mutant-enriched polymerase chain reaction analysis and allele-specific oligonucleotide hybridization. *American Journal of Pathology*, **143**, 545–554.

Hruban, R. H., *et al*. (1998). The genetics of pancreatic cancer: from genes to families. *Surgical Oncology Clinics of North America*, **7**, 1–23.

Hruban, R. H. and Wilentz, R. E. (2000) Pancreas. In: Weidner, N., Cote, R. J., Suster, S. and Weiss, L. M. (eds), *Modern Surgical Pathology*. (W.B. Saunders, Philadelphia).

Jaffee, E. M. (1998). A phase I clinical trial of lethally irradiated allogeneic pancreatic tumor cells transfected with the GM-CSF gene for the treatment of pancreatic adenocarcinoma. *Human Gene Therapy*, **9**, 1951–1971.

Jaffee, E. M., *et al*. (2001). Novel allogeneic granulocyte-macrophage colony-stimulating factor-secreting tumor vaccine for pancreatic cancer: a phase 1 trial of safety and immune activation. *Journal of Clinical Oncology*, **19**, 145–156.

Klöppel, G., *et al*. (1996). *World Health Organization International Histological Classification of Tumors*. (Springer, Berlin).

Klimstra, D. S., *et al*. (1992). Acinar cell carcinoma of the pancreas: a clinicopathologic study of 28 cases. *American Journal of Surgery and Pathology*, **16**, 815–837.

Klimstra, D. S., *et al*. (1995). Pancreatoblastoma. A clinicopathologic study and review of the literature. *American Journal of Surgery and Pathology*, **19**, 1371–1389.

Lubensky, I. A., *et al*. (1996). Allelic deletions on chromosome 11q13 in multiple tumors from individual MEN1 patients. *Cancer Research*, **56**, 5272–5278.

Molina, M. A., *et al*. (1999). Increased cyclooxygenase-2 expression in human pancreatic carcinomas and cell lines: growth inhibition by nonsteroidal anti-inflammatory drugs. *Cancer Research*, **59**, 4356–4362.

Moskaluk, C. A., *et al*. (1997). *p16* and K-*ras* gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. *Cancer Research*, **57**, 2140–2143.

Mulder, I., *et al*. (1999). The impact of smoking on future pancreatic cancer: a computer simulation. *Annals of Oncology*, **10**, S74–S78.

Offerhaus, G. J. A., *et al*. (1988). Gastric, pancreatic and colorectal carcinogenesis following remote peptic ulcer surgery. Review of the literature with the emphasis on risk assessment and underlying mechanism. *Modern Pathology*, **1**, 352–356.

Ozcelik, H., *et al*. (1997). Germline BRCA2 6174delT mutations in Ashkenazi Jewish pancreatic cancer patients. *Nature Genetics*, **16**, 17–18.

Ritts, R. E. and Pitt, H. A. (1998). CA 19-9 in pancreatic cancer. *Surgical Oncology Clinics of North America*, **7**, 93–101.

Rozenblum, E., *et al*. (1997). Tumor-suppressive pathways in pancreatic carcinoma. *Cancer Research*, **57**, 1731–1734.

Schutte, M., *et al*. (1995). Identification by representational difference analysis of a homozygous deletion in pancreatic carcinoma that lies within the BRCA2 region. *Proceedings of the National Academy of Sciences of the USA*, **92**, 5950–5954.

Schutte, M., *et al*. (1997). Abrogation of the *Rb/p16* tumor-suppressive pathway in virtually all pancreatic carcinomas. *Cancer Research*, **57**, 3126–3130.

Solcia, E., *et al*. (1997). *Atlas of tumor pathology: Tumors of the pancreas.* (Armed Forces Institute of Pathology, Washington, DC).

Stolzenberg-Solomon, R. Z., *et al*. (1999). Pancreatic cancer risk and nutrition-related methyl-group availability indicators in

male smokers. *Journal of the National Cancer Institute*, **91**, 535–541.

Su, G. H., *et al*. (1999). Germline and somatic mutations of the *STK11/LKB1* Peutz–Jeghers gene in pancreatic and biliary cancers. *American Journal of Pathology*, **154**, 1835–1840.

Suehara, N., *et al*. (1997). Telomerase activity in pancreatic juice differentiates ductal carcinoma from adenoma and pancreatitis. *Clinical Cancer Research*, **3**, 2479–2483.

Tersmette, A. C., *et al*. (2001). Increased risk of incident pancreatic cancer among first-degree relatives of patients with familial pancreatic cancer. *Clinical Cancer Research*, **7**, 738–744.

Westra, W. H., *et al*. (1998). K-*ras* oncogene mutations in osteoclast-like giant-cell tumors of the pancreas and liver: genetic evidence to support origin from the duct epithelium. *American Journal of Surgery and Pathology*, **22**, 1247–1254.

Whitcomb, D. C., *et al*. (1996). Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nature Genetics*, **14**, 141–145.

Wilentz, R. E., *et al*. (1998). Inactivation of the *p16* (*INK4A*) tumor-suppressor gene in pancreatic duct lesions: loss of intranuclear expression. *Cancer Research*, **58**, 4740–4744.

Wilentz, R. E., *et al*. (1999). Pathologic examination accurately predicts prognosis in mucinous cystic neoplasms of the pancreas. *American Journal of Surgery and Pathology*, **23**, 1320–1327.

Yeo, C. J., *et al*. (1995). Pancreaticoduodenectomy for cancer of the head of the pancreas. 201 patients. *Annals of Surgery*, **221**, 721–733.

Zhang, L., *et al*. (1997). Gene expression profiles in normal and cancer cells. *Science*, **276**, 1268–1272.

FURTHER READING

Hruban, R. H., *et al*. (1997). Pathology of incipient pancreatic cancer. *Annals of Oncology,* **10**, S9–S11.

Hruban, R. H., *et al*. (1997). Pancreatic cancer. In: Vogelstein, B., Kinzler, K. W. (eds), *Genetic Basis of Human Cancer.* (McGraw-Hill, New York) 603–613.

Hruban, R. H., *et al*. (1998). The genetics of pancreatic cancer: from genes to families. *Surgical Oncology/Clinics of North America,* **7**, 1–23.

Hruban, R. H., *et al*. (1998). Tumor-suppressor genes in pancreatic cancer. *Journal of Hepatobiliary and Pancreatic Surgery,* **5**, 383–391.

Hruban, R. H., *et al*. (1999). Screening for pancreatic cancer. In: Kramer, B., *et al*. (eds) *Cancer Screening Theory and Practice.* 441–459 (Marcel Dekker, New York).

Hruban, R. H., *et al*. (1999). Molecular pathology of early pancreatic cancer. In: Shrivastava, D. E., *et al*., (eds), *Molecular Pathology of Early Cancer.* 289–299 (IOS Press, Amsterdam).

Hruban, R. H. and Wilentz, R. E. (2001). Pancreas. In: Weidner, N., *et al*. (eds), *Modern Surgical Pathology.* (W.B. Saunders, Philadelphia).

Solcia, E., *et al*. (1997). *Atlas of Tumor Pathology: Tumors of the Pancreas.* (Armed Forces Institute of Pathology, Washington, DC).

Wilentz, R. and Hruban, R. H. (1998). Pathology of cancer of the pancreas. In: Pitt, H. A. (ed.), W.B. Saunders, Philadelphia, PA. *Oncology Clinics of North America*, **7**, 43–65.

Web Sites

The Johns Hopkins Pancreatic Cancer Web Site:
http://pathology.jhu.edu/pancreas

The Lustgarten Foundation for Pancreatic Cancer Research:
http://www.lustgartenfoundation.org

Pancreatic Cancer Action Network:
http://info@pancan.org

American Cancer Society:
http://www.cancer.org

National Cancer Institutes PDQ for Health Care Professionals:
http://www.oncolink.upenn.edu/pdq_html/1/engl/

Chapter 41

Endocrine Organs

Sylvia L. Asa
University Health Network, Toronto, Canada

Shereen Ezzat
Freeman Centre in Endocrine Oncology, Mount Sinai Hospital, Toronto, Canada

CONTENTS

- Normal Development and Structure
- Tumour Pathology
- Epidemiology
- Aetiology
- Screening and Prevention
- Gross and Histopathology, Immunohistochemistry and Ultrastructure
- Molecular Genetic Findings
- Prognostic Factors
- Overview of Clinical Management

NORMAL DEVELOPMENT AND STRUCTURE

The endocrine organs represent a group of tissues that have, as their primary function, the production and secretion of hormones. They are generally classified into three broad categories:

1. peptide hormone-producing;
2. steroid hormone-producing;
3. thyroid hormone-producing.

Most endocrine cell types fall into the first category. This group of endocrine tissues is composed of cells that have a characteristic neuroendocrine morphology (DeLellis and Tischler, 1998). They have sufficient neural differentiation structurally and functionally that they have been called 'paraneurons' and have been classified as the APUD (amine precursor uptake and decarboxylation) system. In earlier literature it was suggested that they derive from the neural crest embryologically; however, this has not been proved for all members of this group of cells, many of which arise from the primitive endoderm. Nevertheless, functionally they act as neuron-like cells that secrete peptides that are often also produced by neurons. The relationship between these cells and neurons is rather like the comparison between wireless and conventional communication. Neurons produce messengers that are released at synapses and activate receptors in adjacent cells, while neuroendocrine cells produce the same types of messengers that are released into the bloodstream to activate cells throughout the body. The wide array of peptide hormones that they produce is essential for regulation of most metabolic and reproductive functions. These cells are found in classical endocrine organs, such as the pituitary, parathyroid and adrenal medulla, and as members of the dispersed endocrine system scattered within other organs, such as the calcitonin-secreting C cells of the thyroid and the endocrine cells of the lung, gut and pancreas.

The steroid hormone-producing cells are primarily found in the adrenal cortex and the gonads (Sasano, 1998). They also have a distinct morphology that reflects their primary function of conversion of cholesterol into the various mineralocorticoid, glucocorticoid, androgenic and oestrogenic hormones. They are of mesodermal origin arising from the coelomic epithelium that gives rise to the adrenal and the genital ridge.

The thyroid hormone-producing cells are modified epithelial cells derived from the oral endoderm that invaginate from the base of tongue (Murray, 1998). They are specifically involved in the synthesis of thyroglobulin and its iodination to form thyroid hormones.

TUMOUR PATHOLOGY

Tumours of the endocrine system reflect their origin in the three types of endocrine cells.

Tumours of neuroendocrine cells arise either in classical neuroendocrine tissues, such as pituitary,

parathyroid or adrenal medulla, or in other tissues where the dispersed cells reside, such as thyroid, lung, gut or pancreas. These lesions exhibit a wide spectrum of biological behaviours. They may be slowly growing, well-differentiated neoplasms that are considered benign (adenomas), because they do not metastasize. The most aggressive neoplasms are poorly differentiated (small-cell) carcinomas that are rapidly lethal. Many tumours fall into intermediate categories and the prediction of outcome can be very difficult. The term 'carcinoid,' meaning 'carcinoma-like,' was originally introduced by Oberndorfer in 1907, and the terminology has been applied to well-differentiated neuroendocrine tumours and to tumours that result in the classical 'carcinoid syndrome' that results from serotonin excess. The use of this terminology, however, has caused great confusion because of the wide diversity of hormone activity and biological behaviour among these tumours that cannot all be conveyed by this classification. Since many of these ultimately prove to be malignant, this terminology has fallen out of favour. These tumours may be clinically silent in terms of hormone function, but they are almost always found to produce and store hormones. Some elaborate hormones that give rise to colourful clinical syndromes of hormone excess; the pattern of hormone production may be eutopic to the tissue of origin or ectopic, reflecting derepression of genes that are expressed in related cells.

Tumours of steroid hormone-secreting cells usually arise in the adrenals or gonads and very rarely arise in other sites where embryological remnants are found. They are generally classified as benign adenomas or malignant carcinomas based on features of differentiation, hormone production and invasion. Well-differentiated and generally benign tumours express mature steroid hormones. Tumours that are less well differentiated and exhibit malignant behaviour tend to lose the complex enzymatic pathways required for mature hormone production, but often produce hormone precursors of various types. Nevertheless, the functional behaviour of these tumours is not strict enough to allow classification as benign or malignant. These tumours are usually limited to the production of steroid hormones and almost never produce peptide hormones ectopically.

Tumours of thyroid follicular cell derivation are the most common neoplasms of the endocrine system. They include benign follicular adenomas, well-differentiated papillary or follicular carcinomas, poorly differentiated 'insular' carcinomas and dedifferentiated anaplastic carcinomas. Among human malignancies, they include the most benign and nonlethal occult papillary microcarcinomas that are found incidentally in up to 24% of the adult population, and one of the most rapidly lethal malignancies, the anaplastic carcinomas that frequently result in death by strangulation in less than 6 months.

EPIDEMIOLOGY

Tumours of endocrine differentiation are considered to be rare and, as such, the epidemiological data are weak. There are, however, several statistics of note.

Pituitary tumours are reported to be found in about 20% of the general population (Asa, 1998). Many of the studies have reported the identification of these lesions as incidental findings at autopsy, or as radiological findings in the asymptomatic 'normal' population. The true incidence of clinically significant lesions is not known. Some forms of pituitary neoplasia, including corticotroph adenomas causing Cushing's disease and prolactinomas, are more common in women than in men, but overall there is no sex predilection of pituitary neoplasia. These lesions tend to increase with age and are rare in children (Asa, 1998).

Primary hyperparathyroidism is most often due to parathyroid neoplasia and is reported to occur in 1% of the adult population (Apel and Asa, 2001). The true incidence of parathyroid adenomas is not known, however. Parathyroid carcinomas are rare. Benign lesions are more common in women than in men and are primarily found in middle-aged to elderly women. In contrast, carcinoma does not have a predilection for women and some studies indicate onset about one decade earlier than benign parathyroid tumours.

Pheochromocytomas of the adrenal medulla have a reported incidence of 2–8 per million per year and extra-adrenal paragangliomas are even rarer. These lesions have no sex predilection and are rare in children (Lack, 1997; Tischler, 1998).

Well-differentiated tumours of the dispersed endocrine system are rare. Tumors of thyroid C cells, medullary thyroid carcinomas, represent about 5% of thyroid cancers that predicts a prevalence of about 1–2 per 100 000 (LiVolsi, 1990; Moley, 2000). Tumours of the endocrine pancreas have an estimated prevalence of 1 in 100 000 (Klöppel *et al.*, 1998). These lesions show no sex predilection and are very rare in children. Small-cell carcinoma of the lung, the most poorly differentiated endocrine neoplasm of this type, represents one of the four major types of lung cancer, the second most common cancer in men and women and the number one cancer mortality site (Greenlee *et al.*, 2000); this variant has an annual incidence of almost 10 per 100 000 population.

Although adrenal cortical nodules are identified as incidental findings in 0.6–1.3% of asymptomatic individuals, clinically significant adrenal neoplasms are more rare and adrenal cortical carcinoma has an estimated incidence of only about 1 case per million population (Lack, 1997). There is a slight female preponderance. The incidence has a bimodal distribution in the first and fifth decades.

As indicated above, thyroid cancer is the commonest endocrine malignancy, representing 1–2% of all cancers

(LiVolsi, 1990; Murray, 1998; Asa and Bedard, 2000). It is about three times more common in women than in men and currently represents the tenth most common malignancy in women (Greenlee *et al*., 2000).

AETIOLOGY

The aetiology of most endocrine tumours is not known. A small minority is due to inherited genetic defects. The genes responsible for the multiple endocrine neoplasia (MEN) syndromes, *MEN-1* and *MEN-2*, have been cloned and characterized, and the mutations have clarified our understanding of mechanisms of disease. MEN-1 is a classical example of germ-line inheritance of a mutant tumour-suppressor gene (TSG), menin (Komminoth, 1999). It is an autosomal dominant disorder with variable penetrance; the variability of tumour development in pituitary, parathyroids, pancreas and occasionally other sites of the dispersed endocrine system in individual patients is due to the requirement for loss of the intact allele encoding the tumour suppressor. In contrast, MEN-2 is the best example of inheritance of a mutant proto-oncogene. The gene responsible for this disease encodes the transmembrane receptor tyrosine kinase Ret (Mulligan and Ponder, 1995). The identification of an activating Ret mutation in members of kindreds is now accepted as an indication for prophylactic thyroidectomy in early childhood, since these individuals will develop medullary thyroid carcinoma that can metastasize and is lethal in more than half of patients. Moreover, distinct *ret* mutations are associated with distinct clinical phenotypes. Mutations in exons 10 and 11 that encode the extracellular domain of the Ret protein are implicated as the cause of familial medullary thyroid carcinoma alone. Specific mutations, usually in exon 11 involving codon 634, are associated with MEN-2A and specifically codon 634 mutations replacing cysteine with arginine are more often associated with parathyroid disease and pheochromocytoma that characterize this disease complex. Activating mutations in exon 16 that replace a codon 918 methionine with threonine alter the tyrosine kinase domain of Ret and result in MEN-2B, a more aggressive variant of MEN-2 with mucosal neuromas and a marfanoid habitus in addition to tumours of thyroid C cells, parathyroids and adrenal medulla.

The aetiology of thyroid carcinomas of follicular epithelial cells is not entirely known, but there is evidence of a causal role for radiation (LiVolsi, 1990; Murray, 1998; Asa and Bedard, 2000). This is true of radiation therapy, for example in patients who have received external beam radiotherapy for malignancies of the head and neck and also for cosmetic therapy for facial acne. It is also true in populations exposed to radioactive fallout from nuclear disasters, such as in Japan after the nuclear bomb disasters and in Ukraine and Belarus after the Chernobyl episode. The exposure to radioactivity has its highest impact in the young, and the disease is more often multifocal than in sporadic cases; however, the prognosis in patients who have been exposed to radiation does not appear to differ from those with no history of radiation (Brierley and Asa, 2001). Radiation has been implicated as causative of *ret/PTC* gene rearrangements that are thought to play a role in the genesis of thyroid carcinomas (Nikiforova *et al*., 2000). Diet has also been implicated in the development of thyroid cancer; populations with low iodine intake develop goitres and have a higher incidence of follicular carcinoma, but most investigators now recognize dietary iodine insufficiency as the reason for a higher incidence of follicular as opposed to papillary carcinomas. In other words, there is no evidence that it increases the incidence of cancer; rather it alters the morphological variant of well-differentiated thyroid carcinoma.

SCREENING AND PREVENTION

Screening for endocrine tumours has really only been applied systematically in cases of familial disease. Members of kindreds have traditionally been screened using biochemical analysis of hormone hypersecretion. More recently, the addition of genetic information has allowed earlier and definitive identification of carriers in families with known mutations. Patients with MEN-1 have variable disease patterns, so that genetic identification results in continued and careful surveillance of affected family members. Patients with activating mutations of the *ret* proto-oncogene will almost certainly develop medullary thyroid carcinoma, a disease that is lethal if not detected early or prevented, and therefore current guidelines recommend prophylactic thyroidectomy in childhood, usually by age 5 years for those with MEN-2A or FMTC, and at or around age 1 year for those with the more aggressive mutation of MEN-2B. Some families have unidentified mutations and hormonal screening remains the standard mechanism of tumour detection, with the addition of radiological investigation where indicated.

Screening for thyroid tumours is usually part of the physical examination since the thyroid is readily accessible on palpation of the neck. When a thyroid nodule is detected, the most valuable technique to evaluate these lesions is the cytological examination of a fine-needle thyroid aspirate (Murray, 1998; Asa and Bedard, 2000). It can clearly identify some patients who will require surgery and most of those who are unlikely to require surgery. Cytological examination can very quickly impart a diagnosis of *papillary carcinoma*, *medullary carcinoma*, *lymphoma*, *anaplastic carcinoma* or *metastatic carcinoma*. In contrast to these clearly malignant lesions, the majority of thyroid aspirates yield a benign diagnosis, which is either *thyroiditis* or *colloid nodule*, a benign hyperplastic process

with abundant colloid storage. Nevertheless, there is a population of patients whose thyroid aspirates yield abundant follicular epithelial cells, sometimes with atypia; these lesions cannot be classified based on cytology alone and surgery is required to allow thorough analysis of these *follicular lesions* to identify or exclude invasive behaviour that distinguishes benign from malignant neoplasms.

GROSS AND HISTOPATHOLOGY, IMMUNOHISTOCHEMISTRY AND ULTRASTRUCTURE

Tumours of Neuroendocrine Cells

The morphology of endocrine tumours varies with the type and location of the lesion.

Tumours of neuroendocrine cells are generally well-delineated but unencapsulated lesions that have a characteristic histopathology. They are composed of small nests, trabecula or sheets of epithelial cells in a highly vascular stroma (**Figure 1a; see colour plate section**). They occasionally form gland-like structures. The stroma may, in some instances, form amyloid. The tumour cells usually have poorly defined cell borders and abundant cytoplasm that may contain eosinophilic, amphophilic or basophilic granules. Characteristically, the nuclei of tumour cells are bland; nuclear pleomorphism that generally defines malignancy in other epithelial tumours is not a reliable indicator of aggressive behaviour. The more poorly differentiated carcinomas have less cytoplasm, lack granularity and have larger more hyperchromatic nuclei.

These lesions are readily classified by the immunohistochemical localization of common markers of neuroendocrine differentiation (**Table 1**) (Wick, 2000). They almost uniformly stain for synaptophysin, a 38-kDa molecule that is associated with synaptic vesicles of neurons and neuroendocrine cells (**Figure 1b; see colour plate section**). Most contain chromogranins, proteins associated with secretory granules. There are two families of chromogranins, A and B; to classify these lesions appropriately one needs to identify both chromogranins. Moreover, chromogranin immunoreactivity is directly related to the number of secretory granules that may be scarce in poorly differentiated tumours. Other markers of neuroendocrine differentiation include CD57 (Leu7), neural cell adhesion molecule (NCAM; CD56), neuron-specific enolase (NSE) and Protein Gene Product 9.5 (PGP 9.5) that stain variable subpopulations of endocrine lesions and some, such as NCAM and NSE, also stain some non-endocrine tumours. Of course, the structure–function correlations of these lesions are best defined by their immunoreactivity for specific peptide hormones or, in the case of hormones that cannot be localized (such as adrenaline and noradrenaline), the enzymes involved in hormone production (such as tyrosine hydroxylase).

The ultrastructure of these lesions is highly characteristic (**Figure 1c; see colour plate section**). The tumour cells have well developed rough endoplasmic reticulum, reflecting the levels of peptide hormone synthesis, prominent Golgi complexes that are responsible for packaging of hormones for secretion and membrane-bound secretory granules that store hormones for secretion in response to stimulation. The development of these organelles varies with cell differentiation and hormonal activity; the numbers of secretory granules reflect the balance between synthesis, storage and secretion. The morphology of secretory granules is generally reflective of cell type and hormone content, and experts in the field of electron microscopy can classify neuroendocrine cells based on these ultrastructural parameters.

Pituitary tumours (Asa, 1998) and tumours of the adrenal medulla and extraadrenal paraganglia (Lack, 1997; Tischler, 1998) are generally considered benign unless there is evidence of metastatic spread. In contrast, tumours of the dispersed endocrine system, including medullary thyroid carcinomas and endocrine tumours of the lung and gut, are not reliably considered benign (Capella *et al.*, 1995); even in the absence of invasion, there may be delayed recurrence or metastasis. It can be difficult to distinguish tumours of the parathyroid glands from hyperplasia (Apel and Asa, 2001). In this tissue, the diagnosis of malignancy in a neoplasm usually relies on identification of vascular invasion or metastasis.

The current approach to the classification of tumours of the dispersed neuroendocrine system is based on three principles: (i) the diagnosis should be based on the microscopic features of the lesion but must incorporate useful immunohistochemical data; (ii) tumours should be distinguished according to the differentiated cell type and site of origin; and (iii) tumours should be subdivided based on biological behaviour, into benign, low-grade malignant and high-grade malignant lesions; the last include poorly differentiated endocrine carcinoma and the so-called ‘small-cell’ or ‘oat-cell’ carcinomas. The determination of malignant potential is based on architectural and cytological features including invasion as outlined below:

- cytological features;
- ploidy, proliferation markers;
- pattern of hormone production
  eutopic vs ectopic, α-subunit;

Table 1 Diagnosis of neuroendocrine tumours

General	Synaptophysin, chromogranins NSE, CD57, NCAM (CD56), PGP 9.5 Hormones Specific Eutopic Ectopic

- invasion;
  capsular, adjacent tissues, perineural, vascular;
- metastases.

The location and hormone content of individual lesions are of importance, especially in the classification of tumours of histological low-grade malignancy, the 'well-differentiated endocrine carcinoma,' where these additional pieces of information aid in predicting the likelihood of metastatic behaviour.

The wide spectrum of clinical symptomatology associated with these lesions is attributed to the numerous peptide hormones that can be elaborated by the cells of the diffuse endocrine system. A limited example is included in **Table 2** that identifies some of the cell types, hormones and syndromes associated with tumours of the gastroenteropancreatic system.

Tumours of Steroid-Hormone-Producing Cells

Tumours of steroid hormone-producing cells are usually solitary, bright-yellow lesions (**Figure 2a; see colour plate section**) with areas of haemorrhage and necrosis (Lack, 1997; Lack, 1998). They are composed of usually monotonous cells with well-defined cell borders and abundant clear cytoplasm (**Figure 2b; see colour plate section**). Immunohistochemistry does not play a major role in identifying these lesions; however, occasionally it is not possible to confirm steroidogenic differentiation of poorly differentiated tumours on histological grounds alone. These lesions characteristically exhibit nuclear reactivity for the nuclear transcription factor steroidogenic factor-1 and cytoplasmic reactivity for inhibin-α. They also stain for various enzymes involved in the steroidogenic pathways. Ultrastructural examination can be helpful; the cells usually have very well developed smooth endoplasmic reticulum and cytoplasmic lipid droplets (**Figure 2c; see colour plate section**). In addition, the mitochondria have tubulovesicular cristae that are a hallmark of steroidogenic cells.

Nodules in the adrenal cortex are very common and may be hyperplastic or attributable to vascular insufficiency; it can be difficult to distinguish hyperplasia from neoplasia; the criteria for this distinction, although not always applicable, are outlined in **Table 3**. When dealing with unequivocal neoplasms, the diagnosis of malignancy can be very difficult. Cytological atypia is not an indicator of malignancy. Size is an important predictor of biological aggressiveness. Again, however, the distinction of benign from malignant lesions usually depends on the identification of invasion, with extension beyond a capsule into adjacent tissue, vascular invasion or metastasis as the features predicting malignant behaviour. These lesions may be associated with hormone excess syndromes; in some cases, hormonal activity can predict biology. It is, for example, extremely rare for adrenal cortical tumours associated with Conn's syndrome to be malignant. In contrast, adrenal tumours associated with feminization in males are invariably malignant. Most malignancies of these tissues are not associated with hormone excess as their primary manifestation; however, it is not unusual to identify production of steroid hormone precursors by tumour cells when this is examined.

Table 2 Classification of gastroenteropancreatic endocrine cells and tumours

Hormone	Cell type	Clinical syndrome
Insulin	B	Hypoglycaemia
Glucagon/glucagon-like peptides (GLP)	A/L	Diabetes mellitus; skin rash
Somatostatin	D	Somatostatinoma syndrome
Gastrin	G	Zollinger–Ellison syndrome
Pancreatic polypeptide	PP	?
Vasoactive intestinal peptide (VIP)	?	Verner–Morrison syndrome
Secretin	S	(watery diarrhoea hypokalemia achlorhydria (WDHA), pancreatic cholera)
Prostaglandins	?	?WDHA
Serotonin	EC	Carcinoid syndrome
Cholecystokinin (CCK)	I	?
ACTH/MSH; CRH	?	Cushing syndrome
Vasopressin	?	Diabetes insipidus
Growth hormone-releasing hormone (GRH); growth hormone (GH)	?	Acromegaly

Tumours of Thyroid Follicular Cell Derivation

Tumours of thyroid follicular cell derivation are common and are probably the best characterized endocrine neoplasms; however, in this field too there is controversy

Table 3 Hyperplasia vs Neoplasia in Endocrine Tissues

Hyperplasia	Neoplasia
Multiple	Solitary
Poorly encapsulated	Encapsulated
Architectural heterogeneity	Uniform architecture
Cytological heterogeneity	Cytological homogeneity
Comparable areas in adjacent gland	Different from surrounding gland
No compression of surrounding gland	Compresses surrounding gland

concerning diagnostic criteria (LiVolsi, 1990; Rosai *et al.*, 1992; Murray, 1998; Asa and Bedard, 2000).

Follicular Nodules of the Thyroid

Follicular nodules may be hyperplastic, benign follicular adenomas or malignant lesions which include follicular carcinoma and follicular variant papillary carcinoma.

Like in parathyroid and adrenal cortex, *hyperplasia* may be extremely difficult to distinguish from *neoplasia*. As shown in **Table 3**, there are rigid criteria for this distinction; however, in many instances the disease is not as defined as classical teaching has suggested. Generally, hyperplasia is a multifocal disorder in which lesions are poorly encapsulated. The nodules of nodular hyperplasia exhibit architectural as well as cytological heterogeneity and this is truly the hallmark of this disease. In contrast, a follicular neoplasm generally represents a solitary encapsulated lesion that is uniform both in cytology and architecture. Architecturally these lesions may be macrofollicular, in which case they are somewhat difficult to distinguish from hyperplastic nodules; microfollicular lesions are more cellular and worrisome but may also represent part of the spectrum of hyperplasia. Clonality studies have indicated that classical teaching may, in fact, be wrong. Although multinodular hyperplasia is expected to be a polyclonal disease, the dominant nodules in multinodular goitres are often monoclonal, raising serious doubts about our diagnostic criteria. Moreover, the evidence of clonal proliferation in sporadic nodular goitre indicates that the thyroid is a site for the hyperplasia–neoplasia sequence. Nevertheless, clinical experience has shown us that most of these lesions remain entirely benign.

The criteria for the diagnosis of *malignancy in a follicular neoplasm* require the presence of vascular or capsular invasion (**Figure 3a; see colour plate section**). Nuclear and cellular atypia and mitotic figures may be present in adenomas and also in carcinomas and therefore cytological characteristics are not helpful in this regard. Follicular carcinomas cannot, therefore, be distinguished from follicular adenomas by fine-needle aspiration cytology. However, there are additional tools that facilitate this diagnosis; for example, a novel marker of malignancy in these lesions is HBME-1, a monoclonal antibody directed against an unknown epitope that is expressed in malignancy but not benign thyroid follicular epithelial cells. More recently, another marker of follicular carcinoma is the detection of PPARγ that is aberrantly expressed due to a gene rearrangement that places PPARγ under the control of the thyroid transcription factor Pax 8 in follicular carcinoma (Kroll *et al.*, 2000).

Follicular carcinomas exhibit a wide range of biological behaviours that are reflected by morphological criteria. Capsular invasion is usually divided into two groups. *Widely invasive follicular carcinomas*, which are usually identifiable as invasive grossly, and certainly are not difficult to recognize as invasive microscopically, carry a poor prognosis with a 25–45% 10-year survival. In contrast, the more common scenario is that of minimal capsular invasion. This requires very careful and thorough examination of the entire capsule of the follicular neoplasm by the pathologist. *Minimally invasive follicular carcinoma* is identified by total thickness invasion through the capsule superficially into the adjacent parenchyma but not widely beyond the capsule. Borderline lesions include those with invasion into the capsule beyond the bulk of the lesion but not through the full thickness of the capsule or situations in which islands of tumour are trapped within a capsule, associated with perpendicular rupture of collagen. The finding of nests, cords or individual tumour cells within a tumour capsule leads some pathologists to the diagnosis of minimally invasive follicular carcinoma; however, this may represent an artifact in a patient who has undergone fine-needle aspiration biopsy, with trapping by fibrosis or displacement of tumour cells into the capsule. The pathologist is therefore advised to search carefully for evidence of fine-needle aspiration biopsy in the adjacent tissue. This would include finding focal haemorrhage, deposition of haemosiderin-laden macrophages and the presence of granulation tissue and/or fibrosis, all of which would indicate a needle biopsy site and the possibility of artifactual invasion rather than genuine invasion.

The careful search for minimally invasive carcinoma is time consuming and difficult for even the most diligent pathologist. Recent data suggest that this search may be unnecessary, since it would appear that patients with minimally invasive carcinoma have almost a 100% 10-year survival rates and therefore some argue that this disease does not need to be distinguished from follicular adenoma. Nevertheless, the investigators who have reported these data have treated their patients for carcinoma rather than for benign disease. Rather than endorsing a cavalier approach that would entail less work for the pathologist, it behoves us to recognize the presence of potential malignancy, to treat the patient appropriately, but also to identify the excellent prognosis that these lesions carry after appropriate management.

Vasculoinvasive follicular carcinomas are aggressive and require management accordingly.

The last few decades have seen a decrease in the incidence of follicular thyroid carcinoma, probably due to dietary iodine supplementation. However, misdiagnosis of this tumour continues. Benign lesions, such as partly encapsulated hyperplastic nodules or nodules exhibiting pseudoinvasion after fine-needle aspiration, are often overdiagnosed as malignant; papillary carcinomas with follicular architecture are often misinterpreted as follicular carcinoma. The clinical features, pathophysiology, and biological behaviour of follicular cancer differ significantly from those of the entities with which it is often

confused. Only careful histopathological classification will allow correct evaluation of treatment options and prognosis.

Papillary Lesions

Papillary nodules of the thyroid are usually malignant but occasional benign papillary tumours are identified.

Papillary Hyperplasia

Papillary lesions are seen focally in *nodular goitre* with or without associated hyperfunction; 'hot' nodules are usually associated with increased uptake on radionucleotide scan.

The so-called '*papillary hyperplastic nodule*' most likely represents a benign neoplasm of the thyroid. These are said to occur most commonly in teenage girls. They present as solitary nodules and may be associated with clinical hyperfunction. *Benign papillary neoplasms* in adults may result in clinical toxicity and a 'hot' nodule on radionuclide scanning. These lesions are encapsulated, often show central cystic change and have subfollicle formation in the centres of broad oedematous papillae. They are distinguished from papillary carcinoma by lack of nuclear atypia (see below).

Papillary Carcinoma

Papillary carcinoma comprises more than 80% of thyroid epithelial malignancies in countries where goitre is not endemic. The name 'papillary carcinoma' is historic and often is misleading. In fact, the architecture of these neoplasms varies from an almost pure papillary pattern to a pure follicular pattern; many tumours have a mixed papillary and follicular pattern. It is now recognized that the diagnosis of papillary carcinoma is based on cytological criteria, as specified in the WHO classification on thyroid tumours, 'a distinctive set of nuclear characteristics' that can be listed as follows:

- crowded, overlapping 'shingle-tile' appearance;
- large, elongated nuclei;
- irregular nuclear outline;
- etched, folded or moulded nuclear membrane;
- nuclear grooves;
- pale vacuolated nucleoplasm;
- peripheral margination of chromatin;
- bare or marginated nucleoli; and
- nuclear cytoplasmic pseudoinclusions.

These characteristics are illustrated in **Figure 3b; see colour plate section**. No one specific feature is absolutely diagnostic of papillary carcinoma; usually, one relies on a constellation or combination of nuclear features for the diagnosis. These nuclear features may be accompanied by more easily recognized psammoma bodies and the presence of high molecular weight cytokeratins, which are readily localized in formalin-fixed paraffin-embedded tissue when microwave antigen retrieval is applied. Identification of high molecular weight cytokeratins confirms the suspected diagnosis of papillary carcinoma in approximately 60% of cases. Other markers of papillary carcinoma include HBME-1 (see Follicular Nodules of the Thyroid, above) and Ret. The Ret tyrosine kinase, that is expressed in thyroid C cells and exhibits somatic activating mutations in medullary thyroid carcinomas and germ-line activating mutations in MEN-2, is usually not expressed in thyroid follicular epithelial cells. However, there are several variants of gene rearrangements, known as *ret/PTC*, that result in expression of the C-terminus of Ret in papillary carcinoma. Detection of such a rearrangement by immunostaining for Ret can facilitate diagnosis in difficult cases.

Thyroid tumours exhibiting the nuclear characteristics of papillary carcinoma diffusely or multifocally should be diagnosed as papillary carcinoma rather than as follicular carcinoma. These lesions share certain clinical characteristics such as biological indolence and an excellent prognosis with a 20-year survival rate of 95% or better. Papillary carcinomas invade lymphatics, leading to a high percentage of regional lymph node metastases. Metastases beyond the neck are unusual in common papillary carcinoma and probably only occur in about 5–7% of cases.

Aggressive Tumours of Follicular Cell Derivation

Specific morphological features can identify tumours that are more aggressive than the usual well-differentiated follicular or papillary malignancies derived from follicular epithelial cells. For example, the *tall cell variant of papillary carcinoma* is recognized to behave in a much more aggressive fashion. These tumours have a height to width ratio that exceeds 3:1. Tumours that exhibit this feature in more than 30% of the tumour mass generally are found in older patients who have extrathyroidal extension and these patients have a guarded prognosis.

The identification of insular dedifferentiation marks a tumour that requires much more aggressive management. *Insular carcinoma*, also known as *poorly differentiated carcinoma*, is a tumour that exhibits a behaviour intermediate between well differentiated thyroid cancer and anaplastic carcinoma. These lesions are identified by their architectural growth pattern; they form solid nests of epithelial cells that resemble neuroendocrine carcinoma more than a follicular lesion (**Figure 3c; see colour plate section**), however, they generally contain thyroglobulin immunoreactivity to characterize their differentiation.

Anaplastic thyroid carcinomas are composed of undifferentiated cells that may exhibit three general patterns but most tumours manifest mixed morphology. The most common type is the *giant cell variant*; as the name suggests, these tumours are composed predominantly of large cells with abundant cytoplasm and bizarre, often multiple,

hyperchromatic nuclei. The *squamoid variant* is composed of large cells that form nests, resembling squamous carcinoma. *Spindle cell anaplastic carcinomas* have a fascicular architecture and dense stromal collagen with spindle-shaped tumour that may resemble fibrosarcoma. In all three variants, mitotic figures and atypical mitoses are frequent. There is usually extensive necrosis and, in some cases, necrosis may be so extensive that the only viable tumour is around blood vessels. Anaplastic carcinomas are highly infiltrative, destroying thyroid tissue and invading skeletal muscle, adipose tissue and other perithyroidal structures. Blood vessel invasion and thrombosis with or without tumour cell involvement are frequent. These lesions usually have no immunoreactivity for markers of thyroid cells and are often a diagnosis of exclusion.

The reported association between well-differentiated thyroid carcinoma and anaplastic carcinoma ranges from 7 to 89% of cases; however, the lower figures are probably underestimates, attributable to inadequate sampling. The data suggest that anaplastic carcinoma originates most often in an abnormal thyroid; the tumour has a higher incidence in regions of endemic goitre and a history of goitre is reported in over 80% of cases. As stated above, nodular goitre is often the site of monoclonal proliferation, the first step in the hyperplasia–neoplasia sequence. However, it is difficult to document transformation of a benign lesion to a malignant tumour. Insular carcinoma appears to be intermediate in the spectrum, and may represent a transition form. The association of papillary carcinoma, particularly the more aggressive tall cell variant, with anaplastic tumours has also been described. The factors underlying dedifferentiation in thyroid tumours remain to be established; age and radiation have been implicated. Clearly, most well-differentiated thyroid lesions do not undergo such transformation. A pattern of genetic mutations resulting in oncogene activation or loss of tumour suppressor gene activity has been proposed to correlate with the stepwise progression from adenoma to carcinoma and through the dedifferentiation process in thyroid. The significance of microscopic anaplastic change is controversial; some people have suggested that focal microscopic anaplastic dedifferentiation does not alter prognosis, but others have shown that this finding alone is statistically significant as a marker of aggressive behaviour.

MOLECULAR GENETIC FINDINGS

The technique of clonality assessment using X chromosome inactivation patterns has evolved from the Lyon hypothesis, which states that only one X chromosome is active in any female somatic cell; the inactivation occurs early in embryogenesis and persists throughout the lifespan of the cell and its progeny. Traditionally, X chromosome inactivation was determined by the phenotypic expression of isoenzymes of glucose-6-phosphate dehydrogenase (G6PD), a protein encoded on that chromosome. Heterozygosity for this gene, however, is present in only a small proportion of black females, limiting the application of this technique. A molecular approach to the determination of clonality takes advantage of X chromosome inactivation patterns using DNA restriction fragment length polymorphisms (RFLPs). Activated genes can generally be distinguished from their inactive counterparts because of differences in the degree of methylation of cytosine (C) residues which are typically hypomethylated in active genes. X chromosome genes which have been utilized for these studies include hypoxanthine phosphoribosyltransferase (HPRT), phosphoglycerate kinase (PGK) and M27β. Molecular analyses have proved that most endocrine neoplasms are monoclonal.

Although the genetic basis of the inherited endocrine tumours is now understood, the genetic abnormalities underlying the far more common sporadic tumours are not clear (Asa and Ezzat, 1998). In fact, mutations of the common oncogenes and tumour-suppressor genes that have been implicated in other malignancies, such as *ras* and *p53*, are rare in endocrine tumours (Ezzat and Asa, 1998).

Proto-oncogenes are normal cellular genes that play an essential role in the proliferation and differentiation of normal cells. They function at each step of signal transduction pathways as growth factors (e.g. c-*sis*), membrane receptors (e.g. c-*erb*-B, c-*neu*, c-*fms*), GTP-binding proteins (e.g. *ras* family) and nuclear proteins (e.g. c-*myc*, c-*fos*). Proto-oncogenes may be activated by point mutations, translocations or increased expression. Genetic alteration in these genes leads to sustained activation of the gene product in the absence of the normal control mechanisms. Activated oncogenes have been associated with a large number of human tumours, e.g. N-*myc* amplification in neuroblastomas and K-*ras* point mutations in colonic carcinomas.

Ras proteins are involved in transducing signals from the cell surface to a number of ligand–receptor complexes. The commonest mutational sites alter the GTP-binding domain (codons 12/13) or more rarely the GTPase domain (codon 61). Point mutations of all three *ras* genes (H-, K- and N-) are rare in endocrine tumours of all kinds.

Although activating mutations of the *ret* proto-oncogene have been implicated as the cause of familial medullary thyroid carcinoma and MEN-2, activating *ret* mutations are not frequently identified as somatic events in sporadic tumours of the dispersed neuroendocrine system. They are found in a minority of sporadic medullary thyroid carcinomas and pheochromocytomas.

The only endocrine oncogene that has been proved in any consistent fashion in sporadic endocrine tumours is the α-subunit of the Gs protein (Lyons *et al*., 1990). G-proteins are heterotrimeric membrane-anchored peptides that play a

central role in transducing signals from the cell surface ligand–receptor complexes to the downstream effectors. The α-subunit dissociates from the β- and γ-subunits of Gs when GTP displaces its bound GDP, stimulates adenylyl cyclase to produce cyclic AMP from ATP. Cyclic AMP (cAMP) in turn activates c-AMP-dependent protein kinases, increases intracellular calcium transport and may potentiate the effect of activated inositol phospholipid-dependent protein kinases. The weak intrinsic GTPase activity of Gsα and the action of GTPase-activating peptides (GAPs) dissociates GTP from Gsα and terminates the response. Additionally, the multiple structural and functional isoforms of adenylyl cyclase underscore the complexity of this redundant system of signal transduction coupling and provides some insight into the array of potential sites of somatic mutations which could alter both cell division and hormone production. Indeed, one of the earlier and most exciting molecular defects to be described in endocrine tumours involved the single point mutations in two critical domains of the Gsα subunit of GTPase codon 201 where Arg is switched to a Cys or codon 227 where Gln is replaced with Arg. Substitutions at these codons (the *gsp* mutation) activate adenylyl cyclase by inhibiting the hydrolysis of GTP and thereby maintaining Gsα in a constitutively activated state. Activating mutations of this protein are reported to occur in 30% of pituitary growth hormone-producing adenomas (Spada *et al*., 1992) and in a large proportion of hyperfunctioning thyroid adenomas (Suarez *et al*., 1991; Goretzki *et al*., 1992). Marked elevation in *Gs*α mRNA levels have been documented in insulin-secreting pancreatic endocrine tumorus. In all of these situations, they predict benign but endocrinologically hypersecretory lesions.

Over-expression of genes which act as inhibitors of the apoptotic process may also act as oncogenes. In some studies, the oncogenes c-*myc*, *bcl-2*, c-*erb*-B-2 and c-*jun* have been shown to be frequently expressed in human gastroenteropancreatic endocrine tumours. The expression of these oncogenes may represent pathogenic events in the generation, malignant transformation and progression of gastroenteropancreatic endocrine tumours.

Products of tumour-suppressor genes (TSGs) can alter transcription factor activity, thereby modulating physiological growth by arresting cell division in the G_1 phase. This delay may allow for repair of genomic damage or may trigger apoptotic cell death. Deletion or reduced expression of TSGs appear to be a commonly shared mechanism in human tumorigenesis.

The p53 protein plays a role in cell cycle regulation; point mutations, deletions or rearrangements in the *p53* gene which result in an altered protein are considered to be among the commonest genetic mutations in human neoplasms and have been implicated in tumour progression in several types of cancer. Progressive transformation to the malignant phenotype may be the result of mutational inactivation of the *p53* TSG. Indeed, p53 mutations appear to play a role in the pathogenesis of some tumours of the dispersed endocrine system, including those arising from the appendix, and carcinomas of the parathyroids and adrenal cortex. They are a late event in thyroid cancer where they are found primarily in anaplastic carcinomas.

In patients with multiple endocrine neoplasia type 1, loss of heterozygosity (LOH) in tumours has been mapped to 11q13. The *MEN-1* tumour-suppressor gene was recently cloned at this chromosomal site. Mutations of menin are not found in most sporadic tumours of the tissues involved in MEN-1: pituitary, parathyroids and pancreas (Komminoth, 1999). *MEN-1* gene mutations have been found in some sporadic tumours of the dispersed endocrine system, primarily gastroenteropancreatic endocrine tumours including 44% of sporadic gastrinomas and 19% of insulinomas, but also in lung endocrine tumours. This region of 11q13 contains several other genes that are also known to be associated with tumorigenesis and may be implicated in the development of tumours that show a high frequency of LOH in this region.

The retinoblastoma (*Rb*) gene has been implicated in the pathogenesis of parathyroid carcinomas (Apel and Asa, 2001). Although animal models suggest that it should be important in pituitary tumorigenesis, there is no evidence for *Rb* loss or alteration in human pituitary adenomas (Asa and Ezzat, 1998). A small number of case reports have emerged describing deletions and possible rearrangements of the *Rb* gene in the very rare insulin-producing carcinomas. Deletions of the putative Wilms tumour and the transformation suppressor gene k-*rev*-1 have also been described in an insulinoma.

The gene conferring predisposition to familial adenomatous polyposis coli (FAP) has been identified (*APC*) and mapped to chromosome 5q21. Somatic mutations in *APC* have also been identified in sporadic pancreatic, colorectal and gastric carcinoma. This gene may play a role in the development of thyroid carcinomas in affected members of families with FAP; however, this remains unproved (Soravia *et al*., 1999). There is no evidence for a pathogenetic role for APC in sporadic thyroid tumours.

Other genomic mutations that have been described in human endocrine tumorigenesis include chromosome 10 or 17 monosomy, and loss of the adhesion molecule DCC (deleted in colonic carcinoma).

Activating mutations of receptors that regulate hormone synthesis and secretion, long anticipated as the molecular solution to the problem of endocrine tumorigenesis, have been likewise disappointing (Ezzat and Asa, 1998). Only in some cases of hyperfunctioning thyroid adenomas have such activating mutations, in this case of the thyrotropin receptor (Porcellini *et al*., 1994; van Sande *et al*., 1995; Krohn *et al*., 1998), proved to be associated with disease.

The molecular genetics of thyroid carcinoma are a model for understanding the role of genetic mutations in tumorigenesis and the application of that knowledge to diagnosis. The unique chromosomal rearrangements

involving the *ret* proto-oncogene (*ret/PTC* gene rearrangements) are found in papillary carcinomas; however, these are early events, found most frequently in occult papillary microcarcinomas that are biologically insignificant lesions. Follicular carcinomas are now thought to involve another novel gene rearrangement involving the thyroid transcription factor Pax-8 and the *PPAR-*γ gene. However, the factors that predict growth and metastasis in these common well-differentiated malignancies are unknown and therefore the rational management of patients remains unscientific. Ras mutations are rare but their impact on biological behaviour is controversial; they may be involved in tumour progression. p53 mutations are late events that have been described only in anaplastic carcinomas.

The data indicate that endocrine tumorigenesis is the result of novel genetic events. The conventional approach of screening tumours for known oncogenes and TSGs has been attempted but with little success. Clearly, these tumours have the potential to clarify new mechanisms of tumour development.

PROGNOSTIC FACTORS

Prognostic factors in neuroendocrine tumours are generally extent of disease and sometimes involve patterns of hormone production. Some genetic factors are also predictive; for example, familial medullary thyroid carcinoma (FMTC), with or without the association of MEN-2A syndrome, has a better survival than that associated with MEN-2B syndrome and has been associated with better survival than sporadic medullary thyroid carcinoma. However, in patients matched for age, extent of tumour and lymph node involvement, survival is similar for those with the hereditary and sporadic medullary thyroid carcinomas. Any differences in survival of hereditary and sporadic cases may be due to earlier diagnosis in high-risk patients that are screened for hereditary medullary thyroid cancers. Although presenting at an earlier age, patients with MEN-2B have more advanced disease and poorer survival than those with MEN-2A. Younger age and female gender are generally reported as favourable prognostic indicators in this disease. The presence of lymph node involvement affects survival adversely, as does extension through the thyroid capsule. The most important predictive factor for survival is biochemical cure, measuring calcitonin after surgery. However, a decrease in the calcitonin level may indicate progression to a poorly differentiated tumour. Carcinoembryonic antigen (CEA) has also been used as a marker for disease progression in medullary thyroid carcinoma; a short CEA doubling time is associated with rapidly progressive disease.

Concerning the prognosis of tumours of steroid hormone-secreting cells, adrenocortical adenomas generally take longer to diagnose than carcinomatous counterparts. These adenomas usually cause Cushing syndrome, but can also result in virilization, Conn syndrome, or no endocrinological symptoms. 'Nonfunctional' tumours probably merely secreted insufficient steroids to cause signs and symptoms. Unfortunately, small tumour size and 'benign' histological features are not sufficient predictors of clinical behaviour. Among malignant tumours, a high proliferative index as identified by MIB-1 or Ki-67 labelling is thought to predict shorter disease-free survival.

The most useful prognostic markers in well-differentiated carcinomas of thyroid follicular epithelium are patient variables, tumour size and extent of disease. Age is the single most important prognostic factor. Patients under the age of 45 years usually have an excellent prognosis; in contrast, those over 45 years of age generally have a poorer outlook. Sex has also been said in the past to be an important determinant of tumour biology but more recent studies have suggested that there is no major difference in the behaviour of these carcinomas in men and women. Tumour size is exceedingly important. Tumours smaller than 1 cm are common and appear to be different biologically than larger tumours; a recent study has shown that occult papillary carcinomas are identified in up to 24% of the population in thyroids that are removed for nonmalignant or unrelated disease. In contrast, tumours larger than 1 cm are thought to be of clinical significance and those larger than 3 cm generally have a poorer prognosis than do smaller tumours. The presence of cervical lymph node metastases, whether microscopic or identified clinically, is thought to increase the risk of recurrence of disease but has been shown to have no impact on mortality. Extrathyroidal extension, in contrast, predicts a worse prognosis and the presence of distant metastases is the hallmark of an aggressive tumour that will bear the potential for high mortality. In patients who have metastatic disease, the site of metastases, the size of metastases and the ability to take up radioiodine are important factors. The value of novel molecular markers in determining the prognosis of differentiated thyroid cancer remains to be established.

OVERVIEW OF CLINICAL MANAGEMENT

In most instances, the diagnosis of neuroendocrine tumours is based on the identification of hormone excess or radiographic identification of a mass lesion. Some neuroendocrine tumours can be difficult to localize and visualize. They have in common, however, the expression of somatostatin receptors as a distinguishing feature. While the subtype of the five distinct receptors varies among different endocrine tumours, they all bind somatostatin. This feature has provided a novel tool for the imaging of these lesions and their metastases, using a radiolabelled somatostatin analogue, [^{111}In]octreotide. It allows the diagnosis of neuroendocrine tumours and localization of occult metastases or, in some patients with metastatic disease, an occult primary lesion.

Surgery remains the treatment of choice for tumours of the dispersed endocrine system. In patients with localized disease, it can be curative.

In many cases, advanced neuroendocrine neoplasms follow an indolent course. Hepatic metastases are common, and although they can cause significant pain or, in some patients, incapacitating hormone hypersecretion, hepatic metastases are usually asymptomatic. The appropriate timing and efficacy of interventions, including hepatic artery embolization (HAE) and/or cytoreductive surgery remain controversial. While some studies have indicated that earlier resection of the primary tumour was associated with prolonged survival, those with liver involvement by tumour were least likely to benefit from surgical resection. Five-year survival rates range from 40 to 80%. Both HAE and surgical resection provide excellent palliation of hormonal and pain symptoms. In selected patients, surgical resection of hepatic metastases may prolong survival, but this is rarely curative.

The expression of somatostatin receptors by these lesions provides a therapeutic tool. Somatostatin tends to inhibit hormone synthesis and secretion, and there is evidence that it may also inhibit cell proliferation. Native somatostatin, however, has a short half-life and cannot be used therapeutically. Several years ago, a long-acting analogue, octreotide, became available for clinical use as a therapeutic tool; this agent required daily subcutaneous injection. More recently, a longer acting repeatable (LAR) preparation of the somatostatin analogue has been developed with activity lasting nearly 28 days. Its therapeutic efficacy, tolerability, and safety in patients with pituitary and gastroenteropancreatic neuroendocrine tumours have been established. In patients with acromegaly due to pituitary growth hormone excess and in some with functional carcinoid tumour, Zollinger–Ellison syndrome and glucagonoma syndrome the analogue normalizes hormone levels and reduces symptoms. Tumour size does not change significantly and may increase in the rare patient on long-term therapy. The side-effect profile includes transient gastrointestinal upset. Longer term complications include possible worsening of glucose intolerance into frank diabetes mellitus and/or gallstone formation. Another long-acting somatostatin analogue, lanreotide, has also been evaluated, and when compared with the older subcutaenous octreotide preparation, both agents appear to be equally efficacious in terms of symptom control. Thus, it appears that longer acting somatostatin analogues have good therapeutic efficacy, tolerability and safety in the treatment of neuroendocrine tumours. The current analogues, however, lack tumoricidal properties, which limits their overall impact on disease burden. The decision to incorporate this therapy in patients with symptoms related to neuropeptide hyersecretion is usually straightforward. The role of these agents alone or as part of chemotherapeutic agents in patients without neuropeptide hyersecretion symptoms is far less clear.

It has also been suggested that high activities of a radiolabelled somatostatin analogue may have a radiotherapeutic effect. In one recent study, patients with known disseminated neuroendocrine tumours were administered between 1.3 and 4.6 GBq of [^{111}In]octreotide for up to five doses over a 1-year period. The treatments were well tolerated. Further work with this and other radioisotopes including yttrium are now being performed to assess efficacy.

Neuroendocrine tumours are generally very resistant to chemotherapy and therefore chemotherapy is rarely recommended for control or palliation. A larger number of chemotherapeutic agents have been used as single agents or in combination. Streptozotocin (STZ), or glycosamine nitrosourea originally derived from streptomyces has been in use for three decades since initially being found to be active in pancreatic endocrine tumours. Current schedules in common use are the combination of STZ with doxorubicin or 5-fluorouracil (5-FU). Response rates vary according to the type of neuroendocrine malignancy; for example, rates as high as 80% have been reported with metastatic gastrinoma.

Second-line regimen incorporating VP-16 and cisplatin have also been evaluated with useful activity and palliation in some patients.

Tumours of steroid hormone-secreting cells are usually treated surgically. Patients with adrenocortical adenoma are generally cured by surgical tumour resection. Occasional patients with carcinoma experience long survival despite incomplete resection of their tumours, but most patients succumb to recurrent disease with metastases noted within the first 2 years following diagnosis. The adrenocorticolytic medication mitotane provides some temporary relief from cortisol excess in some patients.

Tumours of thyroid follicular epithelium are also most amenable to surgical resection but aggressive surgery is not usually indicated. These lesions have a unique affinity for iodine uptake and therefore microscopic residual disease is usually ablated by administration of radioactive iodine. As thyroid tumours lose their differentiation, one of the manifestations is reduction or loss of expression of the sodium iodide symporter that that is responsible for the success of this form of therapy. Therefore, lack of radioiodine uptake can occasionally mean that disseminated disease is of a less differentiated type. In patients with poorly differentiated or anaplastic carcinomas, or those with gross residual disease of differentiated carcinoma that is not amenable to complete surgical excision, there is a role for external beam radiotherapy.

REFERENCES

Apel, R. L. and Asa, S. L. (2001) The parathyroid glands. In: Barnes, L. (ed.), *Surgical Pathology of the Head and Neck*. 1719–1992 (Marcel Dekker, New York).

Asa, S. L. (1998). *Tumors of the Pituitary Gland. Atlas of Tumor Pathology*, Third Series, Fascicle 22 (Rosai, J., ed.) (Armed Forces Institute of Pathology, Washington, DC).

Asa, S. L. and Bedard, Y. C. (2000). Fine-needle aspiration cytology and histopathology. In: Clark, O. M. and Noguchi, S. (eds), *Thyroid Cancer. Diagnosis and Treatment.* 39–104. (Quality Medical Publishing, St. Louis, MO).

Asa, S. L. and Ezzat, S. (1998). The cytogenesis and pathogenesis of pituitary adenomas. *Endocrine Reviews*, **19**, 798–827.

Brierley, J. D. and Asa, S. L. (2001). Thyroid Cancer. In: Gospodarowicz, M. K. (ed.), *Prognostic Factors in Cancer.* 195–208 (Wiley-Liss, New York).

Capella, C., *et al*. (1995). Revised classification of neuroendocrine tumors of the lung, pancreas and gut. *Virchows Archiv, Abteiluns A: Pathologische Anatomie und Histopathologie*, **425**, 547–560.

DeLellis, R. A. and Tischler, A. S. (1998). The dispersed neuroendocrine cell system. In: Koracs, K. and Asa, S. L. (eds), *Functional Endocrine Pathology.* 529–549 (Blackwell, Boston).

Ezzat, S. and Asa, S. L. (1998). Molecular genetics of endocrine neoplasia. In: Koracs, K. and Asa, S. L. (eds.), *Functional Endocrine Pathology.* 967–983 (Blackwell, Boston).

Goretzki, P. E. *et al*. (1992). Mutational activation of RAS and GSP oncogenes in differentiated thyroid cancer and their biological implications. *World Journal of Surgery*, **16**, 576–582.

Greenlee, R. T., *et al*. (2000). Cancer statistics, 2000. *A Cancer Journal for Clinicians*, **50**, 7–33.

Klöppel, G. *et al*. (1998) The endocrine pancreas. In: Koracs, K. and Asa, S. L. (eds), *Functional Endocrine Pathology.* 415–487 (Blackwell, Boston).

Komminoth, P. (1999). Review: multiple endocrine neoplasia type 1, sporadic neuroendocrine tumors, and MENIN. *Diagnostic Molecular Pathology*, **8**, 107–112.

Krohn, D. *et al*. (1998). Clonal origin of toxic thyroid nodules with constitutively activating thyrotropin receptor mutations. *Journal of Endocrinology and Metabolism*, **83**, 180–184.

Kroll, T. G. *et al*. (2000). PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma. *Science*, **289**, 1357–1360.

Lack, E. E. (1997). *Tumors of the Adrenal Gland and Extra-adrenal Paraganglia. Atlas of Tumor Pathology,* Third Series, Fascicle 19 (Rosai, J., ed.) (Armed Forces Institute of Pathology, Washington, DC).

Lack, E. E. (1998). The Adrenal Cortex. In: Koracs, K. and Asa, S. L. (eds), *Functional Endocrine Pathology.* 596–636. (Blackwell, Boston).

LiVolsi, V. A. (1990). *Surgical Pathology of the Thyroid* (W.B. Saunders, Philadelphia).

Lyons, J., *et al*. (1990). Two G protein oncogenes in human endocrine tumors. *Science*, **249**, 635–639.

Moley, J. F. (2000). Medullary thyroid carcinoma. In: Clark, O. H. and Noguchi, S. (eds), *Thyroid Cancer. Diagnosis and Treatment.*279–308 (Quality Medical Publishing, St. Louis, MO).

Mulligan, L. M. and Ponder, B. A. J. (1995). Genetic basis of endocrine disease. Multiple endocrine neoplasia type 2. *Journal of Clinical Endocrinology and Metabolism*, **80**, 1989–1995.

Murray, D. (1998). The thyroid gland. In: Koracs, K. and Asa, S. L. (eds), *Functional Endocrine Pathology.* 295–380 (Blackwell, Boston).

Nikiforova, M. N., *et al*. (2000). Proximity of chromosomal loci that participate in radiation-induced rearrangements in human cells. *Science*, **290**, 138–141.

Porcellini, A., *et al*. (1994). Novel mutations of thyrotropin receptor gene in thyroid hyperfunctioning adenomas. Rapid identification by fine needle aspiration biopsy. *Journal of Endocrinology and Metabolism*, **79**, 657–661.

Rosai, J., *et al*. (1992). *Tumors of the Thyroid Gland. Atlas of Tumor Pathology,* Third Series, Fascicle 5 (Rosai, J. ed.) (Armed Forces Institute of Pathology, Washington, DC).

Sasano, H. (1998) Steroid-producing tissues. In: Kovacs, K. and Asa, S. L. (eds), *Functional Endocrine Pathology.* 707–714. (Blackwell, Boston).

Soravia, C., *et al*. (1999). Familial adenomatous polyposis-associated thyroid cancer. *American Journal of Pathology*, **154**, 127–135.

Spada, A., *et al*. (1992). G protein oncogenes in pituitary tumors. *Trends in Endocrinology and Metabolism*, **3**, 355–360.

Suarez, H. G., *et al*. (1991). Gsp mutations in human thyroid tumors. *Oncogene*, **6**, 677–679.

Tischler, A. S. (1998). The adrenal medulla and extra-adrenal paraganglia. In: Kovas, K. and Asa, S. L. (eds), *Functional Endocrine Pathology*. 550–595. (Blackwell, Boston).

van Sande, J., *et al*. (1995). Genetic basis of endocrine disease. Somatic and germline mutations of the TSH receptor gene in thyroid disease. *Journal of Endocrinology and Metabolism*, **80**, 2577–2585.

Wick, M. R. (2000). Immunohistology of neuroendocrine and neuroectodermal tumors. *Seminars in Diagnostic Pathology*, **17**, 194–203.

FURTHER READING

Asa, S. L. (1998). *Tumors of the Pituitary Gland. Atlas of Tumor Pathology*, Third Series, Fascicle 22 (Rosai, J. ed.) (Armed Forces Institute of Pathology, Washington, DC).

Clark, O. H. and Noguchi, S. (eds) (2000). *Thyroid Cancer. Diagnosis and Treatment* (Quality Medical Publishing, St. Louis, MO).

DeLellis, R. A. (1993). *Tumors of the Parathyroid Gland. Atlas of Tumor Pathology*, Third Series, Fascicle 6 (Rosai, J. ed.) (Armed Forces Institute of Pathology, Washington, DC).

Kovacs, K. and Asa, S. L. (eds) (1998). *Functional Endocrine Pathology* (Blackwell, Boston).

Lack, E. E. (1997). *Tumors of the Adrenal Gland and Extra-adrenal Paraganglia. Atlas of Tumor Pathology,* Third Series, Fascicle 19 (Rosai, J., ed.) (Armed Forces Institute of Pathology, Washington, DC).

LiVolsi, V. A. (1990). *Surgical Pathology of the Thyroid* (W.B. Saunders, Philadelphia).

Rosai, J., *et al*. (1992). *Tumors of the Thyroid Gland. Atlas of Tumor Pathology*, Third Series, Fascicle 5 (Rosai, J., ed.) (Armed Forces Institute of Pathology, Washington, DC).

Chapter 42

Breast

Shahla Masood and Darian Kameh
University of Florida Health Science Center at Jacksonville, FL, USA

CONTENTS

- Normal Development and Structure
- Epidemiology and Risk Factors
- Aetiology
- Screening
- Diagnosis
- Common Benign Lesions
- Ductal Carcinoma *In Situ* (DCIS)
- Paget Disease of the Nipple
- Lobular Carcinoma *In Situ* (LCIS)
- Invasive Ductal Carcinoma, Not Otherwise Specified (NOS)
- Invasive Lobular Carcinoma
- Other Tumours of the Breast
- Special Studies/Prognostic Indicators
- Treatment

NORMAL DEVELOPMENT AND STRUCTURE

The female breast (**Figure 1**) is a heterogenous structure which overlies the pectoralis major and minor muscles covering the chest wall. The resting mammary gland consists of 5–10 major duct systems which are arranged in a segmental, roughly radial pattern. These duct systems are subdivided into lobules, which are the functional units of the mammary parenchyma. Each ductal system drains through an individual lactiferous sinus. Successive branching leads to terminal ducts distally, which end blindly prior to puberty. With the onset of menarche (first appearance of the menstrual cycle), the terminal ducts proliferate distally, giving rise to lobules consisting of a cluster of epithelium-lined ductules or acini. Each terminal duct and its ductules compose the terminal duct lobular unit (TDLU) (**Figure 2; see colour plate section**) (Lawrence, 1992; Cotran *et al*., 1999).

The majority of the breast stroma consists of dense fibroconnective tissue admixed with adipose fatty tissue, called interlobular stroma, containing elastic fibres supporting the large ducts. The lobules themselves are enclosed by loose delicate myxomatous stroma that is hormonally responsive and contains scattered lymphocytes (intralobular stroma) (Cotran *et al*., 1999).

Histologically, the intact ducts are composed of an inner layer of cuboidal to low columnar epithelial cells, surrounded by a discontinuous layer of myoepithelial cells, enclosed by a basement membrane (**Figure 3**). It is the inner layer of epithelial cells within the TDLU that gives rise to the common forms of breast cancer.

Lobule
Terminal duct
Segmental duct
Lactiferous duct
Lactiferous sinus
Nipple

Figure 1 Normal adult female breast (simplified).

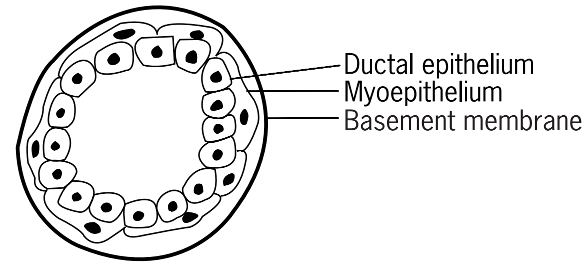


Figure 3 Breast duct histology (cross-section).

The lymphatic drainage of the breast is primarily to the axilla, while a minor portion of it courses to the internal mammary nodes (Lawrence, 1992).

EPIDEMIOLOGY AND RISK FACTORS

The incidence of breast cancer in developed countries is 200–250 per 100 000 women per year. One of nine women in the United States will develop breast cancer in her lifetime; one-third of these women will succumb to the disease, resulting in almost 50 000 deaths yearly (Cotran *et al*., 1999). The incidence increases with age with an average age at diagnosis of approximately 65 years. Age-adjusted incidence rates are rising, probably due in large part to an increase in mammographic screening and the changing epidemiological profile of women (White *et al*., 1990). Recognizing risk factors can help identify high-risk patients who need more intense monitoring as well as provide insight into the pathogenesis of the disease (Masood, 1996a). **Table 1** lists the best recognized risk factors. Women in North America and European countries have the highest rates of breast cancer, and those in Asian and African countries have the lowest (Kelsey and Horn-Ross, 1993). Breast cancer only rarely occurs in males.

About 10–15% of cases of breast cancer have a family history of breast or ovarian cancer (Thompson, 1994). With the identification of the breast cancer susceptibility genes *BRCA1* and *BRCA2* on chromosomes 17 and 13, respectively, a genetic predisposition clearly exists. However, many women with a family history of breast cancer do not carry mutations in one of these genes. The gene products of *BRCA1* and *BRCA2* are multifunctional proteins involved in maintaining genomic stability and the response to DNA damage, transcriptional regulation and cell proliferation (Unger and Weber, 2000; Welsch *et al*., 2000). Mutations in these genes are associated with a marked increase in the overall risk of developing breast cancer and a marked decrease in the age of onset. Specific mutations are more common in some ethnic groups (such as the Ashkenazi Jews) (Masood, 1996a; Cotran *et al*., 1999).

AETIOLOGY

Specific cellular changes occur in the progression to breast cancer. The earliest genetic lesions have been identified in morphologically normal breast epithelium near the site of a carcinoma, and usually large chromosomal deletions (Deng *et al*., 1996). In addition, one of the earliest detectable changes is the loss of normal regulation of cell number, resulting in epithelial hyperplastic lesions. Next, genetic instability occurs in multifocal small clonal populations of cells that can be recognized histologically as atypical hyperplasia. Complete replacement of the normal ducts by atypical cells characterizes carcinoma *in situ*; at this point, there are often (but not always) alterations in specific genes, such as the oestrogen receptor gene, *TP53*, *HER-2/neu/erbB-2*, and cyclin D (Devilee *et al*., 1994). Invasion of the surrounding stroma indicates an invasive neoplasm; although this process must be marked by alterations in specific genes leading to dysregulation of proteases, cell adhesion factors and angiogenic factors, the genetic lesions are not well understood. Similarly, the changes seen in breast cancer cells present in metastatic lesions are not well understood, although a recent report suggests that up-regulation of specific chemokines may account for specific organ propensities for harbouring metastases from breast carcinoma (Muller *et al*., 2001). It is hoped that microarray studies will shed light on these processes (Perou *et al*., 1999; Sgroi *et al*., 1999).

SCREENING

Education and diligent breast self-examination improve the possibility of early breast cancer detection. However, regular mammography is the cornerstone of effective screening for breast cancer. Early breast cancers have an excellent prognosis, with increased long-term survival. The American Cancer Society recommends an annual physical examination, beginning at age 40 years, accompanied by screening mammography at 1–2 year intervals until the age of 50, and yearly at age 50 and older (Masood, 1996a). Minimally invasive techniques such as fine-needle

Table 1 Risk factors for developing breast cancer

High risk (>4 times general population)	Moderate risk (2–4 times general population)	Slight risk (1–2 times general population)
Personal history of prior breast cancer	First-degree relative with history of breast cancer	Onset menarche <12 years of age
Family history of bilateral, premenopausal or familial cancer syndrome	Upper social/economic class	Moderate alcohol intake
Proliferative breast disease with atypia	Prolonged uninterrupted menses	
	Postmenopausal obesity	
	Personal history of prior carcinoma of ovary or endometrium	
	Proliferative breast disease with no atypia	

aspiration biopsy (FNAB) (**Figure 4; see colour plate section**) and core-needle biopsy can be used further to classify and assess lesions deemed radiographically 'suspicious' for malignancy. These procedures can be performed with or without radiographic guidance and can provide valuable information regarding which patients require surgical excision, radical treatment or simple follow-up (Masood, 1996b).

DIAGNOSIS

Generally, palpable breast lesions are felt by a patient or discovered by a physician during physical examination. Breast cancer often presents as an irregular, rock-hard mass with ill-defined margins that invade adjacent tissues (**Figure 5; see colour plate section**). The mass usually is not moveable and is inseparable from adjacent tissues. Rarely, patients may show diffuse induration of the breast skin due to the presence of tumour in dermal lymphatics; this is called inflammatory carcinoma. Non-palpable lesions are detected by imaging techniques (mammography with or without ultrasound). Physical exam and radiologic studies can, at best, establish a clinical suspicion of cancer; however, a firm diagnosis of cancer is made only by the pathologist. Pathologic studies include cytologic evaluation of FNAB material or histologic evaluation of tissue obtained by core needle biopsy or surgical excision (Lawrence, 1992; Masood, 1996a,b). (See also chapter *Breast Imaging and Diagnosis*.)

COMMON BENIGN LESIONS

Benign 'masses' or 'mass-like areas' occur far more frequently than cancer, at a rate of almost 10 : 1. Atypical clinical presentation of benign lesions leads to a clinical suspicion of carcinoma, for which pathological confirmation is warranted (Lawrence, 1992). Several benign lesions can mimic the clinical presentation of cancer, as follows.

Fibroadenoma

This is the most commonly occurring benign tumour of the female breast. These lesions are seen in women who are between the ages of 20 and 35 years, and are composed of hormonally sensitive glandular and stromal tissue. Grossly, these tumours are well circumscribed and have a solid, whorled, greyish white, bulging cut surface, with slit-like spaces. Histologically, these benign growths are composed of fibroblastic stroma resembling intralobular stroma, along with epithelial-lined ductal structures that tend to be cystic, elongated or flattened (**Figure 6; see colour plate section**). These tumours lead to a mild increase in the risk of subsequent breast cancer. Surgical excision is curative (Rosai, 1989; Lawrence, 1992; Masood, 1996b; Cotran *et al*., 1999).

Fibrocystic Change

This represents the single most common disorder of the breast and it accounts for more than half of all surgical operations on the female breast. It occurs most frequently between the ages of 25 and 45 years, with a peak incidence just before menopause, and rarely developing after menopause. Hormonal imbalances are considered to be basic to the development of this disorder. Grossly, the involved tissue has a whitish, poorly circumscribed fibrotic appearance, alternating with cystic structures. Histologically, there are alternating areas of fibrosis, cyst formation, apocrine metaplasia, ductal hyperplasia and adenosis (**Figure 7; see colour plate section**). In the absence of significant ductal hyperplasia (also termed proliferative breast disease), fibrocystic changes do not elevate the risk of developing cancer. Once there is histological confirmation of the diagnosis, no further treatment is warranted (Rosai, 1989; Lawrence, 1992; Masood, 1996b; Cotran *et al*., 1999).

Other Benign Lesions

Between lesions such as fibroadenoma and fibrocystic change on the one hand and cancer on the other, there exists a range of breast lesions that impart an increased risk of subsequent carcinoma. These include proliferative breast disease (e.g. ductal epithelial hyperplasia), sclerosing adenosis and papillomatosis (**Figure 8; see colour plate section**), to name a few. These lesions are beyond the scope of this chapter and the reader is referred to the Further Reading section for further discussion of these entities.

DUCTAL CARCINOMA *IN SITU* (DCIS)

The number of cases of DCIS has increased rapidly over the past 20 years with the advent of mammographic screening. The lesion consists of a malignant population of epithelial cells that are confined by the basement membrane. However, these cells can spread throughout a regional ductal system, producing extensive segmental lesions, or later develop into invasive cancer. DCIS may or may not be grossly apparent (Rosai, 1989). Two general classes of DCIS are recognized, as follows.

Comedo DCIS

This is characterized by distended ducts, solid sheets of high-grade malignant cells and central necrosis (**Figure 9; see colour plate section**). The necrotic

material often calcifies and is radiographically detectable. This form of DCIS has a greater malignant potential than other forms.

Non-comedo DCIS

This is the more common form and can have nuclear grades ranging from low to high. The malignant cells have a monomorphic appearance, necrosis and calcifications may be present. Histological variants include solid (**Figure 10; see colour plate section**), cribriform, papillary and micropapillary (Masood, 1996b; Cotran *et al*., 1999). For further discussion of these variants, see the Further Reading section.

PAGET DISEASE OF THE NIPPLE

This is a variant of carcinoma *in situ* in which the nipple skin is involved. Clinically, the patient often presents with a scaling or eczematous lesion of the nipple. Histologically, cancer cells are seen in between normal epithelial cells of the epidermis of the skin. Paget disease of the nipple is the result of intraepithelial spread from an underlying intraductal carcinoma in nearly all cases; in about half of cases there is an associated infiltrating carcinoma.

LOBULAR CARCINOMA *IN SITU* (LCIS)

LCIS is also known as lobular neoplasia. It usually is not grossly evident and is often an incidental finding in breast tissue removed for other reasons. Histologically, the lobules are distended and completely filled by relatively uniform, round, small- to medium-sized cells. Marked atypia, pleomorphism and mitotic activity are usually absent. LCIS is often multifocal in nature and is thought of more as a 'marker' of breast cancer rather than a direct precursor. This is due to the fact that women diagnosed with LCIS have a marked increase in the risk of developing invasive lobular or ductal carcinoma in their lifetimes (Rosai, 1989; Masood, 1996b; Cotran *et al*., 1999).

INVASIVE DUCTAL CARCINOMA, NOT OTHERWISE SPECIFIED (NOS)

This type of carcinoma represents the 'prototype' breast cancer and includes the majority of carcinomas (70–80%) that cannot be classified under any other subtype. Usually presenting as a fixed firm mass with a 'stellate' or 'spiculated' appearance on mammography, it grossly is tan-white with a firm, often gritty consistency and raggedly infiltrating, poorly defined borders (**Figure 11; see colour plate section**).

Histologically, the infiltrating edge of the tumour extends beyond what is visible grossly, warranting ample excision of normal tissue. The tumour itself consists of malignant ductal epithelial cells disposed in cords, solid nests, tubules, anastomosing masses and mixtures of all these patterns (**Figure 12; see colour plate section**).

The cells invade into the surrounding fibrofatty tissue and by elaborating certain chemical signals, stimulate the adjacent fibrous tissue to proliferate (desmoplastic response), giving the tumour its firm consistency. The histological grading of the tumour is based on mitotic count, cytological atypia and degree of tubule formation. Well-differentiated (low histological grade) tumours often display minimal atypia, low mitotic activity and have prominent tubule formation. Poorly differentiated tumours often have marked cytological atypia, prominent mitotic activity and a solid pattern of growth.

Other distinct histological variants of invasive ductal carcinoma include medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma (**Figure 13; see colour plate section**) and papillary carcinoma (Rosai, 1989; Cotran *et al*., 1999). For further discussion of these variants, see the Further Reading section.

INVASIVE LOBULAR CARCINOMA

This type of invasive cancer comprises only 5–10% of primary breast carcinomas. It tends to be multicentric within the same breast, with a diffusely infiltrative pattern and tends to be bilateral far more frequently than other subtypes. Grossly, the tumour can be rubbery and poorly defined, or gritty hard with an infiltrating border, similar to ductal carcinoma. Histologically, the 'classical' appearance is that of small epithelial cells with little pleomorphism, that invade in strands only one cell thick (often referred to as 'Indian filing' (**Figure 14; see colour plate section**)). Signet ring cells (i.e. cells in which intracytoplasmic mucin pushes the nucleus to the periphery, resembling a signet ring) are common (Masood, 1996b; Cotran *et al*., 1999).

OTHER TUMOURS OF THE BREAST

Cystosarcoma phyllodes is a rare mixed stromal and epithelial neoplasm that may be benign or malignant. The size varies from 1 cm to >10 cm. The epithelial component is similar to that seen in fibroadenoma, whereas the stromal component is more cellular and has greater mitotic activity. The degree of malignancy is estimated by the degree of cellularity, mitotic rate and cytological atypia. Rarely, benign soft tissue tumours and sarcomas may occur as

primary tumours in the breast. In addition, malignant lymphomas may present in the breast.

SPECIAL STUDIES/PROGNOSTIC INDICATORS

By far the most important prognostic indicator is the stage of the tumour, particularly the lymph node status. When axillary nodes are uninvolved, 10-year disease-free survival approaches 80%, but falls to as low as 10% if 10 or more nodes are involved. Increased size of the metastatic deposit and invasion of the metastasis through the lymph node capsule convey a worse prognosis. A poor prognosis is also associated with locally advanced disease (i.e. skin and skeletal muscle involvement), increased tumour size, histological evidence of lymphovascular invasion and high histological grade (more poorly differentiated) tumours. Some specific histological subtypes of invasive cancer (colloid, tubular, medullary, lobular and papillary) actually are associated with a better prognosis than the 'not otherwise specified' (NOS) subtype (Masood, 1996a; Cotran *et al*., 1999).

Special studies can be performed on cytological or solid tissue specimens submitted to the pathology laboratory. The proliferation rate can be assessed by flow cytometry to determine the fraction of cells actively synthesizing DNA and by immunohistochemistry to detect cellular proteins (such as Ki-67) expressed during the cell cycle. A higher proliferative rate is associated with tumours of more aggressive biologic behaviour. DNA content can be determined by flow cytometry analysis or image analysis. Aneuploid tumours have a slightly worse prognosis. The presence of oestrogen and progesterone hormone receptors within tumuor cells is associated with a better prognosis than those tumours that are receptor-negative. This is due to the high regression rate of receptor-positive tumours in response to hormonal manipulation (i.e. tamoxifen, discussed below). Hormone receptor studies can be performed by immunohistochemical methods (**Figure 15; see colour plate section**).

Loss of function of tumour-suppressor genes such as *TP53* (often detected by an increase in immunostaining of inactivated p53 protein) and increased expression of oncogenes such as *Her-2/neu* are commonly seen in association with other poor prognostic factors. Overexpression of *Her-2/neu* has been correlated with a shorter disease-free interval and shorter overall survival in node-positive or node-negative breast cancers. Patients can also be offered Herceptin (see below), based on *Her-2/neu* tissue studies. p53 and Her-2/neu testing can be performed by immunohistochemical methods (**Figure 16; see colour plate section**) or fluorescent *in situ* hybridization (FISH) studies (Rosai, 1989; Cotran *et al*., 1999; Masood and Bui, 2000).

TREATMENT

Previously, the long-held theory of orderly regional to systemic progression resulted in aggressive local treatment, i.e. radical mastectomy (which included the resection of both pectoralis major and minor muscles). Now, a further understanding of the biology of breast cancer has supported an opposing theory, which states that breast cancer has a nonorderly progression to early systemic spread from inception (Lawrence, 1992). Studies have shown equivalent survival data when modified radical mastectomy (which only removes the pectoralis minor muscle) is compared with lumpectomy (excision of the lesion, with a surrounding rim of normal breast tissue) with or without adjuvant therapy (Masood, 1996a). That being said, surgical therapy still remains the foundation of breast cancer treatment. The extent of the initial surgical procedure is determined by the histological grade of the tumour, clinical stage at presentation and the patient's desire for breast conservation and reconstruction, if appropriate. Axillary nodes are routinely sampled or completely dissected to determine the presence of local–regional metastasis (Lawrence, 1992). Radiation therapy has been employed as a postoperative adjunct (especially in connection with the more limited surgical resections), sometimes as the primary treatment and for the control of locally advanced or recurrent disease (Rosai, 1989). Adjuvant chemotherapy is used in the control of regional (axillary) or systemic metastatic disease. Chemotherapy has also been used in the preoperative treatment of larger, high histological grade breast cancers with or without the presence of axillary metastasis. Tamoxifen, which functionally blocks oestrogen receptors in tumour cells which express them, is used in similar settings to other chemotherapeutic agents (Masood, 1996a; Lawrence, 1992). There also appears to be a role for tamoxifen even in the absence of nodal metastasis. The use of the drug herceptin, which targets cells that overexpress Her-2/neu, has found widespread use in patients with recurrent disease, whose tumours over express the *Her-2/neu* oncogene (Masood and Bui, 2000).

REFERENCES

Cotran, R. S., *et al*. (1999). *Robbins Pathologic Basis of Disease*, 6th edn. 1093–1119. (Lippincott-Raven, Philadelphia).

Deng, G., *et al*. (1996). Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science*, **274**, 2057–2059.

Devilee, P., *et al*. (1994). Recent developments in the molecular genetic understanding of breast cancer. *Critical Reviews in Oncology*, **5**, 247–270.

Kelsey, J. L. and Horn-Ross, P. L. (1993). Breast cancer: magnitude of the problem and descriptive epidemiology. *Epidemiological Reviews*, **15**, 7–16.

Lawrence, P. F. (ed.) (1992). *Essentials of General Surgery*, 2nd edn. 272–283 (Williams and Wilkins, Baltimore).

Masood, S. (1996a). Breast health: challenges and promises. *Journal of the Florida Medical Association*, **83**, 459–465.

Masood, S. (1996b). *Cytopathology of the Breast*, 1st edn. 1–5, 78, 167–180, 203–271 (American Society of Clinical Pathologists, Chicago).

Masood, S. and Bui, M. (2000). Assessment of HER-2/neu overexpression in primary breast cancers and their metastatic lesions: an immunohistochemical study. *American Clinical Laboratory Science*, **30**, 259–265.

Muller, A., *et al*. (2001). Involvement of chemokine receptors in breast cancer metastasis. *Nature*, **410**, 50–56.

Perou, C. M., *et al*. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proceedings of the National Academy of Sciences of the USA*, **96**, 9212–9217.

Rosai, J. (1989). *Ackerman's Surgical Pathology*, 7th edn. 1193–1267 (C.V. Mosby, St. Louis).

Sgroi, D. C., *et al*. (1999). *In vivo* gene expression profile analysis of human breast cancer progression. *Cancer Research*, **59**, 5656–5661.

Thompson, W. D. (1994). Genetic epidemiology of breast cancer. *Cancer*, **74**, 279–287.

Unger, M. A. and Weber, B. L. (2000). Recent advances in breast cancer biology. *Current Opinions in Oncology*, **16**, 69–74.

Welsch, P. L., *et al*. (2000). Insights into the functions of BRCA1 and BRCA2. *Trends in Genetics*, **16**, 69–74.

White, E., *et al*. (1990). Evaluation of the increase in breast cancer incidence in relation to mammography use. *Journal of the National Cancer Institute*, **82**, 1546–1552.

Fisher, E. R. (1984). The impact of pathology on the biologic, diagnostic, prognostic and therapeutic considerations in breast cancer. *Surgical Clinics of North America*, **64**, 1073–1093.

Hutter, R. V. P. (1984). Pathological parameters useful in predicting prognosis for patients with breast cancer. In: McDivitt, R. W., *et al*. (eds), *The Breast* (Williams and Wilkins, Baltimore).

Osborne, M. P. (1991). Breast development and anatomy. In: Harris, J. R., *et al*. (eds), *Breast Diseases*, 2nd edn. 1–13 (Lippincott, Philadelphia).

Page, D. L. and Anderson, T. J. (1988). *Diagnostic Histopathology of the Breast*, Churchill Livingstone (New York).

Romrell, L. J. and Bland, K. I. (1998). Anatomy of the breast, axilla, chest wall and related metastatic sites. In: Bland, K. I. and Copeland, E. M. (eds), *The Breast*, 2nd edn, Vol. 2 (W.B. Saunders, Philadelphia).

Rosen, P. P. (1979). The pathological classification of human mammary carcinoma: past, present and future. *Annals of Clinical Laboratory Science*, **9**, 144–156.

Rosenbloom, A. L. (1998). Breast physiology: normal and abnormal development and function. In: Bland, K. I. and Copeland, E. M. (eds), *The Breast*, 2nd edn, Vol. 2 (W.B. Saunders, Philadelphia).

Silverberg, S. G. and Masood, S. (1997). The breast. In: Silverberg, S. G., *et al*. (eds), *Principles and Practice of Surgical Pathology and Cytopathology*, 3rd edn, Vol. 2 (Churchill Livingstone, New York).

World Health Organization (1981). *International Histological Classification of Tumors, No. 2, Histologic Typing of Breast Tumors*, 2nd edn (WHO, Geneva).

FURTHER READING

Elston, C. W. and Ellis, I. O. (1998). *The Breast*, 3rd edn, Vol. 13 (Churchill Livingstone, Nottingham).

Chapter 43

Female Reproductive System

Beth Euscher, Carl Morisson and Gerard Nuovo
Ohio State University Medical Center, Columbus, OH, USA

CONTENTS

- Normal Development and Structure
- Pathology of the Cervix
- Pathology of the Endometrium
- Uterine Stromal Tumours
- Pathology of the Ovary

NORMAL DEVELOPMENT AND STRUCTURE

The female genital tract consists of the ovaries, fallopian tubes, surrounding adnexa, uterus, vagina, and vulva. Despite their proximity and their similarities in many aspects, such as response to oestrogen and progesterone, the female genital tract actually represents the combination of three distinct regions during embryogenesis. These include the ovaries, which begin as midline structures that migrate to the peritoneal cavity, the Mullerian system, which gives rise to the endometrium, myometrium, cervix and outer part of the vagina, and the ectodermal system, from which the vulva and part of the vagina originate. The mesonephric system, which plays a role in filtration of impurities in early foetal development, is also represented in the female genital tract as rests which are often found in the adnexa around the fallopian tube and in the lateral wall of the cervix; these rarely cause clinically relevant pathology.

PATHOLOGY OF THE CERVIX

Tumour Pathology

Squamous cell carcinoma is the most common tumour of the cervix, accounting for nearly 75% of the tumours from this region. This, of course, is consistent with the observation that the strong majority of premalignant lesions of the cervix are derived from squamous cells, termed squamous intraepithelial lesions (SILs). There are several histological subtypes within the category of squamous cell carcinoma of the cervix. The most common type is moderately well differentiated squamous cell carcinoma, where rare cells show the keratinization diagnostic of this tumour. Less common is poorly differentiated squamous cell carcinoma, where keratin formation is very difficult to find with routine haematoxylin and eosin stain. Interestingly, extremely rare in the cervix is well differentiated squamous cell carcinoma, where keratin pearls and individual cell dyskeratosis are abundant. This is in sharp contrast to the vulva, where well differentiated squamous cell carcinoma is very common. In the vulva, this type is rarely associated with infection by human papillomavirus (HPV) (Nuovo *et al*., 1991). This may explain why this type is so rare in the cervix for, as will be discussed below, over 98% of cervical cancers contain HPV DNA (Crum and Nuovo, 1991; Nuovo, 1994).

Another rare type of carcinoma found in the cervix is small cell carcinoma (**Figure 1**). This very rare variant has a poor prognosis, and is histologically equivalent to the very common small cell carcinoma of the lung. Such tumours are called neuroendocrine tumours, as they are commonly associated with the production of proteins that are part of the endocrine system, such as synaptophysin or chromogranin. (See also chapter *Endocrine Organs*.)

The other type of carcinoma found in the cervix is adenocarcinoma. About 25% of cancers of the cervix are of this type. Although the premalignant variant (adenocarcinoma *in situ*) is a well recognized entity, it is common to find SIL associated with adenocarcinoma of the cervix. This implies that the two entities share a common origin, which is indeed the case as will be discussed in the section Aetiology. Another important point that highlights the relatedness between adenocarcinoma and squamous cell carcinoma of the cervix is the fact that they both originate in the exact same location – the transformation zone of the cervix (Crum and Nuovo, 1991; Nuovo, 1994). There are several subtypes of adenocarcinoma of the cervix. The most common type is a moderately well differentiated

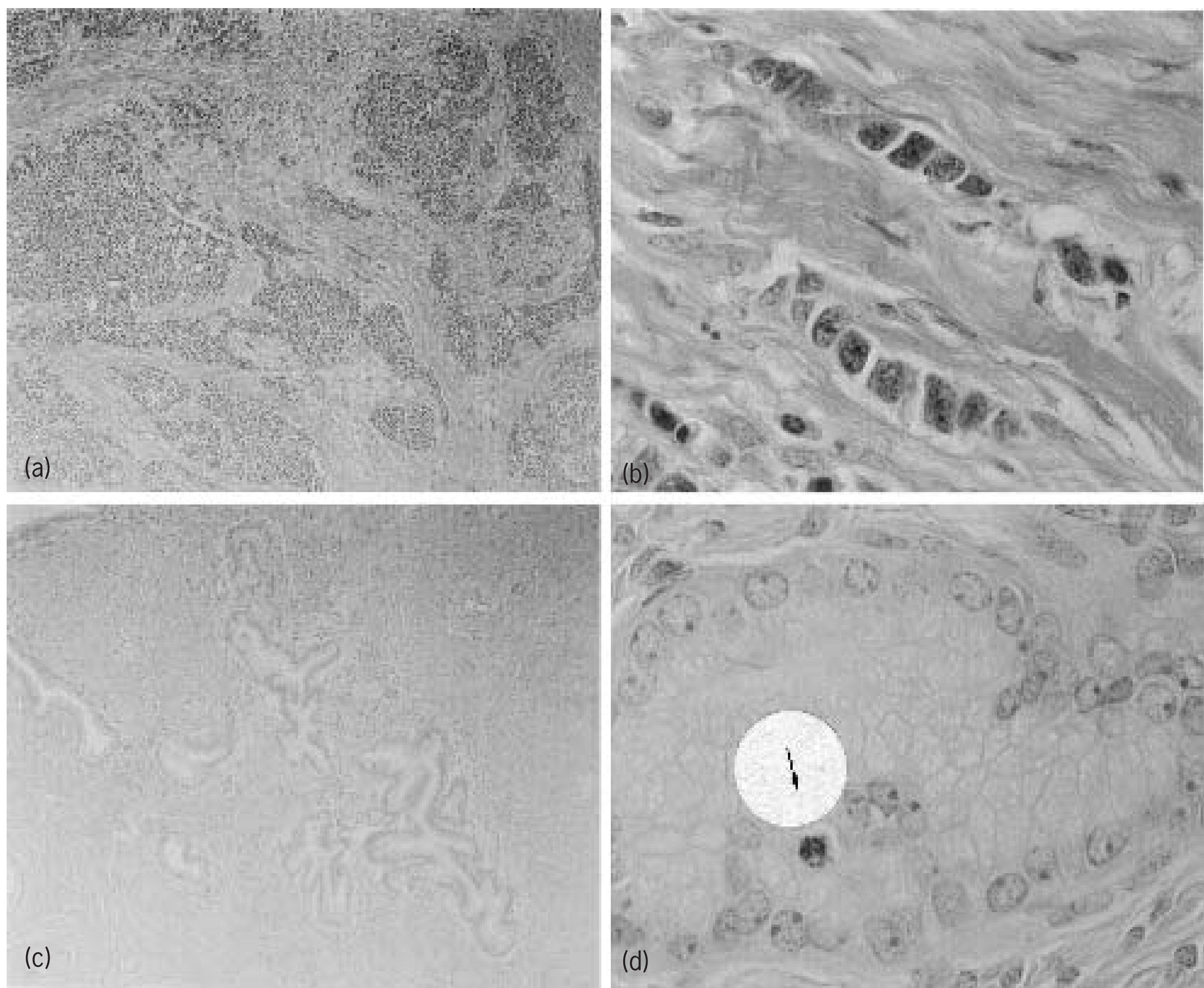


Figure 1 Histology of unusual carcinomas of the cervix. Panel (a) shows the pattern of growth of a small cell carcinoma. At higher magnification (b) note the lack of cytoplasm and nuclear molding. Panel (c) shows the well-formed glands which, however, are invading the underlying stroma. Note the mitotic figure ((d), arrow), characteristic of this tumour, called adenoma malignum. Both lesions contained HPV and showed areas of squamous cell dysplasia.

lesion in which mucous production is easily documented by special stains, such as the mucicarmine or PAS stain (**Figure 2**). At times, such tumours will show cribriform formation, similar to what is evident in endometrial adenocarcinomas. Further, like adenocarcinoma of the endometrium, squamous metaplasia is often identified. As might be expected, it may be difficult – indeed impossible – to differentiate a primary endocervical cancer from an endometrial cancer that has invaded into the endocervix. Similarly, it may be very difficult to differentiate a primary endometrial adenocarcinoma from an endocervical primary that has invaded into the endometrium. These are important clinical distinctions, as they relate to the stage of the disease, which ultimately is the most important indicator of the prognosis of the woman. There is, however, a very simple way to distinguish between endocervical and endometrial adenocarcinoma. This can reliably be done with HPV testing, as endocervical cancer invariably contains HPV whereas endometrial cancer does not contain the virus (**Figure 2**). There are a few other types of endocervical adenocarcinoma to recognize. One is the clear cell variant, which looks histologically identical with the much more common renal cell carcinoma. Interestingly, this is the only type of adenocarcinoma of the cervix that is not found in the transformation zone or associated with HPV. It has been associated with DES exposure *in utero*; it is worth stressing that this is a very rare variant. Also, one should recognize the papillary variant of endocervical adenocarcinoma. On histological analysis, it looks identical with the common papillary serous carcinoma of the ovary (described later). It carries a poor prognosis. Finally, one should realize that the endocervix is the site of a rare variant of adenocarcinoma, very well differentiated, called adenoma malignum. The term adenoma stresses the

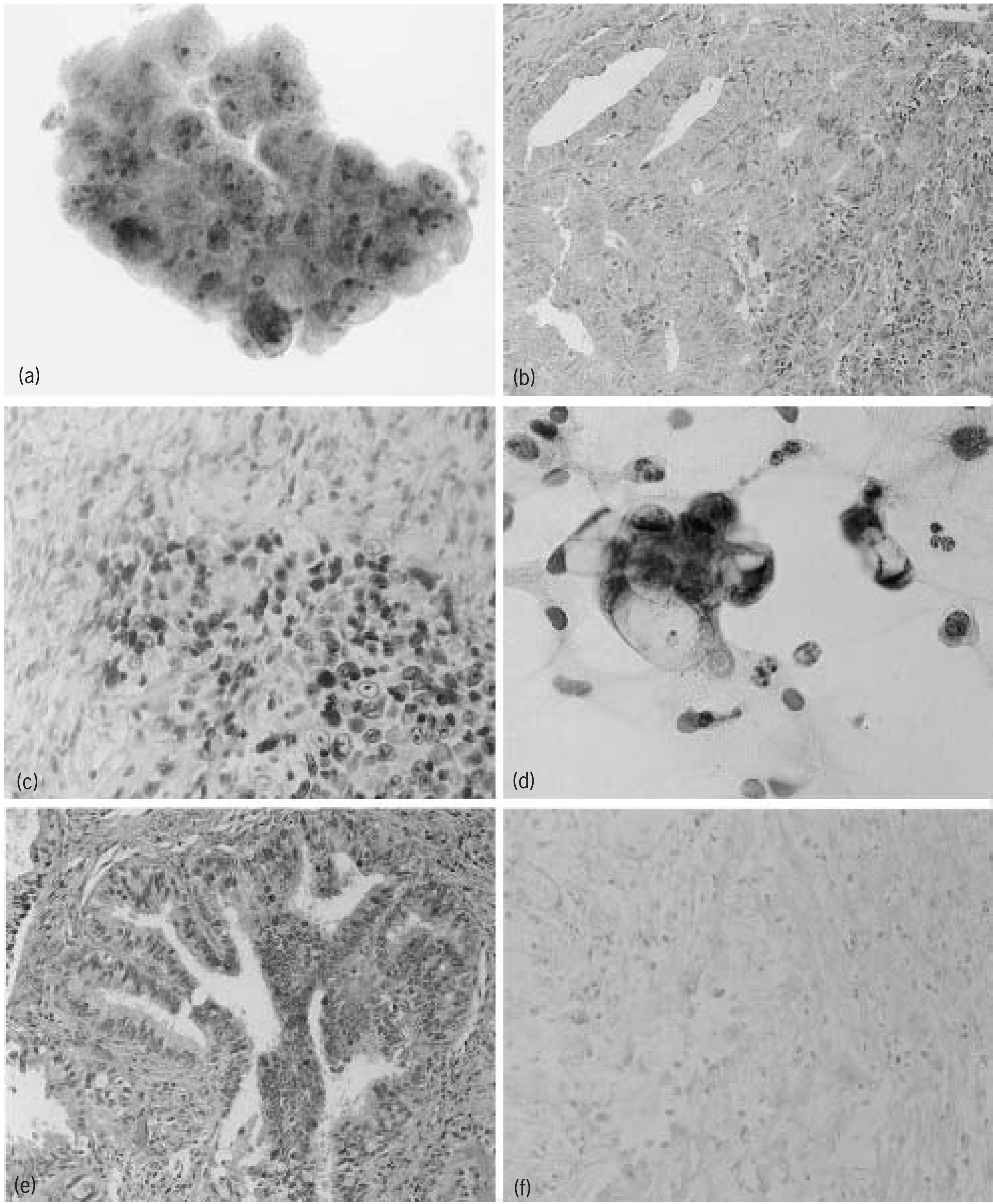


Figure 2 Differentiation between adenocarcinoma of the cervix and endometrium. Panel (a) shows the cytology of an adenocarcinoma from a Pap smear; note the variable cell density. The histology shows similar disorganisation (b); the lesion contains HPV 18 (panel (c), dark-staining cells are positive) confirming that the lesion is from the cervix. Compare this to the cytology ((d), Pap smear of the cervix) and histology (e) of a similar appearing tumour that was in the endometrium. HPV testing showed this to be HPV negative (f), confirming that it was an endometrial cancer.

fact that the glands are remarkably bland in appearance. Two features, mitotic activity (which is a very good marker of adenocarcinoma of the cervix) and branching, or claw-like glands infiltrating the stroma, are what allow one to differentiate adenoma malignum from normal endocervical glands (**Figure 1**).

Epidemiology

The cervix is unique in how well the epidemiology of cancers and precancers at this site are understood. Simply put, the epidemiological data for years have pointed to a sexually transmitted factor. Many years ago, before molecular techniques were so widely available, nearly every agent of sexually transmitted diseases (STDs) was implicated. In the early 1980s herpes simplex virus was a leading candidate (Crum and Nuovo, 1991; Nuovo, 1994). It would take the advent of molecular cloning and hybridization to realize that the actual agent was a virus, HPV. HPV is a small DNA virus – it contains only about 8000 base pairs – which cannot be grown in the laboratory. Its fastidious nature and small size are very well suited to study by molecular hybridization techniques. HPV will be discussed in detail in the Aetiology section. As noted above, it is well documented that the primary risk factor for a woman for developing cervical cancer is her number of sexual partners. This does not relate just to the actual number of her sexual partners, but also the number of female partners for the males with whom she has had sexual relations; this is the so-called 'high-risk male.' Clearly, if a man has had many partners this would put any given woman at higher risk of acquiring any sexual disease as compared with a man who has no other sexual partner. It also follows that groups where monogamy is strictly practiced, as is true of certain religious sects, or women who have not had sexual intercourse, such as nuns, are at very low if not zero risk for developing cervical cancer. Many other factors have been related to increased risk of SIL or cervical cancer, including cigarette smoking, but none has shown the strong correlation related to venereal transmission. It should be added that viruses related to HPV are common in many mammals, including monkeys, where a clear sexual spread has been documented (Crum and Nuovo, 1991; Nuovo, 1994).

Whereas HPV is very tropic for squamous epithelium and it cannot survive long without contact with such tissue, it follows that barrier methods of contraception may be useful in preventing the spread of this disease. Although this has been documented, the protection does not appear to be complete. Given the strong correlation between HPV infection and sexual transmission, an obvious question is whether any sites other than the genital tract show a high relationship between squamous tumours and HPV infection. Indeed, there are two other sites where the association is as strong as in the cervix, namely the periungual region (fingernail bed) of the fingers and, to a lesser extent, the toes and the conjunctiva (Eliezri *et al.*, 1990). Whether cancers at these sites are acquired sexually is unclear, although they do contain the same HPV types as found in the cervix and cases of co-existent cervical dysplasia and periungual dysplasia have been documented. Although tumours of the oral cavity and head and neck region in general have been associated with HPV infection, the association is much less, as is the correlation between sexual risk factors (oral sex) and tumours at these sites. In head and neck tumours, cigarette smoking and alcohol use are the most important risk factors.

Aetiology

As indicated above, the aetiological agent of cervical cancers and SILs is HPV. Of course, not all women exposed to HPV develop cervical SILs or cancer, but infection by the virus is prerequisite for SILs (and ultimately cancer) to develop. A great deal of attention has been given to the function of the different genes (called open reading frames (ORFs)) of HPV, to understand better what is the actual mechanism whereby this virus induces cervical SILs and cancers. HPV, as noted above, is a small virus and has seven early ORFs (E1–E7) and two ORFs that appear later in the infectious cycle (L1 and L2). As might be expected, the late ORFs are involved with production of the protein capsid coat that covers the virus immediately before it leaves the cell. Several of the early ORFs have been shown to be essential for the oncogenic effects of HPV using *in vitro* models. Specifically, ORFs E6 and E7 are capable of transforming normal squamous cells into cells that look identical with dysplastic squamous cells. However, E6 and E7 are not capable of making the normal squamous cells become invasive, that is, malignant. For this to happen, other molecular events must occur, as will be discussed below under Molecular Genetic Findings. It is now clear how E6 and E7 function. They both are capable of binding to and thus inhibiting the action of two important tumour suppressor gene products, p53 and Rb. Both p53 and Rb function to keep a cell's growth in an organized mode. If these proteins are blocked, the cell starts to proliferate at a much higher rate, which is evident under the microscope as hyperplasia. Further, the cells show enlarged, hyperchromatic nuclei that are the features of dysplasia on microscopic examination (Crum and Nuovo, 1991; Nuovo, 1994).

As noted above, there are about 20 HPV types that may be found in cervical SILs. It can be seen from **Table 1** that these types predominate in low-grade SILs. Fewer HPV types are found in the high-grade lesions and invasive cancers. This has an important implication when analysing for HPV.

Table 1 Correlation of HPV type and histology (%)

	HPV 6/11	HPV 16	HPV 18	HPV 31/35/51	HPV other[a]
Cervix					
Low grade SIL	18	34	3	28	17
High grade SIL	0	77	2	18	3
Cancer:					
Squamous cell	0	55	30	10	5
Adeno	0	31	66	2	1
Vulva/penis					
Low grade	96	1	0	0	3
High grade	0	92	0	7	1

[a] HPV other refers to HPV 33, 40, 41, 42, 43, 44, 45, 51, 52, 56, 68 and 70.

Screening and Prevention

Screening for cervical SIL is an excellent example where routine prevention has greatly reduced the incidence of a cancer. In the United States where cytological screening is common, cervical cancer is a relatively rare disease, with an incidence of about 15 000 cases per year, which is far below the rate for other cancers in women, such as breast cancer and lung cancer. However, in other countries where routine cytological screening is not done, such as in Columbia, cervical cancer is the leading cause of cancer death in women. Of course, the screening test for cervical SILs is the Papanicolaou smear (Pap smear). However, a newer test has been proposed for the purpose of screening for cervical SILs, namely HPV testing (Crum and Nuovo, 1991; Nuovo, 1994).

The function of the Pap smear is to have a cytotechnologist examine under the microscope a large sample of the surface cells of the cervix. It is very important to understand the histology and dynamics of the cervix. The cervix contains two distinct types of epithelium, the squamous cell and the glandular cell. The glandular cell layer is one cell thick, whereas the squamous cell layer is many cells (usually around 10–15 layers) thick. The thicker the epithelium, the whiter it appears on gross or macroscopic examination, owing to the masking of the underlying blood vessels. Hence the outer portion of the cervix (called the ectocervix), lined by squamous cells, is white, whereas the inner portion of the cervix (called the endocervix), lined by glandular cells, is pink. The key concept to remember is the transformation zone, which is the area where the squamous epithelium and glandular epithelium meet. At the transformation zone, the squamous cells replace the glandular cells owing to a process called squamous metaplasia (**Figure 3**). In that HPV usually initially infects the metaplastic squamous cells, it follows that cervical SILs originate in the transformation zone. With this realization, one can see why it is so important that the clinician take a sample from the transformation zone when a Pap smear is done. To document that this has happened, the cytotechnologist looks for either the metaplastic squamous cells or the endocervical glandular cells, because if he/she sees the latter, it is assumed that the clinician must have also sampled the transformation zone. Failure to see either cell type raises the question of whether the transformation zone was not sampled, and if so it is more likely that if a SIL was present, it may have been missed (Crum and Nuovo, 1991; Nuovo, 1994).

Figure 3 depicts the normal epithelial cells that are seen on a Pap smear. These include two types of mature squamous cells, called the superficial and the intermediate cells. These terms refer primarily to cytoplasmic features, such as colour, which are under the influence of oestrogen and progesterone, in the cervix and to a greater degree in the vagina. The parabasal cell is rarely seen in the Pap smear, as in most cases it is not present near the surface and hence not likely to be sampled. The parabasal cell can be seen in atrophy, where the mature squamous cells are much reduced in number owing to the marked reduction in the amount of oestrogen and progesterone. The metaplastic squamous cell was described above. Finally, the glandular cells, often present in well-defined groups that resemble a honeycomb, should be evident on a Pap smear (**Figure 3**). Other cell types are commonly found in the Pap smear. These include several types of inflammatory cells, such as the neutrophil, lymphocyte and macrophage. Large numbers of neutrophils at times indicate acute cervicitis, which can be due to many causes including infection by *Candida* or *Trichomonas*. HPV does not cause an acute infection.

Most laboratories that deal with Pap smears taken as a screening test from primarily premenopausal women report that around 90% of the smears are within normal limits or negative for malignant cells. It is important to stress that findings such as inflammation and the associated squamous cell changes (**Figure 3**), often called reparative or reactive, are considered within normal limits at least in the context of precancers of the cervix (Crum and Nuovo, 1991; Nuovo, 1994). Of course, if there is severe inflammation and an organism such as *Trichomonas* or *Candida* is identified, this needs to be reported, but even such cases may be signed off as ‘negative for malignant

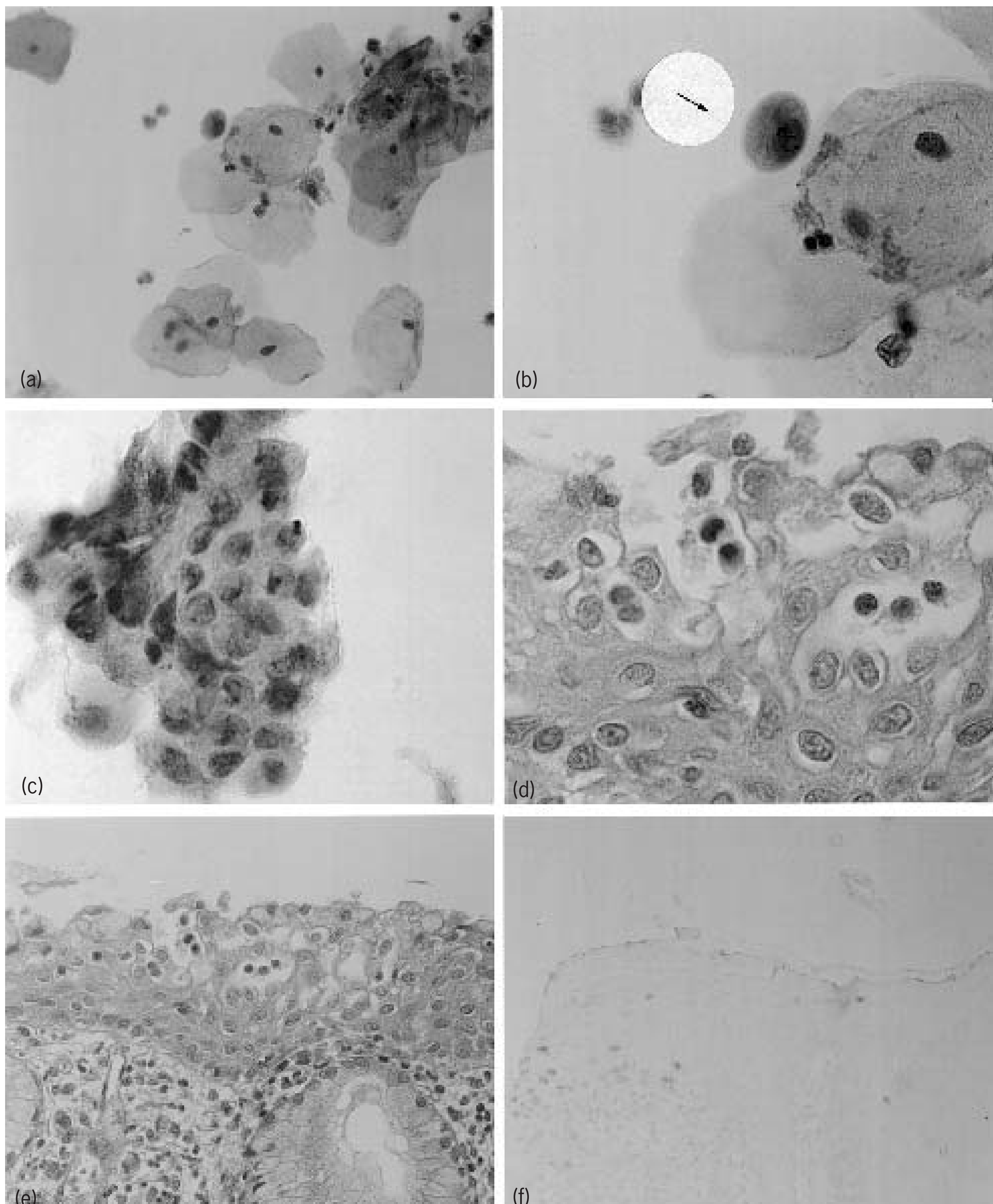


Figure 3 Morphological findings in the normal cervix. Panel (a) shows the squamous cells of a normal Pap smear; note the large amount of cytoplasm and small nucleus. Panel (b) shows the benign metaplastic cell with its small vacuoles (arrow). In panel (c), the honeycombed appearance of benign endocervical cells is apparent. The corresponding histology shows squamous metaplasia ((d) and, at lower magnification, (e)) which, as expected, are negative for HPV using *in situ* hybridization (f). Note that the nuclei are very uniform in squamous metaplasia.

cells,' with a comment that severe inflammation and a specific organism are identified. About 5–8% of Pap smears will be called ASCUS (atypical squamous cells of undetermined significance). This means that the cytological changes suggest but are not diagnostic of SIL. The remainder of Pap smears are diagnosed as SIL, most of these being low-grade SILs.

The diagnosis of low-grade SIL on a Pap smear is made on the basis of two cytological changes: (1) a large, well-defined perinuclear halo surrounded by a thin, clear-cut rim of cytoplasm; (2) enlarged nuclei that show hyperchromaticity (that is, increased darkness on the Pap stain). An example of low-grade SIL is provided in **Figure 4**. Note that the nucleus, although increased in size, is still surrounded by ample cytoplasm.

The diagnosis of high-grade SIL on a Pap smear is made on the basis of two cytological changes: (1) a high nuclear to cytoplasmic ratio (2) a marked increase in the chromaticity of the nucleus. An example of high-grade SIL is provided in **Figure 5**. Note that the nucleus shows an irregular distribution of chromatin and that, overall, there is a substantial increase in chromaticity. It should be added that irregularities in the contour of the nuclear membrane are also considered a diagnostic feature of high-grade SILs, and may also be seen in low-grade lesions. It is considered that the increased pressure that the cells are subjected to in the preparation of the ThinPrep makes such nuclear membrane irregularities common, including in many cases of cells that are clearly not SIL (G. J. Nuovo, unpublished observations). Hence, although abrupt changes or notches in the nuclear membrane are a feature of SIL in the conventional Pap smear, it is considered that they are not a reliable marker of such in the ThinPrep smear.

When the classical changes of low-grade or high-grade SIL are evident on a Pap smear, the diagnosis can be made easily and without equivocation. The difficulty arises when the Pap smear shows features suggestive but not diagnostic of SIL, i.e. ASCUS. There are four conditions which are the cause of most cases of ASCUS: reactive changes due to inflammation in mature squamous cells (mimic of low-grade SIL), reactive changes due to inflammation in immature metaplastic cells (mimic of high-grade SIL), SIL (usually low-grade) and atrophy. Let us examine each of these conditions.

Reactive changes in mature squamous cells are probably the most common cause of ASCUS. Inflammation in the cervix is common and can be due to many agents, such as *Candida* and *Trichomonas*, although in many cases a specific aetiological agent cannot be identified. Whenever there is inflammation, the mature squamous cells will usually demonstrate two cytological changes: a small perinuclear halo which does not show a clear-cut condensed rim of cytoplasm and a slightly enlarged nucleus. However, the enlarged nucleus does not show enough hyperchromaticity to warrant a diagnosis of low-grade SIL. Often, the nucleus will show a nucleolus and/or a nuclear groove, which are useful clues that one is dealing with reactive changes and not true SIL. The difficulty arises when inflammation occurs in the setting of a SIL. It takes experience to be able to 'factor out' the reactive changes in the cells and decide if there still is adequate atypia to warrant a diagnosis of SIL.

Reactive changes in immature metaplastic cells are also a common cause of ASCUS, especially in the ThinPrep sample. It will be recalled that squamous metaplasia is a ubiquitous process in the cervix. When there is associated inflammation, which is also common, the metaplastic cells will demonstrate an increased nuclear to cytoplasmic ratio. It is important to realize that metaplastic cells commonly show darkened nuclei, which of course is also a feature of high-grade SILs. The most reliable way to differentiate high-grade SIL from reactive immature squamous metaplasia is to realize that the latter will show a strong uniformity in the nuclear chromaticity from one cell to the next (**Figure 3**). It is just as important to realize that there are cases when it can be difficult to decide whether the chromaticity is uniform or variable enough to warrant a diagnosis of high-grade SIL versus reactive squamous metaplasia – in such cases, of course, the diagnosis of ASCUS is completely appropriate.

Atrophy is characterized by an increase in the nuclear to cytoplasmic ratio, but not perinuclear halos. Thus, it can mimic a high-grade SIL. The key is to realize that atrophic cells will not have hyperchromatic nuclei.

A common cause of ASCUS is SIL. In some cases, there simply are not enough dysplastic cells for the cytopathologist to make an unequivocal diagnosis of SIL. This may well be a sampling issue, where either the lesion was too small and not enough cells were present for a definitive diagnosis or the SIL was not adequately scraped during the procurement of the Pap smear. In other cases, the cytological changes are not considered to be clear-cut enough for this diagnosis; this often is the case for low-grade lesions. Indeed, it has been well documented that low-grade SILs due to the so-called benign HPV types 6 and 11 often do not show the classical cytological features of low-grade SIL but rather may be interpreted as normal cells, or more commonly, ASCUS (Nuovo, 1994). How does one differentiate true SILs, either low or high-grade, from their mimics? Although some groups have tried to make this distinction on cytological criteria, or using immunohistochemistry for nonspecific markers of increased cell proliferation, in the authors' opinion these variables are not sensitive enough. There is a very sensitive and specific way to make the distinction of ASCUS – benign versus true SIL – namely by HPV testing.

HPV testing is often done with a method called the hybrid capture technique. This is a very sensitive test that is similar to the older dot blot hybridization assays. In these tests, the cells are destroyed and their DNA is retrieved. A sample of the DNA is then hybridized with HPV DNA or RNA probes, and the complex is detected in a variety of

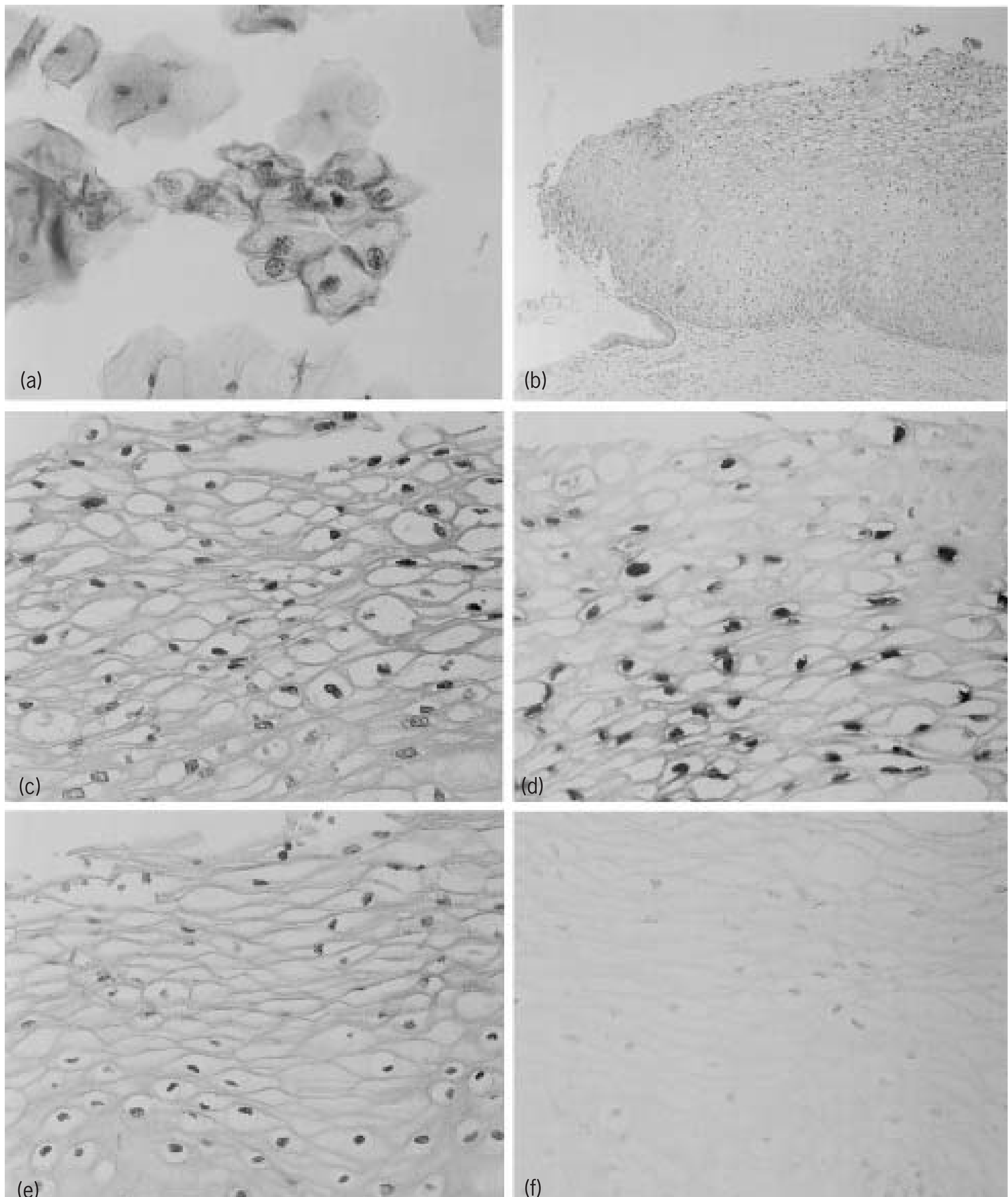


Figure 4 Morphological findings in low-grade SIL. Panel (a) shows the classical large, well-defined perinuclear halos and enlarged nuclei of a low grade SIL. Panels (b) and (c) (higher magnification) show the variable cell density, varying sized and shaped halos and nuclear variability towards the surface that are characteristic of the disease. The lesion contains a large amount of HPV 51 ((d), dark-staining cells mark HPV). Note the well-ordered appearance of the adjacent epithelium (e) that is HPV negative by *in situ* hybridization (f).

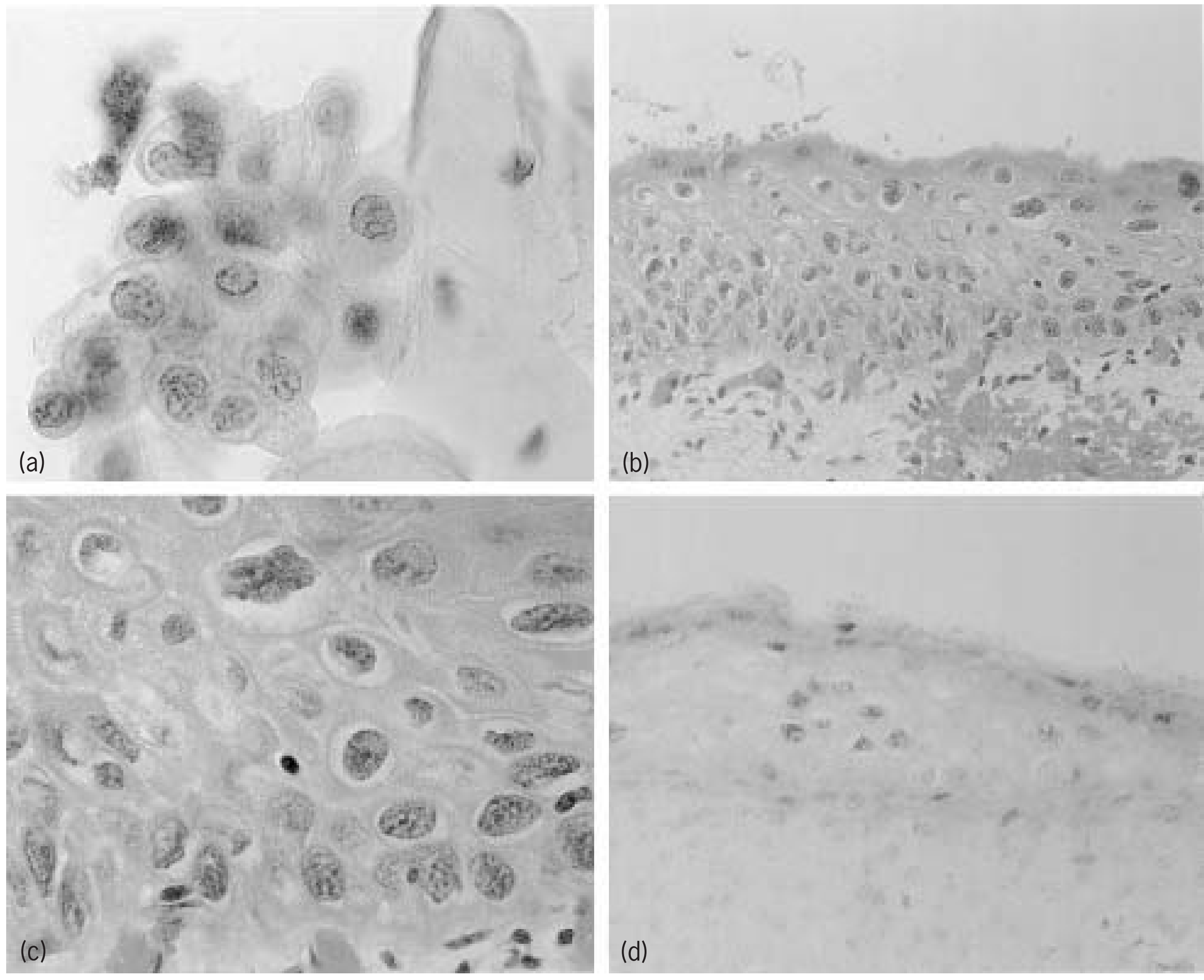


Figure 5 Morphological findings in high-grade SIL. Panel (a) shows the Pap smear; note the hyperchromaticity of the cells and the sharp, abrupt indentations in the nuclear membrane. The biopsy shows a relatively thin epithelium (b) that contains highly atypical cells showing variable cell density and concomitant nuclear atypia (c). HPV 16 is present (d), but in smaller numbers compared with the low-grade SIL (**Figure 4**).

ways (formerly using radioactivity, now more commonly using enzymes complexed to the HPV probes that either cause a colour or light emission to mark the presence of HPV in the sample) (Nuovo, 1997). If enough HPV probes are included to detect most of the 20 HPV types that can be found in the cervix (see **Table 1**), then this system will detect HPV in over 90% of cases of SIL, even if they are called ASCUS. The problem with this test is that they also detect HPV in the absence of ANY cytological changes. That is, about 15–20% of women will have HPV detected with this test, even if there Pap smears are completely normal. Thus, such tests have relatively low specificity. In that it is not clear what the significance is of detecting HPV in the setting of a normal cervix (i.e. a cervix without SIL), and most clinicians do not recommend treatment when HPV is found in this setting, it would be useful to have a test with high sensitivity, but better specificity so as not to detect HPV when no SIL is present.

HPV detection by *in situ* hybridization offers certain advantages over the more sensitive techniques of polymerase chain reaction (PCR) or dot blot hybridization/hybrid capture. First, the HPV is directly detected within intact cells, and thus one can correlate the cytological findings with the viral results. Second, with a probe cocktail that can detect most HPV types, over 90% of low-grade SILs will be positive. Finally, it is very rare (less than 1%) to detect HPV by *in situ* hybridization in a normal-appearing cell from the cervix (Crum and Nuovo, 1991; Nuovo, 1994). We undertook a study that tested all cases of ASCUS for HPV by *in situ* hybridization where there was a biopsy within 6 months of the Pap smear. The data are presented in **Table 2**. Note that the detection of HPV by *in situ* hybridization was an excellent way to differentiate cases of ASCUS associated with a biopsy where only benign reactive changes were seen (HPV negative) from those where an actual SIL was evident on biopsy (HPV positive).

Table 2 Correlation of detection of HPV by *in situ* hybridization with the cytology and, for cases of ASCUS, clinical outcome

Pap smear diagnosis	HPV detection
Normal	1/25 (4%)
SIL	22/25 (88%)
ASCUS (Total)	20/40 (50%)
ASCUS with biopsy proven SIL within 6 months	17/21 (81%)
ASCUS with biopsy negative for SIL within 6 months	3/19 (16%)

Gross/Histopathology/ Immunohistochemistry/EM

The gross findings of the cervix are viewed under magnification using a device called the colposcope. Colposcopy typically involves looking at the cervix from 4× to 25× after applying a weak solution of acetic acid. The acetic acid solution will remove the mucous and highlight any areas where there is increased cellular density which will appear dark white (hence acetowhite lesions). Recall that the outer part of the cervix nearest the vagina (the portio) is white and the endocervix pink because the former is covered by a thick multilayer of squamous cells which masks the underlying blood vessels, which is responsible for the pink colour of the endocervix. It is important to realize that any process which causes a focal increased density of cells in the cervix, be it SIL, inflammation or squamous metaplasia, will produce an acetowhite patch. Hence the presence of such is simply an indicator for the colposcopist as to where to biopsy, and is not diagnostic of a SIL. To diagnose a SIL, a biopsy must be done.

The histological findings of low- and high-grade SIL are shown in **Figures 4** and **5**. Note that in either case, one uses the presence of a disorganized growth pattern (i.e. variable cell density, where some of the cells are closely packed and in other areas some of the cells are widely spread apart) as the most useful feature to diagnose a SIL on biopsy. It is important to stress that this is a low-power (4× or 10×) pattern to view under the microscope. Newcomers to the field often make the mistake of focusing their attention on the cytological details at high magnification. Although useful to see hyperchromatocity, one can usually make the diagnosis on histological grounds at low magnification, using the organisation of the cells as the key variable.

The immunohistochemical analysis of SILs is for the most part non-specific. That is, one can detect in greater number certain proteins that are indicative of the increased cellular proliferation evident in SILs, such as Ki-67. As indicated above, there is one highly sensitive and specific immunohistochemical stain for SILs, namely HPV detection. However, HPV detection via the protein coat of the virus is not very sensitive, as in many cases the virus does not produce enough of its capsid for detection by this method, especially in high-grade SILs and cancers (Crum and Nuovo, 1991; Nuovo, 1994). HPV DNA detection by *in situ* hybridization on biopsies, as on Pap smears, is very sensitive and specific. **Figures 4** and **5** show examples of the detection of HPV by *in situ* hybridization in SILs and cervical cancer. Two important points need to be stressed. First, there are cases where the histological changes are suggestive but not diagnostic of low-grade SIL. These are often biopsies where there is a lot of inflammation and where the Pap smear is called ASCUS. In these instances, HPV testing by *in situ* hybridization is a very reliable way to differentiate those tissues that actually are SIL from its mimics. The importance of this cannot be overstated, given that an over-diagnosis of SIL can cause serious emotional distress for the patient. In the authors' opinion, such equivocal cases should be signed off with the HPV *in situ* test; positive is low-grade SIL and negative is nonspecific reactive changes, negative for SIL (Crum and Nuovo, 1991; Nuovo, 1994). Second, the number of virus particles in high-grade SILs and, especially, invasive cancers is much reduced relative to low-grade SILs. To detect the virus reliably in invasive cervical cancer, one must use either PCR or, if one wishes to localize the virus to a specific cell type, PCR *in situ* hybridization.

Electron microscopy findings in SILs are often nonspecific; actual viral protein is detected occasionally in low-grade SILs as it is in this condition where numerous viral particles may be made. In low-grade SILs, electron microscopy shows that the perinuclear halo consists of a clear zone where the cytoplasm and cell organelles have been pushed to the outer aspect of the cell. Electron microscopy is not a useful method for detecting HPV infection of the cervix.

Molecular Genetic Findings

The molecular genetic findings of SILs centre around two important onco-proteins made by the virus, called E6 and E7. These proteins are capable of inhibiting two important proteins, specifically p53 and Rb, respectively. The E6 and E7 proteins from the oncogenic HPV types, such as HPV 16 and 18, are much more avid in their inhibition of p53 and Rb than the benign HPV types, such as HPV 6 and 11. However, it is now clear that other suppressor gene products, that are not apparently inhibited at all by HPV, are more commonly inhibited in the evolution of cervical cancer. Specifically, p16 is inactivated in nearly all cells in high-grade SILs and cancers, whereas such lesions contain many cells where p53 and Rb are still active, as the virus did not produce either E6 or E7 in those specific cells (Nuovo *et al*., 1999). Finally, it is now clear

that certain host RNAs are selectively increased or decreased in production as lesions progress from SIL to cancer, and again this appears to be independent of HPV interactions. Matrix metalloproteases (MMPs) are enzymes that help cells digest collagen, as is needed in healing of tissue damage. Cervical cancer cells make more MMPs and less of their inhibitors (TIMPS) as they progress from microinvasive cervical cancer (with its good prognosis) to invasive and metastastic cervical cancer (Crum and Nuovo, 1991; Nuovo, 1994).

Prognostic Factors

The primary prognostic marker of cervical cancer is its stage, which is how far it has infiltrated local tissues. It is important to stress that neither HPV type nor histological/cytological findings can accurately predict which SILs will progress and which will regress. The only exception to this statement is the observation that low-grade SILs that contain HPV 6 or 11 usually, if not always, regress.

Overview of Present Clinical Management

About 10 years ago, the treatment for cervical SILs was a simple office procedure called cryotherapy. The key clinical factor was whether the entire lesion could be visualized at colposcopy. If it could not, then the transformation zone of the cervix had to be removed surgically (called cone knife cone biopsy) to document that the lesion was not invasive in the canal and to remove it completely. However, if, as in most cases, the SIL could be seen in its entirety with the colposcope, then a simple 5-min office procedure could eradicate the virus and the lesion in about 90% of women. Over the last 10 years, this has shifted from cryotherapy to using laser ablation, which is equally effective but more expensive. Over the last several years, many gynaecologists have switched to a modified cone biopsy (usually smaller) using a metal loop and electrocautery procedure (LEEP). This procedure usually removes the entire lesion, but may have to be done in an operating room under general anaesthesia. The recurrence rates after LEEP are similar to those seen after cryo- or laser therapy (Crum and Nuovo, 1991; Nuovo, 1994).

The management of cervical cancer depends on stage. For microinvasive cancer, a cone biopsy or LEEP is curative. For more deeply invasive cancers, either surgery plus chemotherapy or, at times, radiotherapy are indicated. Death is usually due to renal failure secondary to obstruction of the ureters by the tumour. Studies using vaccines against HPV to treat cervical cancer are too preliminary but do not appear to be very effective, perhaps reflecting the decreased role of HPV at this advanced stage.

PATHOLOGY OF THE ENDOMETRIUM

This part of the chapter will focus on tumours of the endometrium and myometrium. In order to understand endometrial pathology, it is important to appreciate the effects of oestrogen and progesterone on the normal endometrium. It is the balance of these hormones that allows for normal function of the endometrium such as preparation for embryo implantation and normal menstrual cycles. Oestrogen stimulates proliferation (and thus mitotic activity) in the glands and stroma. Progesterone inhibits mitotic activity, and rather stimulates secretions in the glands and changes in the stroma called predecidualization, where the cells acquire more cytoplasm. The key point to remember is that unopposed oestrogen will induce continued mitotic activity and gland growth, which is a central point in the development of endometrial tumours (Deligdisch, 2000).

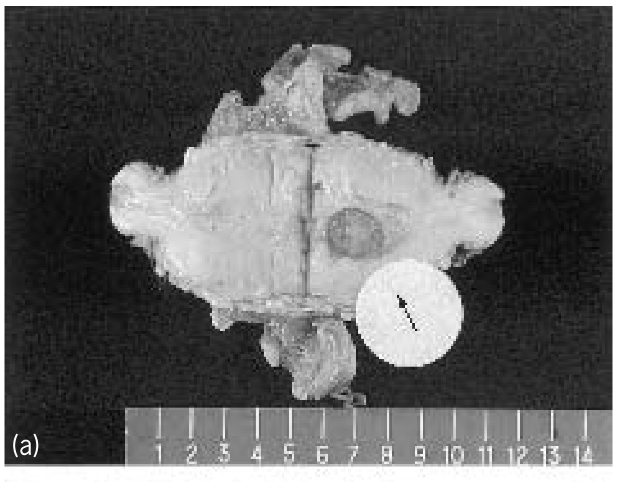
Tumour Pathology

A variety of tumours and tumour-like conditions can occur in the uterus. **Table 3** gives a list of the more commonly encountered tumours in the uterus.

Endometrial polyps are benign, localized lesions which protrude into the endometrial cavity (**Figures 6 and 7**) and represent focal hyperplasia of the endometrium. Variable amounts of endometrial glands, fibrous stroma and blood vessels will be present. If a significant amount of smooth muscle is present, the polyp is referred to as an adenomyomatous polyp. Polyps are common, occurring most frequently in women 40–50 years old; the presenting symptom is often abnormal uterine bleeding. Tamoxifen

Table 3 Commonly encountered uterine pathology

Benign endometrial lesions
Endometrial polyp
Tamoxifen-associated polyp
Adenomyomatous polyp
Decidual polyp
Simple hyperplasia
Premalignant lesion
Complex endometrial hyperplasia
Malignant endometrial lesions
Type I (oestrogen-associated)
Endometrioid adenocarcinoma
Type II (oestrogen-independent)
Papillary serous carcinoma
Clear-cell carcinoma
Benign myometrial tumour
Leiomyoma
Uterine sarcomas
Mixed mullerian tumour
Leiomyosarcoma
Endometrial stromal sarcoma



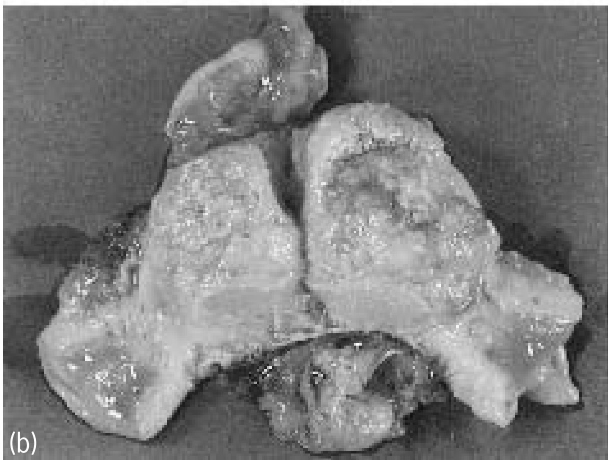


Figure 6 Macroscopic appearance of uterine polyps. Panel (a) shows the macroscopic appearance of a benign polyp (arrow). Panel (b) shows a polypoid endometrial adenocarcinoma that protrudes into the endometrial cavity. Note that the carcinoma is larger, has a less uniform appearance, and has foci of haemorrhage and necrosis.

(an antioestrogen used to treat breast cancer) is clearly linked to the development of endometrial polyps (Nuovo, 1994). Relative progesterone excess can cause a stromal proliferation evident clinically as a decidual polyp (**Figure 7**, which also shows the different histologic features seen in endometrial polyps).

While polyps represent a localized overgrowth of endometrium and stroma, certain precancerous lesions affect the endometrial cavity, diffusely leading to a thickened endometrium, which can simulate a tumour. Endometrial hyperplasia, a condition in which the proportion of endometrial glands is increased relative to endometrial stroma, is a well-defined step in the development of the most common types of endometrial cancer. Endometrial hyperplasia develops under conditions of oestrogen excess. The single most important factor in determining the likelihood of progression of hyperplasia to

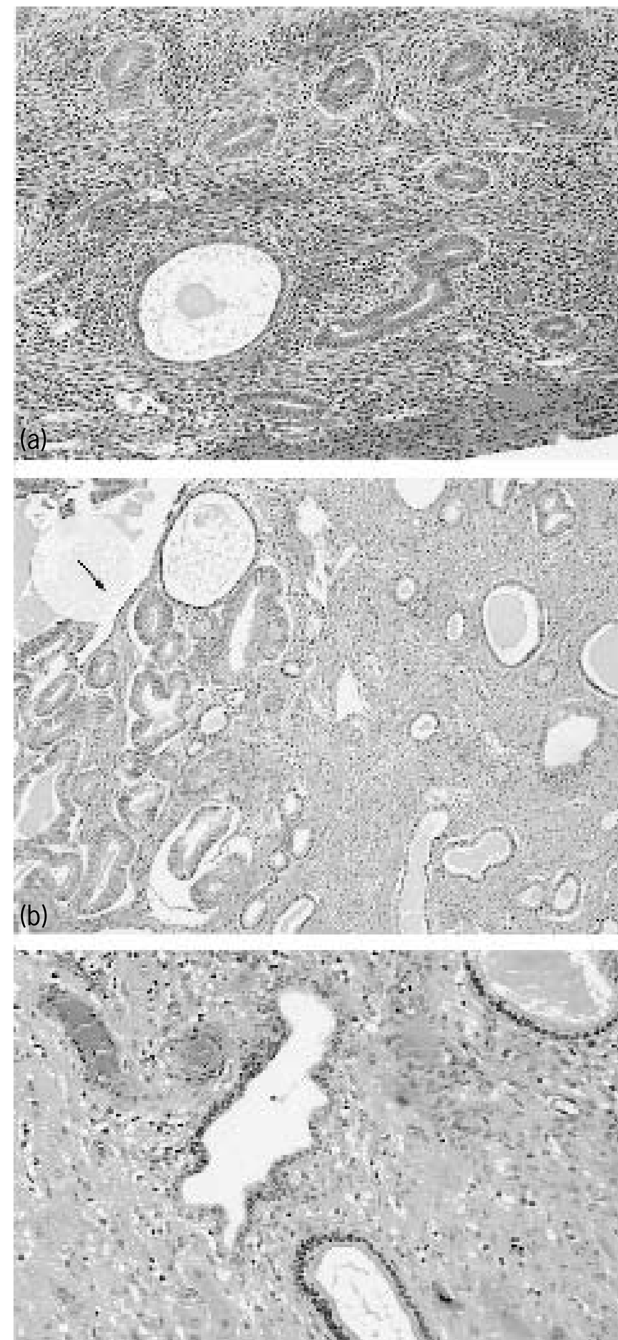
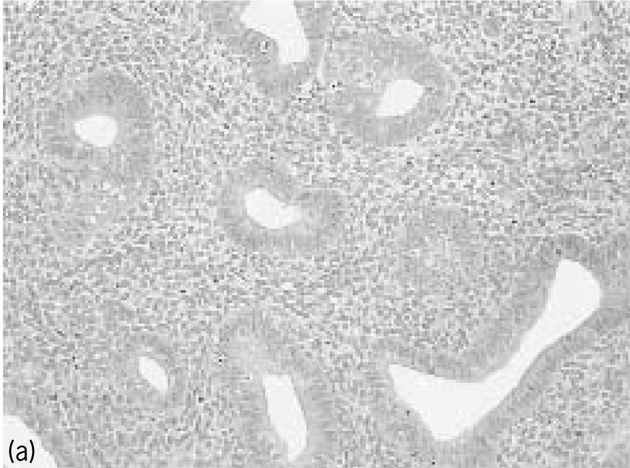


Figure 7 Microscopic appearance of uterine polyps. Note the fibrotic stroma, endometrial glands, which are relatively increased compared to the stroma, and the thick-walled blood vessels characteristic of a benign endometrial polyp (a). In panel (b) (an endometrial polyp from a woman taking tamoxifen for breast cancer), note that the glands are dilated and there is a small focus of endometrioid adenocarcinoma (arrow). This is a common finding in endometrial polyps associated with tamoxifen. Panel (c) shows a so-called decidual polyp, due to the effects of progesterone. Note how the stromal cells are large, plump with ample eosinophilic cytoplasm.

carcinoma is the presence of atypia, both cytological and architectural. Whereas hyperplasia-containing glands with simple architecture with normal gland cytology (simple hyperplasia) has been associated with a 1% cancer progression risk, hyperplastic glands with complex architecture and cytological atypia (complex atypical hyperplasia) has a 29% cancer progression risk (Burke *et al*., 1996). **Figure 8** shows the histological spectrum of endometrial hyperplasia.

Tumours of the endometrium account for 95% of uterine neoplasms (Burke *et al*., 1996). The endometrioid subtype of endometrial adenocarcinoma accounts for nearly 90% of endometrial carcinomas (Burke *et al*., 1996). Endometrioid adenocarcinoma most commonly appears as a large polypoid lesion protruding into the endometrial cavity and may be detectable as an enlarged uterus on bimanual examination. There may be associated haemorrhage and necrosis (**Figure 6**). Histologically, well-differentiated tumours resemble normal endometrium but contain very complex glandular architecture, lack intervening endometrial stroma, and have cytological atypia. Areas of confluent, malignant cells may be seen as tumours become less well differentiated, and it is the percentage of these areas which are evaluated when assigning grade. Tumours with > 50% solid growth are considered poorly differentiated, or high grade. It is not uncommon for endometrioid adenocarcinomas to show areas of squamous or, less commonly, mucinous differentiation. It is important to realize that this does not adversely affect prognosis. Also, areas of squamous differentiation are not included when considering the percentage of solid growth within a tumour (Burke *et al*., 1996). **Figure 9** shows the histological features of endometrioid adenocarcinoma and some of its variants/ mimics.



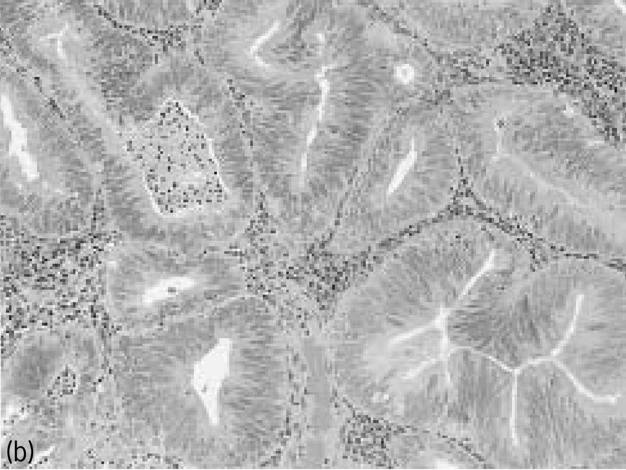


Figure 8 Histological appearance of endometrial hyperplasia. In simple hyperplasia, the endometrial gland to stroma ratio is increased. Glands have varying size, but maintain simple architecture and have no cytological atypia (a). Complex atypical hyperplasia has marked glandular crowding and, compared with (a), a markedly increased gland to stroma ratio. The important clues to the correct diagnosis include architectural atypia (glands within glands) and the associated nuclear atypia (b).

Less common endometrial carcinoma histological subtypes include papillary serous carcinoma (3–10%) and clear cell carcinoma (0.8–5%). These patterns are histologically identical with their counterparts in the ovary and vagina/kidney, respectively. It is important to recognize these patterns as distinct entities because they tend to present at a more advanced stage and are associated with a more aggressive clinical course than endometrioid adenocarcinoma (Cirisano *et al*., 1999). Because these histological patterns have not been associated with a clear premalignant condition or excess oestrogen, have clearly more aggressive behaviour, and have also been shown to have distinctly different molecular alterations compared with endometrioid carcinomas (see below), it is hypothesized that these tumour types have divergent patterns of development. Consequently, endometrial tumours are divided into Type I carcinomas encompassing endometrioid/ oestrogen-dependent cancer and Type II carcinomas encompassing those histological subtypes not associated with oestrogen (Cirisano *et al*., 1999). **Figure 9** shows the histological features of Type II endometrial carcinomas.

Leiomyomas are the prototypical benign tumour of the uterine myometrium and are benign clonal neoplasms of smooth muscle. These tumours grossly appear as well-circumscribed, white, whorled nodules usually without haemorrhage or necrosis (**Figure 10**). Histologically, these tumours typically have uniform, bland spindle cells with few mitoses. This is in direct contrast to uterine sarcomas, which are usually large tumours with haemorrhage and necrosis displaying histological features of marked cellular pleomorphism and abundant mitoses. Uterine sarcomas represent only 5% of uterine malignancies. The malignant counterpart to leiomyoma is the leiomyosarcoma, which represent one third of uterine

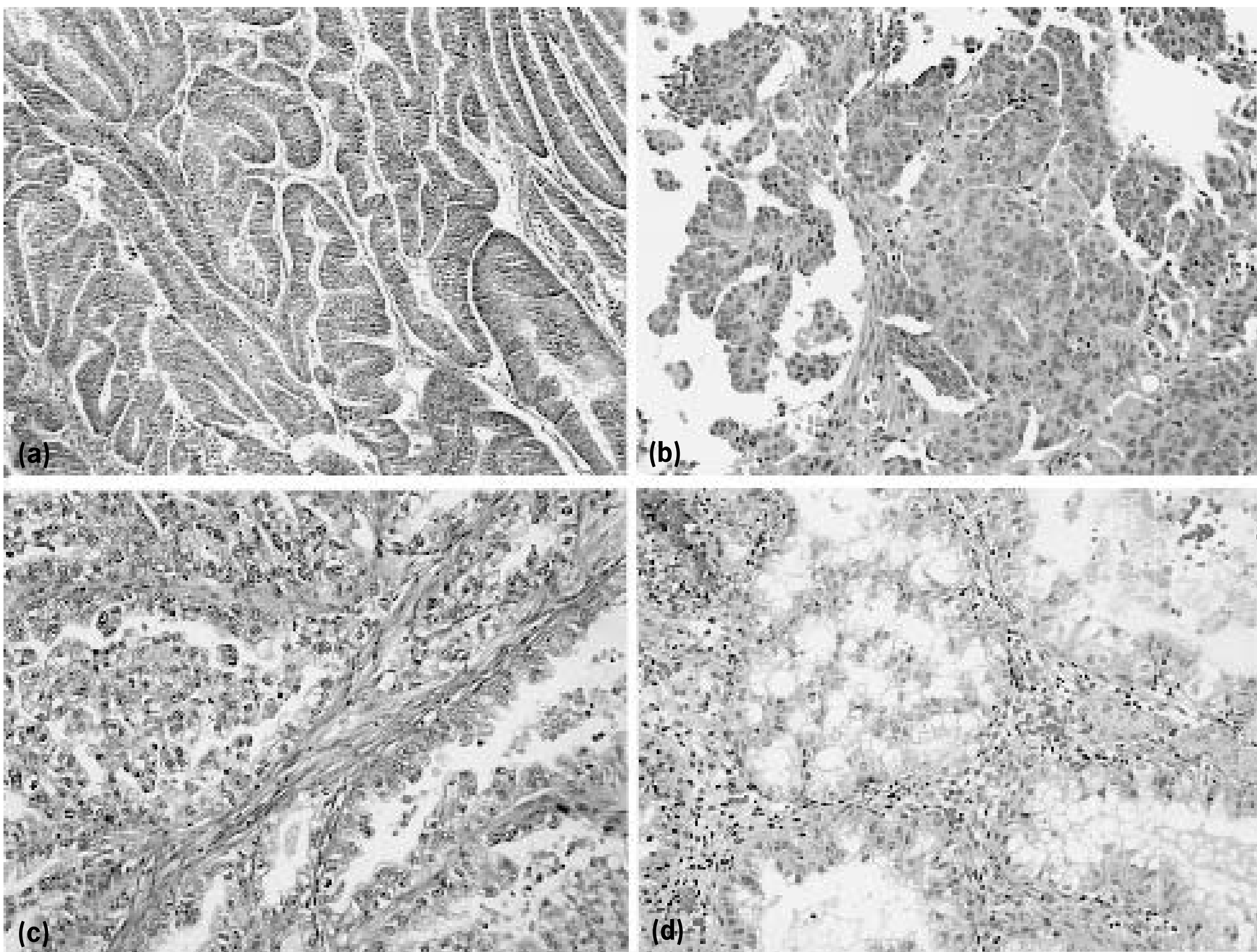


Figure 9 Histological appearance of endometrial adenocarcinoma. In well-differentiated endometrioid adenocarcinoma (grade FIGO I), the tumour is comprised of well-formed glands without solid growth (a). Panel (b) shows a rare and aggressive variant of endometrial carcinoma, papillary serous carcinoma; note the multiple small detached groups of cells characteristic of this entity. A clear-cell carcinoma has in some areas clearing of the cytoplasm with marked cellular atypia; note the 'hobnail' appearance of the nuclei (c). A mimic of clear-cell carcinoma is the Arias Stella reaction (d), characteristic of pregnancy. These cells lack the cytological atypia of clear-cell carcinoma and have larger vacuoles reflecting active secretion.

sarcomas (Levenback, 1996). Other types of uterine sarcoma include the mixed mullerian tumour (most common uterine sarcoma) and endometrial stromal sarcoma. The malignant mixed mullerian tumour (MMMT) is characterized by the presence of a sarcomatous element intimately admixed with a malignant glandular component. This sarcomatous component may be composed of elements normally found in the uterus such as smooth muscle or endometrial stroma (homologous elements), or it may contain elements not present in the uterus such as cartilage and bone (heterologous elements). **Figure 10** shows a comparison of benign and malignant myometrial tumours. Endometrial stromal sarcomas usually show a more bland appearance than leiomyosarcoma or MMMT; they are characterized by increased mitotic activity and invasion of the uterine wall.

Epidemiology and Aetiology

Because endometrial carcinoma, particularly the endometrioid subtype, is by far the most common malignancy of the uterus, the discussion on epidemiology and aetiology will pertain most directly to these tumours. Endometrial cancer is the most common gynaecological malignancy and represents the fourth most common malignancy overall in American women while being the eighth most common cause of cancer death (Burke *et al.*, 1996). The incidence of endometrial cancer peaks at 70–74 years of age, and the lifetime risk of developing endometrial cancer is estimated to be 2%. Into the mid-1980s, the incidence of endometrial carcinoma increased. This has been attributed to increased use of oestrogen replacement therapy; increasing age of the population may also be

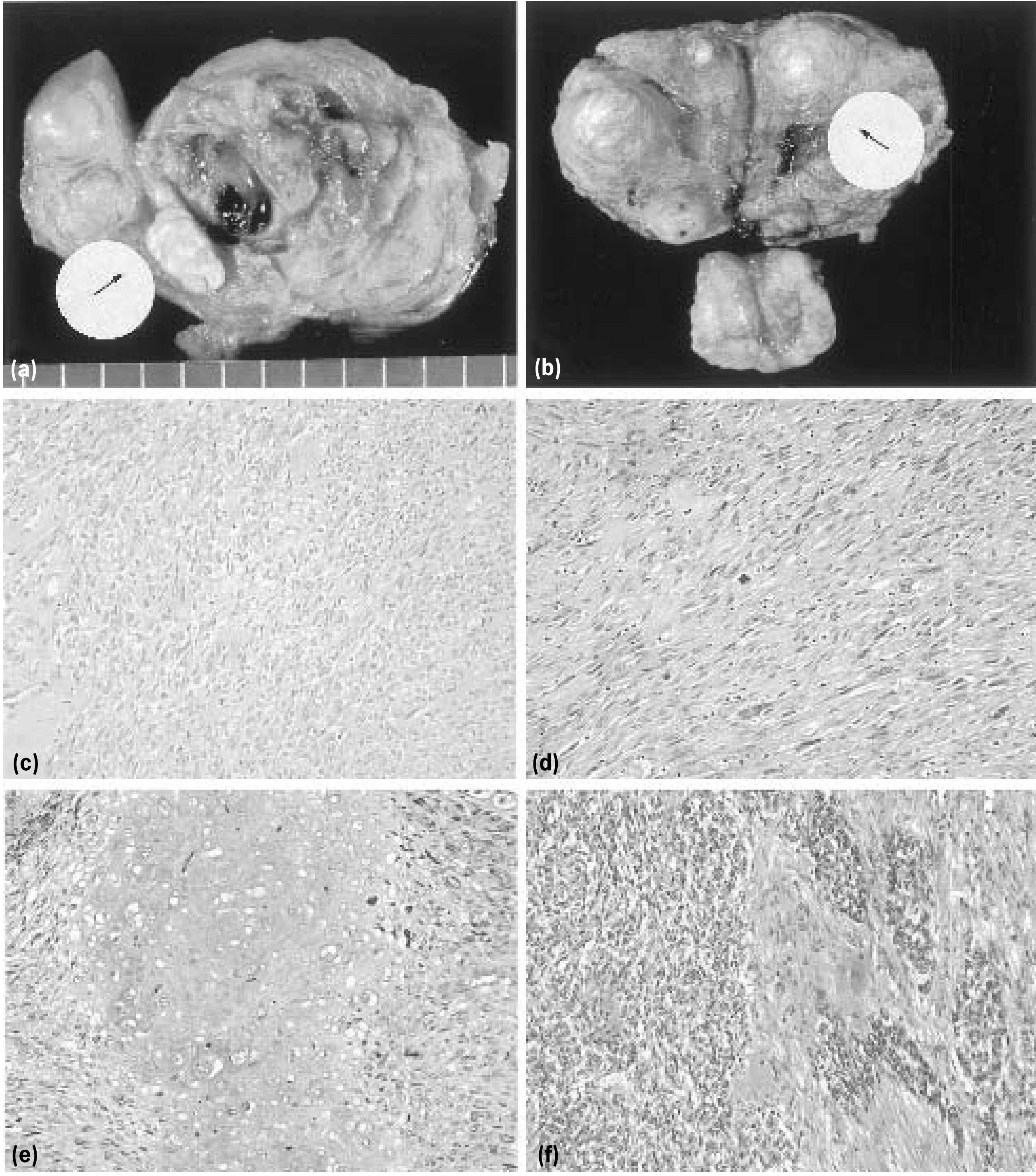


Figure 10 Stromal tumours of the endometrium. Panel (a) shows a benign leiomyoma with a smooth, well-defined tumour border. The cut surface is homogeneously white with a whorled texture (arrow). Panel (b) shows a leiomyosarcoma. In contrast, this tumour appears soft and friable with areas of haemorrhage and lacks a well-defined tumour border (arrow). The corresponding histology of the leiomyoma with bland spindled cells and no mitoses (c) is in contrast to the increased cellularity, marked cellular atypia, and the presence of mitoses of a leiomyosarcoma (d). A malignant mixed mullerian tumour with heterologous elements (cartilage) surrounded by malignant homologous stroma is shown in panel (e). Panel (f) shows an endometrial stromal sarcoma; such tumours usually show less cytological atypia than the other nonepithelial tumours of the endometrium.

related. More recently, a decline in incidence of endometrial carcinoma has been noted, which may be related to improved formulations in oral contraceptives with lower oestrogen content and to the addition of progesterone to menopausal hormonal replacement regimens (Burke *et al.*, 1996).

As previously discussed, a relationship between unopposed oestrogen and the development of endometrioid adenocarcinoma is well established. This was first suggested by the increased incidence of endometrial carcinoma with the advent of single-agent hormone replacement therapy with oestrogen. Other conditions that could expose the endometrium to unopposed oestrogen have been associated with an increased risk of developing endometrial cancer. Nulliparity is associated with the development of carcinoma, as are early menarche and late menopause. These three conditions allow for increased time of exposure of the endometrium to oestrogen. It is important to consider that nulliparity may be related to infertility due to anovulation. It is only after ovulation that progesterone impacts the endometrium. Therefore, anovulatory women have chronic overexposure to excess oestrogen. Obesity is a commonly cited risk factor and may be explained by an increased availability of unopposed oestrogen due to increased production of oestrone from androstenedione. Medical conditions such as diabetes and hypertension are also associated with increased cancer risk. It is not entirely clear whether they represent independent risk factors or whether their association with underlying morbid conditions such as obesity explains the association. A less common association with the development of endometrial cancer, secondary to the rarity of the tumour, is an oestrogen-secreting tumour such as a granulosa cell tumour of the ovary (Burke *et al.*, 1996). Finally, mention should be made of the association of tamoxifen with the development of endometrial carcinoma. Tamoxifen has an antioestrogenic effect in breast tissue and is therefore both a successful adjuvant therapy and a potential chemopreventive agent for breast cancer. In the uterus, however, tamoxifen is a partial oestrogen agonist leading to polyp formation, endometrial hyperplasia and carcinoma in some women (Nuovo, 1994).

Few risk modifiers have been described. Oral contraceptives with relative increased proportion of progestational agents may confer a decreased risk. Smoking has also been shown to decrease the risk of developing endometrial carcinoma and may be related to differences in oestrogen metabolism and earlier menopause (Burke *et al.*, 1996). It should be noted that the moderate decrease in endometrial cancer associated with smoking does not outweigh the risk of lung cancer, which is now the leading cause of cancer death in women in the United States.

The above discussion focused on endometrial cancers with Type I pathology. Cancers with Type II pathology do not have a relationship with unopposed oestrogen, and risk and lifestyle factors for these tumours have not been entirely elucidated. Precursor lesions to these cancer subtypes are currently being established, and certain oncogene mutations have been more strongly associated with the development of Type II than with Type I carcinomas. Family history with documented inherited predispositions to oncogene mutation is emerging as a risk factor for these tumour types.

Screening and Prevention

Although endometrial carcinoma is the most common gynaecological malignancy, no well-developed screening programme exists, as screening has not proven beneficial in the early detection of carcinoma. The majority of women present with postmenopausal bleeding or abnormal uterine bleeding if premenopausal and is subsequently discovered to have uterine carcinoma at an early stage (Burke *et al.*, 1996).

Some have attempted to detect preclinical disease in the asymptomatic patient and have suggested various screening options including the Pap smear, transvaginal ultrasound (TVUS) and endometrial biopsy (EMB). The sensitivity and specificity for the detection of endometrial carcinoma by Pap smear is low. In symptomatic patients, an abnormal Pap smear predicted correctly the presence of carcinoma in only 28% of patients (Nuovo, 1994). TVUS in the asymptomatic patient has no well established cutoffs for which a screen would be considered positive, although endometrial thickness <4 mm usually results in tissue insufficient for diagnosis on subsequent biopsy and is less likely to be associated with significant uterine pathology. Important, too, is the cost of the equipment and time involved (Burke *et al.*, 1996). Endometrial biopsy in the asymptomatic patient may be problematic as many elderly patients have cervical stenosis and an office pipelle biopsy may be technically impossible (Burke *et al.*, 1996). It has been suggested that endometrial biopsy in patients without postmenopausal bleeding be limited to patients on postmenopausal hormone replacement therapy or in those patients with endometrial cells or ASCUS on Pap smear or who are taking tamoxifen. Some advocate the use of tumour markers to detect preclinical disease. The most commonly used marker in gynaecological malignancies is CA-125. Unfortunately, CA-125 may be elevated in inflammatory states and also in benign gynaecological conditions such as endometriosis. It has been reported as a useful marker to detect early disease recurrence in patients who have elevated CA-125 prior to definitive treatment. This is particularly true of papillary serous carcinoma (Kurman *et al.*, 1995).

In patients who have symptoms, i.e. abnormal bleeding, TVUS can be a useful tool. An endometrial stripe thickness of >10 mm predicts the presence of significant uterine pathology in 10–20% of patients. When technically

possible, EMB is most frequently performed to evaluate abnormal bleeding. It has the advantage of being an office procedure, and although only 5% of the endometrium is sampled, EMB has a 97% sensitivity and an 83% specificity (Chen *et al*., 1999). The gold standard for the evaluation of symptomatic women remains hysteroscopy with D&C. However, this requires general anaesthesia. The advantage is that 60% of the endometrium is sampled and there is direct visualization of any polypoid lesions in the endometrial cavity (Chen *et al*., 1999).

Little is written about prevention. Maintenance of ideal body weight and addition of progestational agents in hormone replacement regimens appear to be practical methods of prevention based on currently understood risk factors.

Gross/Histology/Preinvasive/ Immunohistochemistry

The gross and histological features of uterine tumour pathology and preinvasive lesions have been discussed in a previous section. This part of the chapter will focus on ancillary techniques of tumour diagnosis and histological mimics of uterine tumours.

Immunohistochemistry has become a useful tool in detecting markers of cellular differentiation in order to assign a histological subtype when the answer is not obvious by usual microscopic examination. Expression of immunohistochemical markers may also be used as prognostic indicators for tumour behaviour. Because the endometrium is a hormonally modulated by oestrogen and progesterone, it is logical that receptors for these hormones would be present. Oestrogen and progesterone receptors are abundant in benign endometrium but show a decreased presence in endometrioid malignancies by immunohistochemical analysis. This makes sense because malignancies are not under the same regulatory influences as benign tissues. Well-differentiated cancers are more likely to be receptor positive; they more closely resemble their tissue of origin. It has been suggested that progesterone receptor positivity is associated with better-differentiated tumours with less aggressive biological behaviour. In theory, progesterone administration as adjuvant therapy for patients with progesterone positive tumours could be useful in counteracting the effects of oestrogen as in the normal endometrium (Kurman *et al*., 1995).

There are times when endometrial adenocarcinoma may be difficult to distinguish from endocervical adenocarcinoma, particularly when endometrial cancer extends into the cervix. In the past, some have used carcinoembryonic antigen (CEA) as a distinguishing tool, citing that there is positive staining in 80% of endocervical adenocarcinomas and 8% of endometrial carcinomas (Kurman *et al*., 1995). The distinction is important because, of the two, endometrial carcinoma has a better prognosis. Recently, because endocervical cancers have been associated with HPV (particularly subtype 18), studies have been made to determine whether there is a difference in HPV expression and determined that endometrial carcinoma is not associated with HPV, and is thus a more reliable way to differentiate endocervical from endometrial cancers, as discussed in detail above in the section on cervical cancer (Nuovo, 1994).

It is important to be aware of benign conditions that can simulate malignancy. The classical example is the Arias Stella reaction, which can look nearly identical with clear cell carcinoma (**Figure 9**). It is most commonly associated with pregnancy but has been reported to occur with use of certain medications (Kurman *et al*., 1995). A good history and the presence of progesterone-related endometrial stromal changes (decidualization) are invaluable in making this distinction; if the Arias Stella reaction is due to miscarriage, villi may be present.

Molecular Genetic Findings

The molecular aspects of endometrial carcinogenesis continue to be elucidated. Endometrioid tumours are influenced by oestrogen and progesterone, and the development of endometrial hyperplasia is directly related to excess oestrogen. These hormones have effects on cellular proliferation; therefore, it is possible that molecular disturbances in the cell cycle contribute to the development of carcinoma. How transformation from premalignant lesions to cancer at the cellular level is still not known.

p53 has been studied extensively in the endometrium. It is involved in many crucial cellular functions such as cell cycle regulation, DNA repair, cellular differentiation and apoptosis (programmed cell death). It has been shown that there are alterations in the function of p53 in Type II, or nonoestrogen-dependent tumours, probably due to gene mutation. Oestrogen-dependent tumours (Type I) and their precursors rarely demonstrate *p53* mutations. The functional alteration of p53 in Type II tumours along with their lack of association with unopposed oestrogen suggests the existence of an oestrogen-dependent pathway in endometrial tumorigenesis. In non-endometrial tumours, *p53* mutation is associated with poor differentiation, aggressive clinical course and poor prognosis. The fact that most Type II tumours have altered p53 function explains at the molecular level why this subset of endometrial cancer has a more aggressive clinical course than Type I tumours.

Prognostic Factors

Endometrial cancer is currently surgically staged as anatomical extent of disease is felt to be an important indicator of prognosis and in directing adjuvant therapy. **Table 4** shows the current FIGO staging system. Additional

Table 4 FIGO* surgical staging for endometrial carcinoma

I		Tumour limited to the uterine fundus
	IA	No myometrial invasion
	IB	Myometrial invasion $\leq 50\%$
	IC	Myometrial invasion $\geq 50\%$
II		Tumour extends to the cervix
	IIA	Superficial glandular spread
	IIB	Stromal invasion
III		Regional tumour spread to the pelvis
	IIIA	Involvement of the uterine serosa, adnexa or positive peritoneal cytology
	IIIB	Vaginal metastases

Simplified Staging for Ovarian Carcinoma

Stage I Growth limited to the ovary/ovaries

Stage II Growth involving one or both ovaries, but limited to the true pelvis

Stage III Extension of tumour beyond the true pelvis, and/or positive pelvic lymph nodes or implants outside the pelvis

Stage IV Distant metastases

* Federation of International Gynecologic Oncologists.

features in determining prognosis and risk of recurrence include histological grade, histological subtype and age.

Studies of survival based on surgical stage have shown that depth of myometrial invasion is important with a 97% survival in patients with tumours confined to the inner third of the myometrium, decreasing to 70% with outer third involvement. The incidence of lymph node metastasis increases with depth of muscle invasion. Involvement of the cervix (stage II), particularly the cervical stroma, decreases the 5-year survival to 50%, probably related to the increased incidence of para-aortic lymph node metastases in these patients. Grade III histology is a sensitive indicator for poorer prognosis with a higher percentage of pelvic lymph nodes involved in these patients (Ludwig, 1995).

When one considers Type I versus Type II tumours, stage becomes a less important indicator of prognosis. Tumours with Type II histology do worse stage for stage than those with Type I histology. This is reflected in overall survival rates of 33% and 92%, respectively. These patients tend to present with more advanced disease than is initially suspected. Even when disease is at a very early stage, these tumours have a higher recurrence risk (Ludwig, 1995).

Clinically, age is considered to be an indicator of prognosis. Younger women tend to fare better than older women. Younger women who develop endometrial carcinoma usually do so in the context of complex atypical hyperplasia secondary to oestrogen excess. These malignancies tend to be better differentiated and less aggressive. Older women more often have tumours with poor differentiation and are at increased risk for developing tumours with Type II histology (Ludwig, 1995).

Clinical Management

The current first-line therapy for endometrial carcinoma is an exploratory laparotomy with a simple total abdominal hysterectomy/bilateral salpingo-oophorectomy. A lymph node dissection may also be performed as part of the staging procedure based on whether lymph nodes appear enlarged or firm, if high-grade or high-risk histology is known to be present, and when there is $>50\%$ myometrial invasion (Chen *et al*., 1999). The need for adjuvant therapy is based upon final surgical stage. Current modalities include radiation and chemotherapy.

UTERINE STROMAL TUMOURS

The most common benign stromal tumour in the uterus is a leiomyoma. In fact, it is one of the most common tumours encountered in gynaecological pathology with an estimated incidence of 20–40%. These tumours are benign proliferations of smooth muscle. Clinically, they tend to be less important as at least 50% are asymptomatic. Abnormal uterine bleeding (from submucosal leiomyomas) is the most common symptom and can be seen in one third of patients, and one third of patients may have pelvic pain. Local symptoms such as genitourinary and gastrointestinal compression and pelvic pressure may be seen when tumours reach a large size. These tumours may be detected incidentally in hysterectomies for other reasons, or may be discovered on routine bimanual examination. Treatment is often not indicated, but abnormal uterine bleeding, severe pelvic pressure or pain, infertility/habitual abortion due to uterine cavity distortion and compromise of adjacent organs (i.e. hydronephrosis) are all indications for therapy (Barbieri, 1999). Current therapeutic options include hysterectomy or myomectomy (local resection of the leiomyoma). Because these tumours are sensitive to oestrogen and progesterone, drugs to block the actions of these hormones have also been used with some success in an effort to treat these tumours (Deligdisch, 2000).

Uterine sarcomas are the least common uterine malignancy. In general, these tumours represent less than 5% of corpus cancers and have a poorer prognosis. The gross and histological subtypes have been discussed previously. Most patients present with abnormal uterine bleeding, and the tumour may be seen protruding through the cervix (classical for MMMT). The incidence tends to increase with age, and sarcomas may be more common in African Americans. Few risk factors are clearly defined. The most common histological subtype is the malignant mixed mullerian tumour. It also has the worst prognosis with an overall survival of 20%. Leiomyosarcomas account for 30% of sarcomas. It is important to note that most arise independently, and it is now no longer believed that they arise from pre-existing leiomyomas. Survival for patients with disease confined to

the uterus approaches 30%. Endometrial stromal sarcomas are the least common subtype, but if the tumour is low-grade, the prognosis is good with 90% survival overall. The standard therapy for sarcomas is hysterectomy. The benefit of adjuvant therapy in extending survival has not been clearly proven (Levenback *et al*., 1996).

PATHOLOGY OF THE OVARY

Tumour Pathology

Primary tumours of the ovary can be divided into three major groups that are based upon their presumed cell of origin. Tumours derived from the mesothelial-like lining of the ovary comprise the common group of epithelial ovarian tumours, which show various types of epithelial differentiation. Tumours derived from the germ cell elements are referred to as germ cell tumours, and encompass a wide variety of histological types. Tumours derived from the supporting elements of the ovary, which includes granulosa, theca and fibrous cells of the stroma, make up the group of tumours referred to as sex-cord stromal tumours. These tumours generally have less histological variety than germ cell tumours and are noted for their potential for hormonal production. A fourth group of tumours can be included that make up a wide variety of tumour types, including metastatic tumours, pseudoneoplastic tumours or tumour-like conditions, and a rare group of tumours of uncertain cell origin. For the purpose of simplicity, this last group of rare tumours will not be discussed.

Although ovarian tumours encompass a wide variety of neoplasms with numerous histological types and diverse clinicopathological features, general principles can be established regarding their presentation and outcome. In the premenopausal patient ovarian tumours must be distinguished from benign ovarian processes, including various benign types of cysts, endometriosis and inflammatory lesions whose mode of clinical presentation will often be identical with that for ovarian tumours and include vague abdominal symptoms, pelvic pain, urinary symptoms or abdominal distension. Rarely does either group present as dysfunctional uterine bleeding.

The presentation of epithelial ovarian tumours can best be explained by examining their natural course of spread. Tumour dissemination of ovarian carcinomas generally occurs by penetration of the ovarian capsule and seeding of peritoneal surfaces and by lymphatic invasion and spread to contiguous lymph nodes in the pelvic region. Spread by haematogenous routes to distant sites is a late manifestation and signals an ominous prognosis. Implantation of peritoneal surfaces is generally associated with relatively noninvasive growth, but increasing tumour size causes mechanical interference with other vital structures and organs such as the bowel and ureters. Because of this predictable route of dissemination, staging and prognosis of epithelial ovarian tumours can be divided into tumours confined to the pelvis and tumours that have spread beyond the pelvis. In contrast, germ cell tumours frequently disseminate to distant sites by haematogenous and lymphatic routes. Sex-cord stromal tumours, even when of a malignant nature, tend to be localized to the ovary, with metastases to distant sites noted many years after their removal.

Epidemiology

The incidence of ovarian cancer has remained relatively stable over the last 25 years. Ovarian cancer is the sixth most common form of cancer in women, the fourth leading cause of cancer death among women and the leading cause of death among gynaecological malignancies. The overall lifetime risk of a woman dying from ovarian cancer is approximately 1.5%, with ovarian cancer being the cause of death in one of every 90 women. Most malignant ovarian tumours occur in women over the age of 45 years where an ovarian mass has an approximately 30–40% chance of being malignant. Conversely, in women less than 45 years old, or premenopausal women, an ovarian mass is most likely to be a benign with malignant tumours making up less than 15% of this group. Thus, an ovarian mass in a premenopausal woman is approached much differently from an ovarian mass in a postmenopausal woman.

Certain generalizations can be made regarding the incidence of specific tumour types in pre- and post-menopausal woman. The most common ovarian tumour accounting for approximately one quarter of all ovarian tumours and one third of all benign ovarian tumours is the benign dermoid cyst (mature cystic teratoma), which belongs to the group of germ cell tumours. Germ cell tumours as a group are much more common in the first two decades of life. Conversely, the common group of epithelial ovarian tumours are rare in the first two decades of life and become much more prevalent with the onset of reproductive age and in postmenopausal women. Within these two groups, malignant epithelial ovarian tumours (carcinomas) tend to occur after the age of 60 years, whereas benign ovarian tumours and tumours of low malignant potential (borderline tumours) tend to occur in the 40–60 year age group. Sex-cord stromal tumours are most common in women from 40 to 60 years of age.

Most epidemiolgical factors of ovarian cancer relate to the common group of epithelial tumours. This group of tumours is uncommon before the age of 40 years, and peaks in the seventh to eighth decades of life. Most tumours in this group, 70–80%, are of a benign nature with the incidence of malignant tumours being closely related to age. Factors that have been associated with a higher risk of epithelial ovarian cancer include infertility, nulliparity,

multiple miscarriages and the use of clomiphene (a fertility drug); factors associated with a lower risk include multiple pregnancies, breast feeding, tubal ligation and the use of oral contraceptives. Most of these factors can be explained by the hypothesis that ovulation induces an aberrant repair process, and factors which decrease ovulation decrease the incidence of this group of tumours. The use of oestrogen in postmenopausal women has not been shown to increase the risk of ovarian cancer.

Aetiology

Between 5 and 10% of women with ovarian cancer have a family history of ovarian cancer, and approximately half or less of this group have inherited disease. A woman with one affected first-degree relative has an approximately 5% lifetime risk of developing ovarian cancer, while two or more affected first-degree relatives confers a 30–50% lifetime risk of developing ovarian cancer. The most notable hereditary group is the breast–ovarian cancer group linked to *BRCA-1*, a probable tumour-suppressor gene. BRCA-linked ovarian cancer does not show distinct clinicopathological characteristics that distinguish it from sporadic ovarian tumours, although different studies have shown a significantly more favourable outcome. Owing to the much higher incidence of sporadic ovarian cancer versus inherited ovarian cancer, a routinely obtained family history is an unreliable way to identify patients who might be at high risk of developing ovarian cancer. Unlike *BRCA-1*, which confers a relatively high risk for breast and ovarian cancer, the incidence of ovarian cancer with *BRCA-2* appears to be much lower. Other inherited ovarian tumours are usually components of a multi-system genetic syndrome, and usually are not related to the common group of epithelial tumours. These syndromic complexes include Peutz–Jeghers syndrome (sex-cord stromal tumours), gonadal dysgenesis (gonadoblastomas), basal cell nevus syndrome (ovarian fibromas), ataxia telangiectesia, Muir–Torr syndrome, Li–Fraumeni syndrome and Cowden syndrome.

The effect of environmental factors on the incidence of ovarian cancer have been inconclusive with the exception of drugs used to enhance fertility or treat infertility, which generally increase the rate of epithelial malignancies. Factors which have shown conflicting results but may be possibly related to an increased rate of ovarian cancer include increased dietary fat, increased coffee consumption and the use of talc in the perineal area. The effect of oral contraceptive (OC) use has been extensively studied with results consistently showing a lower incidence of ovarian cancer in women who have used OCs for more than five consecutive years. A hypothetical analysis of the protective effect of OCs in nulliparous women with no family history of ovarian cancer has postulated that the risk of ovarian cancer in this group of women is reduced by more than 50% after 5 years of consecutive OC use and is less than the risk in parous women with a comparable family history and no history of OC use. (See also chapter *Inherited Predispositions to Cancer*.)

Screening and Prevention

Screening for ovarian cancer has primarily been directed towards the postmenopausal patient owing to the much higher incidence of malignant tumours in this group. Stage I ovarian carcinomas have a 90% cure rate, but only 25% of such tumours present as localized malignancies. Stage II ovarian carcinomas have an approximately 70% cure rate, while the prognosis for Stage III and IV carcinomas declines dramatically to 15–20%. Various screening tests that have been employed for the detection of ovarian cancer include tumour markers, ultrasound, rectovaginal or bimanual vaginal examination and Pap smears. All of these tests have various problems which has prompted many investigators to combine the various tests to increase sensitivity and specificity. Rectovaginal and bimanual vaginal examination is sensitive with experienced gynaecologists, but lacks specificity owing to the frequent occurrence of benign ovarian masses. Pap smears, although relatively specific, lack sensitivity as only 10–30% of ovarian cancers can be detected by this method (Nuovo, 1994).

Of the various tumour markers studied, including carcinoembryonic antigen, ovarian cyst adenocarcinoma antigen, lipid-associated sialic acid, NB/70K, TAG 72.3, CA 15-3 and CA 125, only the last has received widespread use. CA 125 is not specific for ovarian cancer and can be elevated in 5–40% of benign gynaecological masses including uterine leiomyomas and endometriosis. Other nongynaecological cancers, including those of the pancreas, stomach, colon, and breast, have been associated with elevated CA 125. Among gynaecological cancers other than ovarian cancers, elevated CA 125 has been reported in cervical and fallopian tube malignancies. Although rare, elevated CA 125 has been reported in up to 1% of healthy women with no evidence of cancer. Problems with sensitivity also exist with CA 125 and must be considered with regards to quantitative parameters. CA 125 is determined by radioimmunoassay methods and is reported in units per millilitre, with levels $>35\,\mathrm{U\,mL^{-1}}$ being considered abnormal and levels $>65\,\mathrm{U\,mL^{-1}}$ being indicative of malignancy. CA 125 will be elevated above normal levels in 80–85% of women with Stage III or IV ovarian cancer, but varies in different studies from 30 to 50% in women with Stage I or II ovarian cancer. Most certainly it can be said that CA 125 is much more sensitive and reliable at detecting higher-stage ovarian cancer, but the lack of elevated CA 125 must be regarded cautiously in early ovarian cancer.

Owing to the lack of sensitivity of pelvic examination and serum CA 125 in detecting ovarian cancer a

combination of these tests with ultrasonography has been advocated. A large study involving more than 22 000 postmenopausal women using a combination of CA 125 measurement and ultrasonography showed a specificity of greater than 99% for the detection of ovarian cancer but a positive predictive value of less than 30%. Others have advocated a more elaborate screening protocol first by transvaginal ultrasonography with abnormal results followed by pelvic examination, serum CA 125 determination, Doppler flow sonography and tumour morphological indexing by ultrasonography results. In these and numerous other studies the routine testing of asymptomatic women has shown limited utility in the prevention of ovarian cancer owing to the number of diagnostic laparotomies performed per cancer detected. It has been calculated that screening 100 000 asymptomatic women over the age of 45 years for ovarian cancer would detect 40 cases of ovarian cancer with 5398 false-positive results and 160 complications from diagnostic laparotomy. The NIH Consensus Development Panel in their most recent statement in 1995 concluded that there was no evidence to support routine screening in the general population and no convincing data even in high-risk patients.

Gross/Histopathology/Preinvasive Lesions/Ultrastructure/Immunohistochemistry

Among the three common groups of ovarian tumours, i.e. sex-cord stromal tumours, germ cell tumours and epithelial tumours, the last group accounts for over 90% of all malignant tumours of the ovary and approximately 60% of all ovarian tumours. The epithelial ovarian tumours are divided according to cell type and architectural pattern, with special reference given to any significant contribution by the surrounding ovarian stroma. If the tumour contains a prominent stromal component the suffix-fibroma is attached and the tumour is designated an adenofibroma rather than an adenoma, with this particular entity being almost invariably a benign proliferation. Architectural features are used in describing the location of the tumour as either on the surface or within the ovary, the degree of solid or cystic component, and are important in assessing the risk of benign or malignant behaviour. If the most prominent portion of the tumour is on the outer surface of the ovary, which almost invariably has papillary histological features, these are designated as surface papillary tumours. If the tumour contains a prominent cystic component the prefix cyst- is attached, such as cystadenoma, cystadenofibroma (**Figure 11**) and cystadenocarcinoma. A solid tumour with no cystic change would thus be called an adenoma, adenofibroma or adenocarcinoma. Adenofibromas and cystadenofibromas are generally solid tumours with the cystic component of cystadenofibromas only being recognized microscopically. The majority of ovarian tumours of the epithelial type are grossly cystic and their distinction into specific type or prediction of their biological behaviour is limited by the gross examination. Most ovarian tumours in general cannot be distinguished from each other by their gross pathologic examination and require histologic examination for their final classification.

Architectural features are especially important in predicting the biological behaviour of a given tumour which is generally divided into three main categories of benign, borderline or malignant. Benign architectural features generally imply a tumour, whether it be solid or cystic, has a simple morphological pattern and shows no invasion of the surrounding stroma. Malignant architectural features are the opposite extreme with a complex morphological pattern and obvious invasion of the surrounding stroma. In between are borderline tumours. Each specific tumour type has its own specific criteria for what defines a borderline tumour, but the general concept is that borderline tumours exhibit a morphological pattern between simple and complex and can show minimal invasion of the surrounding stroma. Cytological features are also evaluated in placing a given tumour into one of these three specific categories of biological behaviour and include nuclear atypia, number of mitoses and other parameters such as multinucleation and the amount of cytoplasm present.

The most specific categorization of the epithelial ovarian tumours relates to the type of cellular component present, with the five major groups being serous (resembles fallopian tube), mucinous (resembles cervix), endometrioid (resembles endometrium), clear cell (resembles kidney tumour) and transitional cell (Brenner tumours, resemble urinary bladder). Not infrequently epithelial tumours of the ovary will be associated with numerous small foci of tumour spread throughout the peritoneal lining, which can be either totally benign or malignant. These peritoneal implants generally recapitulate the primary ovarian tumour and their origin is generally considered to be an example of independent primary tumours in a low-grade tumour, while in high-grade tumours they may be either true metastases or independent primary tumours. It is important to realize that these implants do not necessarily indicate malignant behaviour of the primary ovarian tumour and are not an unexpected finding.

The most common of all of the epithelial tumours are tumours of the serous type, which account for over half of this group and from 30 to 50% of all ovarian tumours. Over 70% of serous ovarian tumours are benign, with borderline tumours being relatively rare and malignant types accounting for the remaining 20–25% of these tumours. Serous tumours are usually cystic, generally of moderate size (less than 10 cm) and lined by an epithelium that mimics that of the fallopian tube being ciliated cuboidal or columnar. Microscopic features of benign tumours are single or multiple cysts with simple papillae projecting into the lumen (usually comprising less than 10% of the total cyst wall) that are lined by a single layer of serous

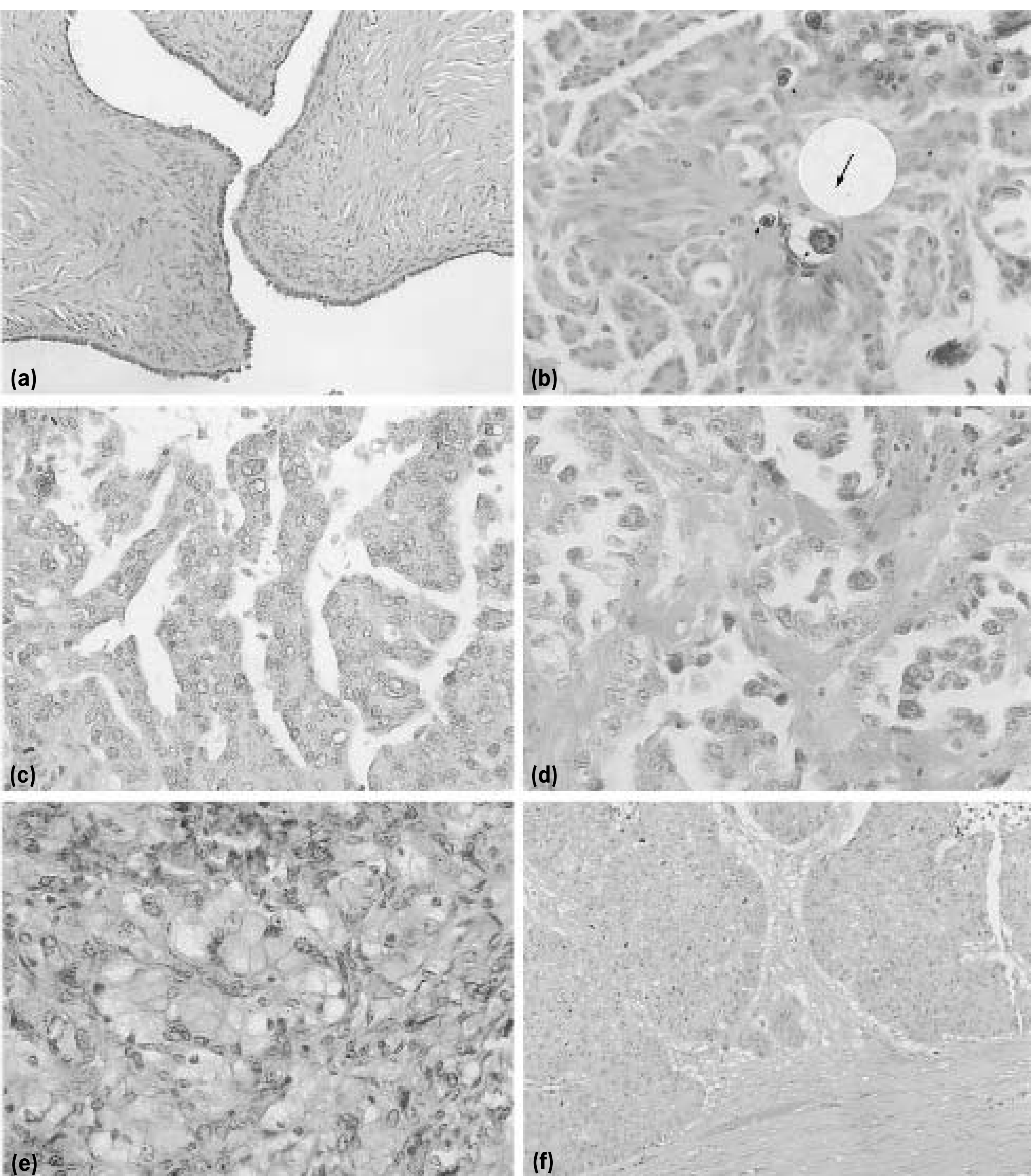


Figure 11 Epithelial tumours of the ovary. Panel (a) shows the large, broad papilla of a serous cystadenofibroma; the simple columnar lining indicates that this is a benign lesion. In panel (b), note that the papilla are more complex and that psammoma bodies are now evident (arrow); this is a serous cystadenoma of borderline malignant potential. In contrast, note the much more complex architecture of a serous cystadenocarcinoma (c). The other panels show other carcinomas of the ovary including clear cell carcinoma (d), note similarity to similar tumour of the endometrium), metastatic carcinoma from the stomach (Krukenberg tumour, where the malignant cells contain ample mucin (e)) and (f) endometrioid carcinoma of the ovary with its characteristic squamous metaplasia.

cells with no atypia. Borderline tumours show more complex papillae with a stratified layer of cells that usually show some atypia. Owing to the complex branching nature of the papillae and multiple layer of cells lining the papillae, histological sections will often show small clusters of cells that appear to be floating freely in the cystic space adjacent to the papillae. As these papillae become more complex they may become condensed into concentric concretions referred to as psammoma bodies, a characteristic feature of serous tumours (**Figure 11**). Some serous borderline tumours may show microinvasion of the surrounding stroma which has not been shown to have any prognostic significance. Serous borderline tumours frequently are associated with peritoneal implants that are generally of low grade and noninvasive. If these implants are benign they are often referred to as endosalpingiosis after their resemblance to fallopian tube epithelium.

Serous carcinomas are the most common malignant ovarian tumour and are most common in women after 65 years of age. They usually are grossly cystic but can appear solid owing to the greater degree of epithelial proliferation and their cystic nature is only recognized microscopically; when cystic most of their inner wall contains papillary structures. Serous carcinomas show more complex branching of the papillae and stratification of the epithelial lining to such a degree that crowding of the papillae results in a more solid appearing tumour with slit-like lumina (**Figure 11**) Areas of the tumour may show solid sheets of cells, making their distinction from endometrioid tumours difficult. Findings that may help in this instance are the presence of multinucleated cells and psammoma bodies and the absence of squamous differentiation and intracellular mucin, the latter two of which are often present in endometrioid carcinomas. Serous carcinomas show obvious stromal invasion and nuclear atypia.

Mucinous tumours represent approximately 25% of epithelial tumours and 15% of all ovarian tumours and are noted for the large size they commonly attain at presentation. Of all ovarian tumours, mucinous tumours are the type that can be most readily recognized grossly owing to their frequent multicystic appearance with prominent mucin production. The majority of mucinous tumours, (>75%), are benign with borderline tumours being extremely rare. As opposed to the fairly constant histological type seen in any given serous tumour, mucinous tumours are noted for histological variability. One histological section of a mucinous tumour may appear of borderline type, while an adjacent histological section shows invasive carcinoma. For this reason it is important to sample mucinous tumours well.

Mucinous tumours are lined by epithelium that resembles that of the endocervical canal and consists of mucin-filled columnar cells. Less often the epithelial lining resembles intestinal-type epithelium with prominent goblet cells. Low-grade mucinous tumours, or mucinous cystadenomas, are generally cystic tumours lined by a single layer of endocervical-type mucinous epithelium. The criteria for endocervical-type borderline mucinous tumours closely resemble those of serous borderline tumours. Mucinous carcinomas vary greatly in their microscopic appearance in any given tumour and are not divided into endocervical or intestinal type, both of which are frequently seen in a single example. Mucinous carcinomas vary from solid nodules to glands, often with scant stroma.

Endometrioid tumours are usually carcinomas, with benign and borderline examples being very uncommon. Although endometrioid tumours comprise less than 5% of all ovarian tumours, they account for 15–20% of malignant ovarian tumours. Endometrioid carcinomas may be cystic or solid and lack gross features that distinguish them from other ovarian tumours. They are lined by epithelium that mimics the endometrium and are associated with endometriosis in 5–10% of all cases and a synchronous carcinoma of the endometrium in approximately 25% of cases. Endometrioid carcinomas classically are characterized microscopically by tubular glands lined by mucin-free psuedostratified epithelium as well as squamous metaplasia. Another feature commonly seen in endometrioid carcinomas is focal squamous differentiation (**Figure 11f**). Owing to their glandular appearance their distinction from metastatic carcinoma, particularly colonic adenocarcinomas, can be difficult. In the case of adenocarcinomas of the colon immunohistochemical stains can help in this distinction. Epithelial tumours of the ovary as a group are positive for cytokeratin 7 and negative for cytokeratin 20 whereas intestinal adenocarcinomas generally show the opposite pattern of immunoreactivity. Another common distinction is from metastatic breast cancers and again immunohistochemical stains can be helpful, with ovarian carcinomas staining positively for CA 125 and metastatic breast cancer staining for gross cystic disease fluid protein 15.

Clear-cell carcinomas represent approximately, 5% of all ovarian carcinomas. Like endometrioid carcinomas they are frequently associated with endometriosis, and in some reports up to 50% of clear-cell carcinomas show such an association. Clear-cell carcinomas, unlike other ovarian carcinomas, more frequently present as Stage 1 tumours prompting earlier investigators to suggest a better prognosis for this class of ovarian carcinomas, but if adjusted for stage at presentation clear-cell carcinomas have a worse prognosis than other epithelial ovarian carcinomas. Grossly, clear-cell carcinomas may be solid or cystic and typically show numerous nodules of solid tumour protuding into cystic cavities. Microscopically clear-cell carcinomas usually show a solid nest of clear cells mixed with small cysts lined by cells with a hobnail appearance (**Figure 11**).

Transitional cell tumours of the ovary are relatively rare, making up less than 2–3% of all ovarian tumours, and of this group practically all are benign. Often referred to as

Brenner tumours, these tumours show urothelial differentiation similar to the lining of the urinary tract. Grossly, Brenner tumours are usually solid tumours but may show cystic change. Microscopically, Brenner tumours are composed of sharply demarcated nests of transitional cells, often with small cysts in their central portion, in an abundant fibromatous stroma. Up to half of Brenner tumours are incidental findings in ovaries removed for other reasons, and are frequently associated with mucinous cystadenomas.

The other common group of epithelial tumours seen in the ovary are metastatic tumours. Owing to the rich vascularity of the ovarian stroma, 10–20% of all tumours of the ovary are metastatic. The clinical history is very valuable in such instances, but other findings that may help in such instances are the presence of bilateral tumours, numerous separate nodules within a single ovary and the presence of tumour on the ovarian surface without peritoneal involvement. The presence of bilateral involvement should be approached cautiously as only 75% or less of metastatic ovarian tumours are bilateral and 10–20% of primary ovarian carcinomas will present as bilateral masses with primary serous tumours being bilateral 50–75% of the time. Up to one third of breast cancer patients and one quarter of colon cancer patients will have ovarian metastases during the course of their treatment. Among metastatic ovarian tumours, those derived from signet ring cell carcinomas of the stomach, i.e. Krukenburg tumours, have received the most attention. Although not the most frequent metastatic tumour to the ovary, Krukenburg tumours have received such attention because they often occur in middle-aged to younger patients without a known primary. Krukenburg tumours can closely mimic clear-cell carcinomas and germ cell tumours owing to their histological appearance (**Figure 11**).

Sex-cord stromal tumours are relatively uncommon tumours that make up approximately 5% of all ovarian tumours. As a general rule this group of tumours occur in a slightly younger population than epithelial tumours and examples in paediatric populations are not rare. Derived from the stroma of the ovary this group of tumours is noted for their response to and production of hormones as they recapitulate their normal role in reproduction. Most ovarian sex-cord stromal tumours are oestrogenic but up to one quarter may be androgenic. The most general histological characteristic of this group of tumours is the presence or absence of the accumulation of lipid in the tumour cells which imparts a particular histological appearance. The most common tumours in this group include fibromas, thecomas and granulosa cell tumours. Grossly they are generally solid tumours but may be cystic, especially in the case of granulosa cell tumours. These tumours can occur in a mixed pattern and are usually solid nests of cells that vary from spindle-shaped cells to round or oval cells. Fibromas and thecomas are generally solid tumours and a sharp distinction between the two is often difficult to distinguish as fibromas may undergo leutinization focally and are referred to as fibrothecomas. Whereas fibromas are generally spindle-shaped cells, thecomas undergo leutinization and classically have pale abundant cytoplasm (**Figure 12**). Granulosa cell tumours are divided into an adult and juvenile types, with the latter having a better prognosis, and both are characterized by round to oval cells with nuclear grooves surrounding small cystic spaces. In the case of adult granulosa cell tumours, these cystic spaces take on a microfollicular appearance classically referred to as Call–Exner bodies (**Figure 12**). Granulosa cell tumours generally present as Stage I tumours and have a 90% survival at 5 years, but are notorious for recurrence at distant metastatic sites 15–20 years later with a 50% mortality rate at this time. Other tumours in this group often referred to as steroid cell tumours, including Sertoli–Leydig cell tumours, are extremely rare.

Germ cell tumours account for approximately 30% of all ovarian tumours, and the dermoid cyst accounts for approximately 95% of this group of tumours. The remaining 5% of germ cell tumours are generally of a malignant tumour type and include dysgerminoma, yolk sac tumour, embryonal carcinoma, choriocarcinoma and immature teratoma. This latter group of tumours occurs almost exclusively in patients less than 20 years of age and as a group account for less than 1% of all ovarian tumours. Malignant germ cell tumours in general are most common in the paediatric population and young women. There are major differences between germ cell tumours occurring in very young female patients and those occurring in adolescents and adults. Germ cell tumours occurring in early childhood tend to be pure yolk sac tumours or teratomas, have no malignant non-neural epithelial components, are euploid or tetraploid and prior to puberty occur almost exclusively in extragonadal sites with primary germ cell tumours of the ovary being very rare. Germ cell tumours occurring in adolescents and adults are generally of mixed histological type or dysgerminomas, may have malignant epithelial components, commonly show isochromosome 12p and aneuploidy by cytogenetic studies and generally involve the ovary with extragonadal tumours being rare.

Germ cell tumours may be composed of a number of different tissue types or show a combination of different germ cell tumour types and are frequently complex histologically. It is the most malignant component present that determines the behaviour of these tumours. At one extreme of this complexity is the mature teratoma that contains tissue representing all three embryonic layers in a uniformly mature fashion. At the other end of the spectrum are immature teratomas and mixed malignant germ cell tumours which show a wide spectrum of histological types and degree of malignancy. In between these extremes of histological complexity are pure malignant germ cell tumours, with dysgerminoma being the most common example (**Figure 12**).

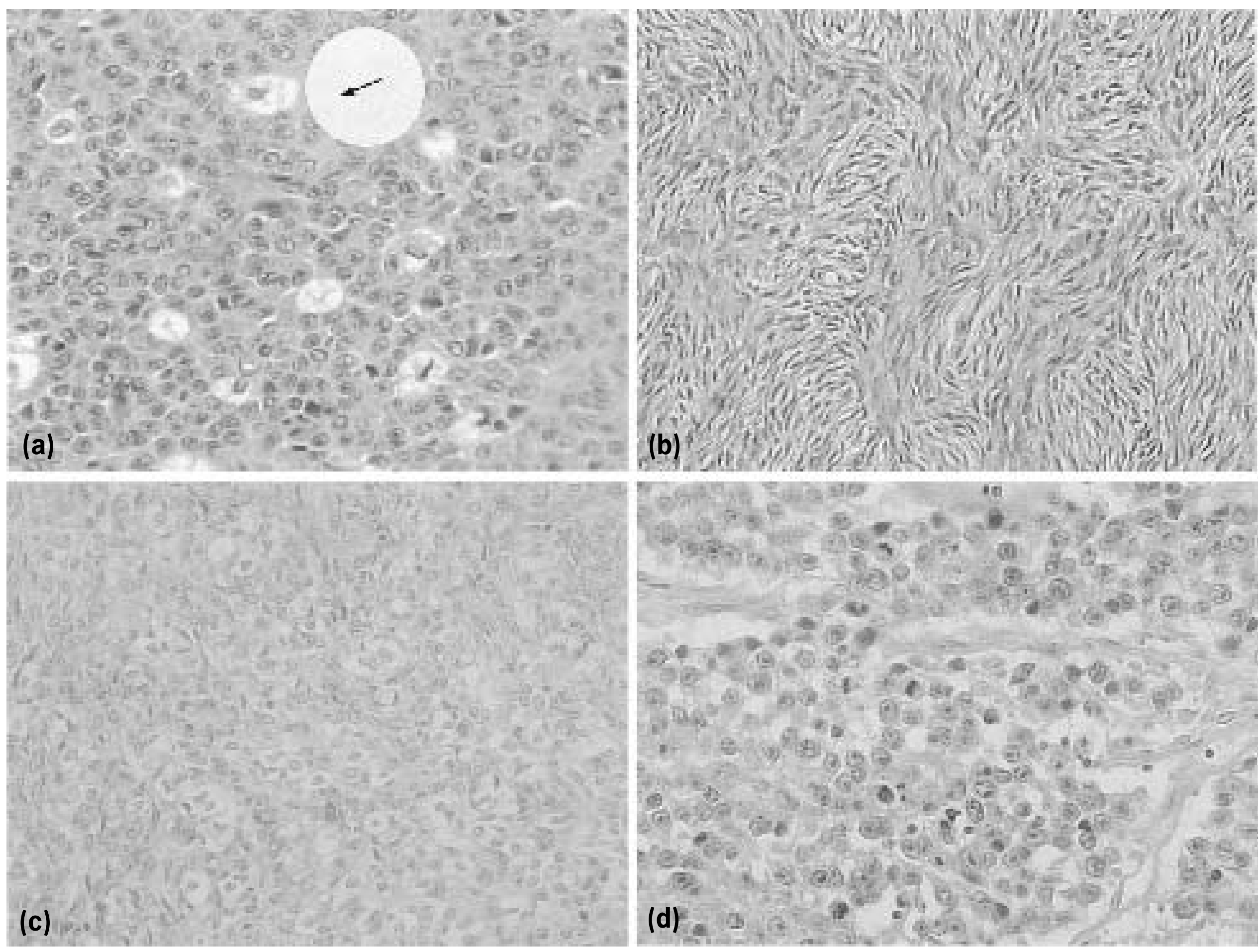


Figure 12 Stromal and germ cell tumours of the ovary. Panel (a) shows a granulosa cell tumour; note the nuclear grooves and the circular areas of degenerated material, called Call–Exner bodies (arrow). Panel (b) shows the interlacing fascicles of a fibroma whereas panel (c) demonstrates the prominent cytoplasm of a luteinized thecoma; this change often indicates hormonal production by the tumour. Panel (d) shows the large, polygonal cells of a dysgerminoma of the ovary; scattered lymphocytes are also characteristic of this tumour.

Molecular Genetic Findings

The most commonly identified gene conferring a high risk of epithelial ovarian cancer is *BRCA-1*, as discussed above. Among the sex-cord stromal tumours, molecular studies have generally been directed at adult-type granulosa cell tumours, where mutations have been identified in the G_I subunit of regulatory G proteins. Similar findings have been identified in individuals with the McCune–Albright syndrome, who show dysregulated ovarian function and isosexual precocity but no associated risk of ovarian cancer. Further studies are needed to identify the role of molecular alterations and the associated molecular genetic findings associated with regulatory G proteins.

The majority of molecular genetic findings in germ cell tumours have centred around cytogenetic findings with single gene mutations not being readily identified. The majority of mature teratomas have been shown to be karyotypically normal whereas immature teratomas show a high frequency of random nonrecurrent chromosomal abnormalities. Other adult germ cell tumours have frequently shown two copies of the short arm of chromosome 12, with this finding being rare in prepubertal patients with similiar tumours.

Prognostic Factors

The most reliable indicator of prognosis is stage at presentation. The overall 5-year survival rate for ovarian cancer confined to the ovary (Stage I) is approximately 75%. The rate declines to approximately 60% for pelvic extension of tumour beyond the ovary (Stage II) and 20% for metastatic disease (Stage III or IV). The size of the tumour in postmenopausal women has important prognostic implications, with tumours less than 5 cm usually being benign (95%), while tumours greater than 10 cm in size are often malignant (60%). Another prognostic factor at presentation is the presence or absence of ascites and the

volume of fluid present, with high-volume ascites having a worse prognosis. Levels of CA 125 both preoperatively and postoperatively have prognostic significance. Older age at presentation is associated with a worse prognosis, although other medical factors associated with the elderly patient make this parameter difficult to judge. A large volume of residual tumour after initial surgery and adjuvant chemotherapy portend a poor prognosis. Histological type among epithelial tumours is most important for clear-cell carcinomas, which usually have a higher recurrence rate and lower survival rate for Stage I tumours than other ovarian epithelial carcinomas. The grade (or cytological details) of an ovarian epithelial carcinoma is generally only important in Stage I tumours, as poorly differentiated tumours as opposed to well or moderately differentiated have a worse prognosis (and thus higher stage) and often prompt more aggressive treatment.

Overview of Present Clinical Management

The initial management of a patient with an adnexal mass is directly related to the age of the patient and the size of the mass. An initial mass greater than 10 cm in any patient regardless of age will usually result in laparoscopy or an exploratory laparotomy. At the other extreme an adnexal mass slightly increased over normal ovarian size, or less than 5 cm, is approached differently in pre- and post-menopausal patients. In women of reproductive age the high incidence of cystic enlargement of one or both ovaries due to the process of ovulation and formation of corpora lutea allows a much more cautious approach. These patients will usually be re-evaluated after two menstrual cycles and if the mass persists will then be evaluated by transvaginal ultrasonography and CA 125 measurement. In the postmenopausal patient any ovarian enlargement will prompt transvaginal ultrasonography and measurement of CA 125. In these patients any abnormal findings by ultrasonography other than a simple cyst will prompt an exploratory laparotomy.

In the event that a primary ovarian cancer is found, the most important parameter that will influence future treatment is accurate surgical staging. A simplified view of ovarian cancer staging determines three basic parameters that include whether or not the tumour is confined to the ovary or ovaries (Stage I) and, if not, then whether the cancer is limited to the true pelvis (Stage II) or extends beyond the true pelvis (Stage III or IV). Generally, ovarian epithelial cancers limited to the ovary do not require adjuvant chemotherapy or radiation and reproductive sparing surgery may be performed with the exception of tumours that are poorly differentiated and clear-cell carcinomas. All other stages of ovarian epithelial cancers require at least a total abdominal hysterectomy and some form of adjuvant chemotherapy and radiation. Of special note is that some borderline tumours, even if bilateral, can be treated with a reproductive sparing procedure.

Proper and accurate staging are particularly important in malignant germ cell tumours where reproductive sparing surgery is sufficent for Stage I tumours, which is the common presentation in this group. With the advent of multiagent chemotherapy, higher-stage germ cell tumours can be cured with a high rate of success (90%), and accurate staging can prevent significant morbidity and mortality in patients requiring such therapy, and just as importantly avoid undesirable side effects in those patients not requiring chemotherapy.

The management of sex-cord stromal tumours is not as well defined as for epithelial and germ cell tumours. Surgical treatment alone is often considered adequate treatment with the choice of adjuvant chemotherapy and/or radiation being of some debate. Historically, granulosa cell tumours have been treated with postsurgical radiation, but the recent recognition of their frequent recurrence 15–20 years later has prompted reconsideration of their treatment.

REFERENCES

Barbieri, R. (1999). Ambulatory management of uterine leiomyomata. *Clinical Obstetrics and Gynecology*, **42**, 197–205.

Burke, T., *et al*. (1996). Endometrial hyperplasia and endometrial cancer. *Obstetric and Gynecological Clinics of North America*, **23**, 411–455.

Chen, L., *et al*. (1999). Endometrial cancer: recent developments in evaluation and treatment. *Oncology (Huntington)*, **13**, 1665–1670.

Cirisano, F., *et al*. (1999). Epidemiologic and surgicopathologic findings of papillary serous and clear cell endometrial cancers when compared to endometrioid carcinoma. *Gynecology and Oncology*, **74**, 385–394.

Crum, C. P. and Nuovo, G. J. (1991). *Human Papillomavirus and Their Relationship to Genital Tract Neoplasms.* (Raven Press, New York).

Deligdisch, L. (2000). Hormonal pathology of the endometrium. *Modern Pathology*, **13**, 285–294.

Eliezri, Y., *et al*. (1990). The occurrence of human papillomavirus DNA in cutaneous squamous and basal cell neoplasms. *Archives of Dermatology*, **23**, 836–842.

Kurman, R., *et al*. (1995). Endometrial carcinoma. In: Kurman, R. (ed.), *Blaustein's Pathology of the Female Genital Tract.* 439–486 (Berlin, Springer).

Levenback, C., *et al*. (1996). Uterine sarcoma. *Obstetric and Gynecological Clinics of North America*, **23**, 457–473.

Ludwig, H. (1995). Prognostic factors in endometrial cancer. *International Journal of Gynaecology and Obstetrics*, 49, Suppl., S1–S7.

Nuovo, G. J. (1994). *Cytopathology of the Female Genital Tract: An Integrated Approach.* (Williams and Wilkins, Baltimore).

Nuovo, G. J. (1997). *PCR In Situ Hybridization: Protocols and Applications*, 3rd edn. (Raven Press, New York).

Nuovo, G. J., *et al*. (1991). Correlation of histology and detection of human papillomavirus DNA in vulvar cancers. *Gynecology and Oncology*, **43**, 275–280.

Nuovo, G. J., *et al*. (1999). *In situ* detection of the hypermethylation-induced inactivation of the *p16* gene as an early event in oncogenesis. *Proceedings of the National Academy of Sciences of the USA*, **96**, 12754–9.

FURTHER READING

Brown, F. M., *et al*. (1999). LSIL biopsies after HSIL smears. Correlation with high-risk HPV and greater risk of HSIL on follow-up. *American Journal of Clinical Pathology*, **112**, 765–768.

Lin, M. C., *et al*. (1998) Patterns of allelic loss (LOH) in vulvar squamous carcinomas and adjacent noninvasive epithelia. *American Journal of Pathology*, **152**, 1313.

Navarro, M., *et al*. (1997). Cytologic correlates of benign versus dysplastic abnormal keratinization. *Diagnostic Cytopathology*, **17**, 447–451.

Nuovo, G. J. (1997). *In situ* detection of PCR-amplified metalloprotease cDNAs, their inhibitors, and human papillomavirus transcripts in cervical carcinoma cell lines. *International Journal of Cancer*, **71**, 1056–1060.

Nuovo, G. J. (1999). Detection of human papillomavirus DNA in Papanicolaou smears: Correlation with pathologic and clinical findings. *Diagnostic Molecular Pathology*, **7**, 158–163.

Quade, B. J., *et al*. (1999). Frequent loss of heterozygosity for chromosome 10 in uterine leiomyosarcoma in contrast to leiomyoma. *American Journal of Pathology*, **154**, 945–950.

Chapter 44

Urinary Tract

Pheroze Tamboli
The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

Rafael E. Jimenez and Mahul B. Amin
Emory University Hospital, Atlanta, GA, USA

CONTENTS

- Introduction
- Normal Development and Structure of Urinary Tract
- Kidney
- Renal Pelvis and Ureter
- Urinary Bladder
- Urethra

INTRODUCTION

A myriad of tumours affect the urinary tract, including numerous benign and malignant types. These tumours arise from the different tissues that comprise the structures of the urinary tract. The malignant tumours originating from the epithelium are referred to as carcinomas and are the most common malignant tumours of the urinary tract (**Figure 1**). Only the most common malignant tumours of the urinary tract are discussed in this chapter. Sarcomas, lymphomas and other assorted tumours affecting the urinary tract are rare and are beyond the scope of this text.

NORMAL DEVELOPMENT AND STRUCTURE OF URINARY TRACT

Normal Development

The urinary tract is almost entirely derived from the mesoderm, except for the most distal part of the urethra, which develops from the ectoderm. The kidney and ureter develop from the intermediate mesoderm, while the urinary bladder and urethra are derived from the urogenital sinus.

Structure

The kidneys are paired organs located in the retroperitoneum, extending from the level of the twelfth thoracic vertebra to the third lumbar vertebra. Each kidney is surrounded by abundant adipose tissue that is covered by the membranous perirenal fascia of Gerota. The renal parenchyma is enveloped by a fibroelastic capsule, which acts as a barrier to the spread of cancer. The kidney is divided into 8–18 (average 14) lobes that are fused together. Each lobe is constituted by the outer cortex and inner medullary pyramid. The nephron is the functional unit, which is composed of the glomerulus and tubule. The tubules empty into collecting ducts that coalesce to form the terminal ducts of Bellini; the latter number 10–25 in each lobe. The ducts of Bellini open at the tip of the papilla, which is the apex of the medullary pyramid. The renal sinus is a concave space at the medial aspect of the kidney where the renal pelvis and calyces, blood vessels and nerves enter the kidney. It is filled with fibroadipose tissue that is continuous with the perirenal adipose tissue and has numerous lymphatic channels and blood vessels.

The renal pelvis is located at the renal hilum and consists of two or three major calyces and 8–18 minor calyces. The minor calyces surround the renal papillae. The ureters arise from the renal pelves, and both are hollow structures with similar features. Their walls are composed of four layers, which from inside out are the urothelium (transitional epithelium) lined mucosa, lamina propria, muscularis propria and adventitia.

The urinary bladder is also a hollow organ that is located deep in the pelvis, behind the pubic bone. The ureters traverse the wall of the urinary bladder to open in the inferior part referred to as the trigone. The superior portion is the dome. The wall of the urinary bladder is composed of four layers, starting from the inside with the urothelium (transitional cell) lined mucosa, followed by the lamina propria, muscularis propria, adventitia and

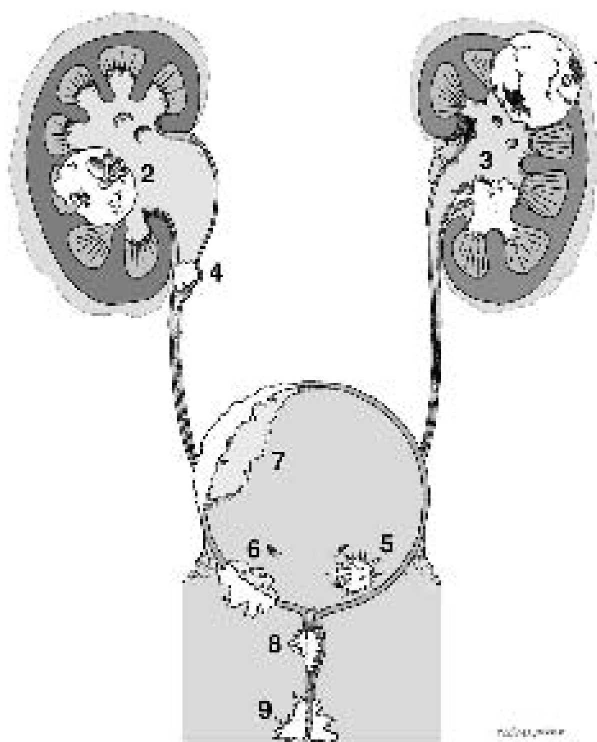


Figure 1 Most common cancers of the urinary tract and their locations. 1, Nephroblastoma (Wilms' tumour) and renal cell carcinoma; 2, collecting duct carcinoma; 3, urothelial carcinoma of renal pelvis; 4, urothelial carcinoma of the ureter; 5, papillary non-invasive urothelial carcinoma of the urinary bladder; 6, invasive urothelial carcinoma of the urinary bladder; 7, squamous cell carcinoma and adenocarcinoma of the urinary bladder; 8, papillary invasive urothelial carcinoma of the urethra; 9, squamous cell carcinoma and adenocarcinoma of the urethra. (Figure courtesy of Dr J. A. Gomez.)

perivesical adipose tissue. The muscularis mucosae is an interrupted layer of smooth muscle in the lamina propria, which needs to be distinguished from the muscularis propria as the staging of urinary bladder tumours is dependent on the depth of invasion.

The urethra is continuous with the neck of the urinary bladder. The male urethra is divided into the prostatic, membranous, penile and bulbar segments. Except for the distal end, which is lined by squamous epithelium, the urethra is also lined by urothelium. The female urethra is shorter and is roughly divided into the proximal one-third, which is lined by urothelium, and distal two-thirds lined by squamous epithelium.

KIDNEY

A variety of different benign and malignant tumours involve the kidney in children and adults, which are listed in **Table 1**. These tumours arise from the different components that constitute the kidney. Malignant tumours arising from the epithelium of the kidney include nephroblastoma (Wilms' tumour), which is the most common renal tumour in children, and renal cell carcinoma, which is the most common renal cancer in adults. Urothelial carcinoma also affects the kidney, arising in the structures lined by urothelium, i.e. the renal pelvis. The discussion in this chapter will concentrate on the most common malignant tumours, i.e. nephroblastoma, renal cell carcinoma and urothelial carcinoma. The other less common tumours will not be described as they are beyond the scope of this text.

Table 1 Tumours of the kidney

Benign tumours	Malignant tumours
Epithelial tumours	*Epithelial tumours*
Renal adenoma	Nephroblastoma (Wilms tumour)[a]
Renal oncocytoma	Renal cell carcinoma
Metanephric (embryonal) adenoma	Urothelial carcinoma of renal pelvis
Nephrogenic adenofibroma	
Mesenchymal tumours	*Mesenchymal tumours*
Angiomyolipoma	Clear cell sarcoma[a]
Mesoblastic nephroma[a]	Rhabdoid tumour[a]
Leiomyoma	Leiomyosarcoma
Juxtaglomerular cell tumour	
Renomedullary interstitial cell tumour	
Haemangioma	
Lymphangioma	
Lipoma	
Myxoma	
Neurogenic tumours	
Benign fibrous histiocytoma	
Solitary fibrous tumour	
Miscellaneous neoplasms	*Miscellaneous neoplasms*
Cystic nephroma	Lymphoma/leukaemia
Solid and cystic biphasic tumour	Plasmacytoma
	Secondary involvement
	Metastasis

[a] Tumours predominantly or exclusively affecting children.

Nephroblastoma (Wilms' Tumour)

Tumour Pathology

Nephroblastoma (Wilms' tumour) is the most common primary renal tumour that affects children, accounting for more than of 80% of renal tumours in children. The eponym Wilms' tumour is used commonly in honour of Max Wilms, who reported seven tumours and reviewed the literature, although he was not the first to describe this tumour. The improvements in treatment and survival of nephroblastoma (Wilms' tumour) patients have in large part been due to the success of multidisciplinary cooperative

groups such as the National Wilms' Tumor Study (NWTS) and the Société Internationale d'Oncologie Pédiatrique (SIOP). A large part of our knowledge of this disease is based on the findings of the NWTS and SIOP studies (Boccon-Gibod, 1998; Neville and Ritchey, 2000).

More than 90% of these tumours develop in children with no other known problems, while the remainder affect children with specific malformations. The presence of an abdominal mass is the most common presenting feature. Other signs and symptoms include pain, haematuria, hypertension, intestinal obstruction and sometimes problems related to distant metastasis.

There are two staging systems for nephroblastoma (Wilms' tumour) based on the extent of spread of tumour; these are listed in **Table 2**. The NWTS and SIOP staging systems are very similar, but the fundamental difference is that in the former the tumour is staged prior to chemotherapy.

Epidemiology

Nephroblastoma (Wilms' tumour) affects one in 8000–10 000 children (0–15 years of age), with over 90% tumours occurring in children under 6 years of age, most commonly between the ages of 3 and 4 years. Patients with unilateral tumours present at a mean age of 41.5 months for males and 46.9 months for females. In contrast, patients with bilateral tumours present at a younger age, with a mean of 29.5 months for males and 32.6 months for females. This tumour is very rare in neonates (0.1% of all nephroblastomas) and rarely affects adults. Girls are affected more commonly than boys, with a male to female ratio of 0.92:1 for unilateral tumours and 0.60:1 for bilateral tumours. In the United States this tumour is most common in African-Americans and least common in Chinese-Americans.

Aetiology

Nephroblastoma (Wilms' tumour) is associated with at least three well-defined congenital syndromes. The WAGR syndrome (Wilms' tumour, Aniridia, Genital abnormalities and mental Retardation) is associated with a deletion in the *WT-1* gene located on chromosome 11p13. The Denys–Drash syndrome (glomerulonephritis, pseudohermaphroditism and nephroblastoma) is also associated with mutations of the *WT-1* gene. Patients with Beckwith–Weidemann syndrome (hemihypertrophy, renal medullary cysts, adrenal cytomegaly and nephroblastoma) have abnormalities of the *WT-2* gene located distal to the *WT-1* gene on chromosome 11p15.5. In up to 1% of patients the tumours are familial. Other possible associations include renal malformations, cutaneous naevi, trisomy 18, genital malformations and neurofibromatosis.

Screening and Prevention

The use of abdominal ultrasound is recommended for routine screening of patients at high risk of nephroblastoma (Wilms' tumour), such as children with the three above-mentioned syndromes. Recommendations for screening include ultrasound examination every 3 months until the age of 7 years, followed by physical examination every 6 months. However, the efficacy of screening programme remains to be determined.

Gross Features

Nephroblastoma usually presents as a solitary tumour that is well encapsulated from the adjacent normal renal parenchyma. The tumour size varies, but the majority are larger than 5 cm and some may exceed 10 cm in greatest dimension. According to data from the NWTS the median weight of these tumours is 550 g. The cut surface shows a mainly

Table 2 Staging systems for nephroblastoma (Wilms' tumour)

Stage	NWTS[a]	SIOP[b]
I	Tumour confined to the kidney and completely resected without rupture	Tumour limited to the kidney and completely excised
II	Tumour extends beyond the kidney (by direct invasion of capsule, extrarenal vascular invasion or tumour spillage without peritoneal contamination), but is completely resected (includes biopsied tumour)	Tumour extending outside kidney, but completely excised, with or without regional lymph node invasion
III	Gross residual tumour, and/or involved surgical margin, and/or tumour in regional lymph nodes	Incomplete surgical excision, without haematogenous metastases, but including biopsy before or at surgery, preoperative rupture, peritoneal implants or nodal metastasis beyond regional lymph nodes
IV	Any local stage with haematogenous metastasis	Distant metastases
V	Bilateral renal tumours at time of diagnosis (tumour in each kidney is separately substaged)	Bilateral renal tumours at time of diagnosis

[a]National Wilms' Tumour Study: tumour staged following surgery and before chemotherapy.
[b]Société Internationale d'Oncologie Pédiatrique: tumour staged following surgery preceded by chemotherapy (unless metastases are present at time of presentation).

solid, uniform, soft, tan or grey tumour. Haemorrhagic and necrotic areas may be present. Cyst formation may be present, sometimes extensively involving the entire tumour.

Microscopic Features

Nephroblastoma (Wilms' tumour) has three components that are present in varying quantities, blastema, epithelium and stroma (**Figure 2, see colour plate section**). Blastema is composed of closely packed small cells without visible cell borders that are arranged as nodules, interconnecting cords or diffuse sheets (Schmidt and Beckwith, 1995). The epithelial component generally forms tubules, glands or glomeruloid structures. The stromal component is made of spindle cells that surround the blastema and epithelial components. The spindle cells may be set within a myxoid background resembling primitive mesenchyme or may have a fibroblastic appearance. The stromal cells may be differentiated into rhabdomyoblasts, bone, cartilage, adipose tissue, mature ganglion cells or neuroglia.

Anaplasia is one of the most important histological features that needs to be determined in every tumour. Anaplasia, which may be either focal or diffuse, is defined as the presence of multipolar mitotic figures and nuclei with increased chromatin that are at least three-fold larger than adjacent cells (Faria *et al*., 1996). Diffuse anaplasia is the presence of anaplasia in more than one portion of the tumour or if found in any extrarenal or metastatic site.

Immunohistochemistry

There are no specific immunohistochemical stains for making the diagnosis of nephroblastoma (Wilms' tumour). The blastema component does stain with vimentin immunohistochemical stain.

Prognostic Factors

Tumour histology and stage are the most important prognostic factors in nephroblastoma (Wilms' tumour). Tumours are divided into those with favourable histology and those with unfavourable histology, depending on the presence or absence of anaplasia. Tumours with unfavourable histology account for 5–6% of all cases. Patients who have tumours with unfavourable histology have a worse prognosis, e.g. nearly half of the patients in the second NWTS who died of disease had unfavourable histology, although they numbered only 12% of all cases. Other prognostic factors that have limited success in predicting prognosis and that are still under investigation include DNA content measured by flow cytometry, nuclear morphometry, serum renin and erythropoietin.

Overview of present Clinical Management

A multimodal approach is used for treating patients with nephroblastoma (Wilms' tumour), which varies in different parts of the world. In most of the United States and Canada the NWTS protocol is followed, which calls for radical nephrectomy (with rare exceptions) followed by chemotherapy and/or radiation therapy, based on the surgical and pathological findings. In most of Europe the SIOP protocols are followed, in which chemotherapy is given prior to nephrectomy.

Renal Cell Carcinoma

Tumour Pathology

Renal cell carcinoma is a malignant tumour that arises from the renal epithelium. The history of renal cell carcinoma is interesting in that it was not until 1960 that Oberling *et al*. laid to rest the controversy regarding the origin of renal cell carcinoma by providing conclusive electron microscopic evidence of its origin from renal tubular epithelium. Prior to the study by Oberling *et al*. there were certain schools of thought that considered renal cell carcinoma to be of adrenal origin. It was Grawitz in 1883 who put forth the theory of renal cell carcinoma originating from adrenal rests in the kidney, rather than from renal epithelium, because of its resemblance to the adrenal gland cortex. Therefore, in the past this tumour was referred to as Grawitz tumour or 'hypernephroma', the latter term being coined by Birch-Hirschfeld in 1892.

A number of renal cell carcinomas are detected incidentally when the patient is investigated for other medical conditions. The tumour may grow substantially before it induces any signs or symptoms. The so-called 'classical triad' of flank pain, flank mass and haematuria is now only seen with advanced-stage tumours, which occur rarely. In addition to the aforementioned, patients may have nonspecific complaints such as fever, weakness, malaise and weight loss. Renal cell carcinoma may also be associated with paraneoplastic syndromes such as polycythaemia, hypertension, Cushing syndrome and hypercalcaemia. In 25%–30% of patients metastasis may be the first indication of the presence of the cancer.

The classification of renal cell carcinoma has changed in the recent past to better embody our understanding of this disease. Two important workshops on the classification of renal tumours were held in 1996 and 1997. The first, entitled 'Impact of Molecular Genetics on the Classification of Renal Cell Tumours' was held in October 1996 in Heidelberg, Germany. The conclusions of this workshop are referred to as the Heidelberg classification of renal tumours (Kovacs *et al*., 1997). The second, entitled 'Diagnosis and Prognosis of Renal Cell Carcinoma: 1997 Workshop', organized by the American Joint Committee on Cancer (AJCC) and Union Internationale Contre le Cancer (UICC), was held in March 1997 (Störkel *et al*., 1997). These two classification systems are listed in **Table 3**. While these two classification systems at first glance may seem to be similar, and they are in the use of similar terms, the major difference is the Heidelberg

classification's emphasis on genetic correlation. Both of these classification schemes are based on the light microscopic appearance of the tumour but are in keeping with our prevailing knowledge of the genetics of these cancers. The diagnostic terminology used is practical, concise and reflective of the morphologic appearance of the tumour; however, some of these terms are not descriptive and do not encompass all the morphologic patterns of a given tumour. There are numerous older classification schemes for subtyping renal cell carcinoma that use different terminology, including terms such as chromophil carcinoma (for papillary renal cell carcinoma), nonpapillary renal cell carcinoma, etc. It is preferable to avoid use of these terms as they may cause confusion.

All renal cell carcinomas are graded according to Fuhrman's nuclear grading system, which is divided into four grades based on the nuclear size, nuclear irregularity and nucleolar prominence (**Table 4**). The grade is based on the highest grade within the entire tumour, and not the predominant grade. It is evaluated at 100× and 400× magnification using a light microscope.

The tumour, nodes and metastasis (TNM) system is the most widely used system for staging renal cell carcinoma (**Table 5**). The other system that is used sometimes is the Robson staging system (**Table 5**). The staging systems are partly based on the extent of invasion by the tumour, with the lower stage tumours (pT1 and pT2) being confined to the kidney, whereas the higher stage tumours (pT3 and pT4) extend beyond the confines of the organ.

Table 3 Classification of renal cell carcinoma

Heidelberg classification[a]	AJCC/UICC classification[b]
Benign tumours	*Benign tumours*
Papillary renal cell adenoma	Papillary adenoma
Renal oncocytoma	Renal oncocytoma
Metanephric adenoma	Metanephric adenoma
Metanephric adenofibroma	Metanephric adenofibroma
Malignant tumours	*Malignant tumours*
Common or conventional renal cell carcinoma	Conventional (clear) renal cell carcinoma
Papillary renal cell carcinoma	Papillary renal cell carcinoma
Chromophobe renal cell carcinoma	Chromophobe renal cell carcinoma
Collecting duct carcinoma	Collecting duct carcinoma
RCC, unclassified	Unclassified renal cell carcinoma

[a]Kovacs *et al.*, 1997.
[b]Störkel *et al.*, 1997.

Epidemiology

Renal cell carcinoma affects approximately 1–3% of all patients with cancer worldwide (McLaughlin and Lipworth, 2000). The highest incidence of renal cell carcinoma is recorded in Scandinavia followed by other parts of northern Europe, the United States and Canada. The lowest incidence rates have been recorded in Central and South America and in Asia. It is estimated that 30 800 new cases of renal cell carcinoma will be diagnosed in the United States in the year 2001, constituting the ninth most common tumour in that country. An estimated 12 100 persons are expected to die of the disease in the United States in the year 2001. The incidence of renal cell carcinoma has been increasing every year in the United States since the 1970s, with a greater increase being noted in the African-American population as compared with Caucasians. This cancer more commonly affects men than women, with a male to female ratio of 2 : 1. Patients of renal cell carcinoma are more likely to be in the late sixth and early seventh decades of life, although these cancers have been reported to occur in children. In addition to the sporadic forms of the cancer, a small percentage are familial, these include renal cell carcinoma associated with Von Hippel–Lindau (VHL) syndrome, familial conventional (clear cell) carcinoma and familial papillary renal cell carcinoma. These cancers usually occur in younger patients, and show autosomal dominant inheritance.

Aetiology

Numerous agents (**Table 6**) are associated with development of renal cell carcinoma (McLaughlin and Lipworth,

Table 4 Fuhrman's nuclear grading for renal cell carcinoma

Grade	Nucleus	Nucleolus
1	Small (10-μm diameter), round, uniform, resembling nucleus of mature lymphocyte	Inconspicuous or absent nucleoli (viewed at 400× magnification)
2	Larger nuclei (15-μm diameter), with slight nuclear irregularity	Small nucleoli (only visible at 400× magnification)
3	Large nuclei (20-μm diameter), with obvious nuclear irregularity	Large, prominent nucleoli (visible at 100× magnification)
4	Same as grade 3 but more bizarre with multilobation and large clumps of chromatin	Large, prominent nucleoli (visible at 100× magnification)

Table 5 1997 TNM staging systems for renal cell carcinoma

Primary tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1	Tumour 7.0 cm or less in greatest dimension, limited to the kidney
T2	Tumour > 7.0 cm in greatest dimension, limited to the kidney
T3	(a) Tumour invades adrenal gland or perinephric tissue, but not beyond Gerota's fascia
	(b) Tumour extension into renal vein(s) or vena cava below the diaphragm
	(c) Tumour invades vena cava above the diaphragm
T4	Tumour invades beyond Gerota's fascia
Regional lymph node (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single regional lymph node
N2	Metastasis in more than one regional lymph node
Distant metastasis (M)	
MX	Presence of distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Stage groupings			
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1-T3	N1	M0
Stage IV	T4	N0–N1	M0
	Any T	N2	M0
	Any T	Any N	M1

Robson staging system	
I	Confined to kidney
II	Perinephric fat invasion but tumour completely resected
IIIa	Renal vein or vena cava involvement
IIIb	Lymph node involvement
IIIc	Vascular and lymphatic invasion
IV	Metastatic disease or involvement of adjacent organs other than the adrenal gland

2000). Cigarette smoking is the most prominent aetiological factor associated with renal cell carcinoma; 20–30% of all renal cell carcinomas affecting men and 10–20% affecting women are attributed to cigarette smoking. There is a strong dose–response relationship associated with cigarette smoking; the relative risk for smokers ranges from 1.2 to 2.3, with a relative risk from 1.9 to 2.5 for heavy smokers. In the familial forms of renal cell carcinoma, genetic alterations play an important aetiological role. Germ-line mutations of chromosome 3p are associated with the VHL syndrome; mutations of chromosome 3p are also associated with familial conventional (clear cell) carcinoma. Germ-line mutations of the *MET* proto-oncogene, located on chromosome 7, are associated with the familial form of papillary renal cell carcinoma. In women, there is a positive association with obesity, which may be hormonally related. Long-term haemodialysis used in the treatment of renal disease leads to an increased incidence of acquired cystic disease, which in turn leads to increased risk of renal cell carcinoma. Although use of the drug phenacetin has been related to urothelial carcinoma of the renal pelvis, there is a less consistent association with renal cell carcinoma. Association with the drug acetaminophen has not been proved, except in one study. Other possible aetiological factors include hypertension, diet, and exposure to asbestos and petroleum products.

Gross Features

Renal cell carcinoma most often occurs as a single solid tumour often located at the periphery of the renal parenchyma (**Figure 3; see colour plate section**). However, collecting duct carcinoma is located more centrally in the medullary region. A bright yellow or light orange colour is most characteristic of conventional (clear cell) renal cell carcinoma, which is the most common type of renal cell carcinoma. In addition, there may be areas of haemorrhage and necrosis resulting in a mottled

Table 6 Aetiological and putative aetiological factors in urinary tract tumours

Renal cell carcinoma	Urothelial carcinoma of renal pelvis and ureter	Urothelial carcinoma of urinary bladder
Cigarette smoking	Cigarette smoking	Cigarette smoking
Germ-line mutations (3p, MET proto-oncogene)	Phenacetin abuse	Arylamines
Obesity (in women)	Thorium exposure	Chemotherapeutic agents
Haemodialysis (long-term)	Balkan nephropathy	Radiation
Hypertension	Urothelial tumours of the urinary bladder	Chronic infection
Sickle cell trait		Schistosomiasis
		Bladder exstrophy
		Artificial sweeteners
		Human papillomavirus
		Gene mutation/deletions (*p15*, *p16*, *p53*, *ras*, *c-myc*)

appearance. Chromophobe renal cell carcinoma is relatively homogeneous with a light beige, tan–brown or brown colouration; however, higher grade tumours may have haemorrhage and necrosis. Papillary renal cell carcinoma is well circumscribed; it may have a similar colouration as the conventional (clear cell) type, or may be light tan to grey, depending on the number of foam cells in the stroma. Foci of necrosis, haemorrhage and cyst formation may be seen in any of the subtypes. Areas of sarcomatoid dedifferentiation are grey–white, more fleshy or firm. Pale yellow or pale tan firm areas of scar formation should not be confused for sarcomatoid dedifferentiation. Foci of calcification may also be present. The majority of the carcinomas are confined to the kidney, but some may show gross invasion into the perinephric adipose tissue or into the renal vein, rarely extending into the inferior vena cava and the right side of the heart.

Microscopic Features

Each of the different subtypes of renal cell carcinoma has distinct morphological features, which are detailed below.

Conventional (Clear Cell) Renal Cell Carcinoma

Conventional (clear cell) renal cell carcinoma is the most common subtype, representing 65–75% of renal cancer in most series. Tumour cells are arranged in sheets, nests, or tubules (**Figure 4; see colour plate section**). Uncommonly they may show a tubulocystic or, rarely, a papillary arrangement. Most tumour cells have clear cytoplasm; however, tumours can have a combination of cells with clear and granular eosinophilic cytoplasm. Tumours almost exclusively composed of cells with eosinophilic cytoplasm are rare. The clear cell appearance of conventional renal cell carcinoma is secondary to the lipid and glycogen content of the cells. The periodic acid Schiff (PAS) histochemical stain, with and without diastase, is the best method for demonstrating the cytoplasmic glycogen. One of the hallmark histological features is the delicate, interconnecting, sinusoidal-type of vasculature, sometimes likened to ‘chicken wire’ (**Figure 4**). This type of vascular pattern is generally not seen in the other subtypes of renal cell carcinoma and is a good clue to making the diagnosis of conventional renal cell carcinoma when cells with granular eosinophilic cytoplasm predominate. Thus, the diagnosis of conventional renal cell carcinoma is based on the architectural pattern and the vascular pattern, rather than the tinctorial properties of the cell cytoplasm.

Papillary Renal Cell Carcinoma

Papillary renal cell carcinoma accounts for about 10–15% of all renal carcinomas. Multifocal and bilateral tumours are more common in this subtype than any other subtype of renal cell carcinoma. Microscopically, it shows fibrovascular papillary cores lined by a single layer of low cuboidal epithelial cells that have scant pale cytoplasm and oval dark nuclei (**Figure 5; see colour plate section**). Some tumours have tall columnar pseudostratified cells with abundant eosinophilic cytoplasm. Tumour cells may have haemosiderin pigment within the cytoplasm. In addition to the papillae, tumour cells may form tubules, tubulopapillary structures and, rarely, solid nests. A characteristic feature, which may not be present in all tumours, is the presence of foamy macrophages within the fibrovascular stalks. Laminated calcifications (psammoma bodies) are also commonly present.

Chromophobe Renal Cell Carcinoma

Chromophobe renal cell carcinoma accounts for about 5% of all renal carcinomas and presents distinctive histological and ultrastructural features that clearly separate it from the other subtypes. There are two morphological variants, typical or classical chromophobe and the eosinophilic variant; this distinction is based on the tinctorial properties of the cytoplasm. These variants are not clinically important, but are important from the diagnosis point of view as the eosinophilic variant may be mistaken for renal oncocytoma, a benign tumour. Tumour cells are arranged in sheets, broad alveoli or nests. There are two types of cells, clear and eosinophilic (**Figure 6; see colour plate section**). Both types are usually present with one type predominating; if clear cells predominate, the tumour is referred to as typical or classical, and if the eosinophilic cells predominate, it is the eosinophilic variant. The clear cells have abundant clear cytoplasm with a frothy, vacuolated or bubbly appearance (soap bubble appearance) and prominent cell membranes, resembling plant cells. The cytoplasm typically forms a rim along the cell membrane that appears darker than the remainder of the cytoplasm. These cells also may have a perinuclear halo. The eosinophilic cells are smaller and have finely granular eosinophilic cytoplasm, with a variable degree of perinuclear clearing. The nuclei are hyperchromatic, frequently binucleate and have a wrinkled nuclear membrane. The nuclear features along with the perinuclear halo and the prominent cell membranes is referred to as ‘koilocytoid atypia,’ as it superficially resembles the changes seen in cells infected by the human papillomavirus. Hale’s colloidal iron stain is the histochemical stain of choice for confirming the diagnosis of chromophobe renal cell carcinoma; this stain shows diffuse, reticular staining.

Collecting Duct Carcinoma

Collecting duct carcinoma is rare, accounting for about 1% of all renal cell carcinomas. Microscopically, three features characterize this renal cancer; a tubulo-papillary arrangement of cells, desmoplastic reaction of the stroma and dysplastic changes in the adjacent collecting ducts. Dilated tubules or solid areas may also be present. In

general, before making a diagnosis of collecting duct carcinoma, metastasis to the kidney should be excluded. Renal medullary carcinoma is a distinctive type of collecting duct carcinoma with an aggressive clinical course, which arises in the renal medulla and is associated with sickle cell trait. In addition to the typical morphological features of collecting duct carcinoma, these cancers have reticular, microcystic areas, which resemble yolk sac tumour. Collecting duct carcinoma shows variable areas of mucin formation, which stain with histochemical stains for mucin, such as mucicarmine and Alcian blue.

Renal Cell Carcinoma, Unclassified

Unclassified renal cell carcinoma is not a distinctive subtype, but rather represents renal cancers that do not fit into one of the above-mentioned categories. Renal cancers in this category include, but are not limited to, those not conforming to any of the known histological subtypes, composites of recognizable subtypes, cancers with extensive necrosis and minimal viable tumour, mucin-producing tumours that are not collecting duct carcinoma and tumours with sarcomatoid dedifferentiation that do not have an epithelial element that can be readily assigned to one of the above categories (Störkel *et al*., 1997).

Renal Cell Carcinoma with Sarcomatoid Dedifferentiation

Sarcomatoid dedifferentiation is seen in approximately 1.5% of renal cell carcinomas. The term sarcomatoid dedifferentiation denotes anaplastic transformation of the renal cell carcinoma into a high-grade biphasic tumour that contains both malignant carcinomatous and mesenchymal elements. These tumours have been referred to in the past as carcinosarcoma or mixed mesodermal tumours. The incidence of sarcomatoid dedifferentiation, in our experience, varies amongst the different subtypes: 8% in conventional (clear cell), 3% in papillary renal cell carcinoma, 9% in chromophobe renal cell carcinoma, 27% in collecting duct carcinoma and 11% in renal cell carcinoma unclassified. The carcinoma component is usually high grade, at least a Fuhrman's nuclear grade 3, but may have any grade. The sarcomatous component may resemble pleomorphic malignant fibrous histiocytoma or an unclassified spindle cell sarcoma, or may show differentiation into bone, cartilage, skeletal muscle or blood vessels. It is important to differentiate a true sarcomatous component from benign spindle cells that are sometimes seen in renal cell carcinoma. The majority of these lesions are high stage at presentation and the prognosis is very poor. When reporting sarcomatoid dedifferentiation in renal cell carcinoma it is important to mention the percentage of the sarcomatoid component and the type of epithelial and sarcomatoid components. The percentage of the sarcomatoid component has been reported to be important for survival; patients with >50% sarcomatoid dedifferentiation in their cancer do poorly.

Immunohistochemistry

All renal cell carcinomas stain positive with immunohistochemical stains for cytokeratin cocktail (AE1/AE3 and CAM 5.2), low molecular weight cytokeratin and epithelial membrane antigen (EMA). Conventional (clear cell) and papillary renal cell carcinoma also stain with vimentin, which is an intermediate filament usually associated with mesenchymal structures. Chromophobe renal cell carcinoma, on the other hand, does not stain with vimentin, which may be utilized for distinguishing the eosinophilic variant of chromophobe renal cell carcinoma from conventional renal cell carcinoma with predominantly eosinophilic cells. Collecting duct carcinoma, however, has a unique staining pattern, reacting with both low and high molecular weight cytokeratins, peanut agglutinin, *Ulex europaeus* lectin and epithelial membrane antigen. This characteristic staining pattern is similar to that of the distal collecting tubule epithelium.

Electron Microscopy

Ultrastructurally, the cells of conventional (clear cell) renal cell carcinoma exhibit a brush border, tend to form microlumina and have a basal lamina that separates groups of cells from each other. Abundant glycogen and lipid are present in the cytoplasm. Chromophobe renal cell carcinoma has characteristic microvesicles, which are probably derived from the endoplasmic reticulum or from mitochondria. Mitochondria also impart the characteristic granularity to the cytoplasm seen by light microscopy.

Molecular Genetics

Sporadic conventional (clear cell) renal cell carcinoma typically (in approximately 80–90%) shows loss of genetic material from the short arm of chromosome 3, in the region 3p14–3p26 that harbours the *VHL* gene at 3p25.3. Mutations within the *VHL* gene region and inactivation of this gene by hypermethylation are common. Sporadic papillary renal cell carcinoma is characterized by trisomies, especially of chromosomes 7 and 17, and loss of the Y chromosome. Other chromosomes that may be involved include 3, 9, 11, 12, 16, and 20; some of these additional abnormalities are speculated to lead to progression to a more aggressive phenotype. Translocation between chromosomes X and 1 has also been reported, and is more common in children. As mentioned earlier, the familial cases of papillary renal cell carcinoma show germ-line mutations of the *MET* proto-oncogene. Chromophobe renal cell carcinomas are characterized by combined losses of multiple chromosomes including 1, 2, 6, 10, 12, 13, 14, 15 and 17. Polysomy of chromosome 7, trisomy 12, 16 and 19, telomeric associations and structural abnormalities of 11q have also been described in these cancers. Another important finding in chromophobe renal cell carcinoma is abnormalities of

mitochondrial DNA, a feature not seen in the other subtypes. Collecting duct carcinoma does not have any distinct genetic alterations. Monosomy of chromosomes 1, 6, 14, 15 and 22, deletions of 8p and13q, and loss of heterozygosity (LOH) in 1q have all been reported in this cancer.

Prognostic Factors

The prognosis of renal cell carcinoma is heavily dependent on the TNM stage of the tumour (Rini and Vogelzang, 2000). The 5-year survival rate of all patients with renal cell carcinoma is 70–75%. The overall 5-year survival rate is $\leq 10\%$ in patients with stage IV disease, 20–40% in patients with stage III tumours, 50–60% with stage II tumours and 60–90% with stage I tumours. The presence of sarcomatoid dedifferentiation also portends a poor prognosis, with disease-specific survival rates of 59 and 22% at 1 and 5 years, respectively. Fuhrman's nuclear grade has also been shown to have limited value in predicting prognosis. Up to 50% of patients with the VHL syndrome die of renal cell carcinoma. Newer markers that are under investigation for determining prognosis include DNA ploidy, proliferation marker MIB-1, p53, vascular endothelium growth factor (VEGF), *VHL* gene mutations, etc. None of these markers has so far been shown to be better at predicting tumour recurrence or death from disease, and all of these markers and molecular genetic alterations need to be studied further.

Overview of Present Clinical Management

Surgical resection in the form of radical nephrectomy remains the most effective treatment for clinically localized tumours. In recent years the use of nephron-sparing surgery has been gaining popularity for the treatment of small, localized tumours that can be easily resected (Wunderlich *et al*., 1998). Solitary metastasis of renal cell carcinoma may be resected, but with limited success. Immunotherapy using interferon and/or interleukin 2 (IL-2) has had partial success in treating metastatic disease, as it remains resistant to systemic chemotherapy. Newer therapies for metastatic renal cell carcinoma that are currently undergoing investigation include tumour-specific vaccines, angiogenesis inhibitors, monoclonal antibodies and dendritic cell therapy.

RENAL PELVIS AND URETER

Tumours of the renal pelvis and ureter are uncommon. Urothelial carcinoma accounts for 90% of the malignant tumours, but represents less than 5% of all tumours arising from the urothelium. The remaining 10% include squamous cell carcinoma and adenocarcinoma. Urothelial carcinomas of the renal pelvis and ureter have epidemiological, clinical, morphological and prognostic features similar to those arising in the urinary bladder (Steffens and Nagel, 1988). There are significantly different aetiological factors (**Table 6**) associated with these tumours including phenacetin abuse, thorium (radiographic contrast medium) exposure and Balkan nephropathy. Another characteristic feature is the association with urothelial tumours of the urinary bladder, which may occur synchronously or metachronously. Tumours arising in the ureter may fill the lumen, leading to obstruction and possibly hydronephrosis. All the different morphological variants of urothelial carcinoma reported in the urinary bladder have also been reported in the renal pelvis and ureter. The grading system is similar to that of the urinary bladder urothelial carcinomas. Also like the urinary bladder, the TNM staging system is based on the depth of invasion (**Table 7**).

Table 7 1997 TNM staging system for urothelial carcinoma of renal pelvis and ureter

Primary tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Ta	Papillary noninvasive carcinoma
Tis	Carcinoma *in situ*
T1	Tumour invades subepithelial connective tissue
T2	Tumour invades the muscularis
T3	(For renal pelvis only): tumour invades beyond muscularis into peripelvic fat or the renal parenchyma
T3	(For ureter only): tumour invades beyond muscularis into periureteric fat
T4	Tumour invades adjacent organs, or through the kidney into the perinephric fat
Regional lymph nodes (N)[a]	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single lymph node, 2 cm or less in greatest dimension
N2	Metastasis in a single lymph node, 2 cm but not more than 5 cm in greatest dimension; or multiple lymph nodes, none more than 5 cm in greatest dimension
N3	Metastasis in a lymph node more than 5 cm in greatest dimension
Distant Metastasis (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Stage Groupings			
Stage 0a	Ta	N0	M0
Stage 0is	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
Stage IV	T4	N0	M0
	Any T	Any N	M0
	Any T	Any N	M1

[a]Laterality does not affect the N classification.

Nephroureterectomy with a bladder cuff remains the treatment of choice for these cancers.

URINARY BLADDER

Up to 95% of urinary bladder tumours (**Table 8**) are of epithelial origin, of which 90% are urothelial neoplasms. The urothelium is a highly specialized epithelium that lines the entire urinary tract and has the ability to modify the number of layers forming it, depending on the level of distention of the organ wall. It has traditionally been referred to as the transitional cell epithelium, but urothelium is currently the preferred term (Epstein *et al*., 1998).

Other types of epithelial tumours arise in the urinary bladder, including squamous cell carcinomas and adenocarcinomas. These are usually diagnosed when the entire tumour is entirely composed of malignant squamous or glandular elements, as foci of squamous or glandular differentiation can be seen in otherwise typical urothelial carcinomas, more commonly in high-grade tumours. Small cell carcinoma and other neuroendocrine carcinomas of the urinary bladder have also been reported.

Table 8 Tumours of the urinary bladder

Benign tumours	Malignant tumours
Epithelial tumours	*Epithelial tumours*
Urothelial papilloma	Urothelial carcinoma (papillary noninvasive, papillary invasive, flat noninvasive, flat invasive, invasive NOS, micropapillary, lymphoepithelioma-like, nested)
Urothelial inverted papilloma	Squamous cell carcinoma
Villous adenoma	Adenocarcinoma
	Small cell carcinoma
	Undifferentiated carcinoma
Mesenchymal tumours	*Mesenchymal tumours*
Leiomyoma	Rhabdomyosarcoma
Neurofibroma	Leiomyosarcoma
Haemangioma	
Inflammatory myofibroblastic tumour	
	Miscellaneous neoplasms
	Lymphoma/leukaemia
	Plasmacytoma
	Malignant melanoma
	Paraganglioma
	Germ cell neoplasms
	Secondary involvement
	Direct extension from adjacent organs (rectum, bladder, uterus)
	Metastasis

Mesenchymal neoplasms, which range from benign proliferations to highly malignant and aggressive tumours, are much less common. The most common benign mesenchymal neoplasm is leiomyoma, while the most common sarcomas are rhabdomyosarcoma in children and leiomyosarcoma in adults. Lymphoma, plasmacytoma and infiltration by leukaemia rarely present in the urinary bladder without a previous history of the disease. Finally, secondary involvement of the urinary bladder by other tumours is rare, but when it occurs it is usually by direct extension from adjacent organs.

Urothelial Carcinoma

Tumour Pathology

Urothelial neoplasms arise by two distinct pathobiological pathways, resulting in the development of either papillary or nonpapillary tumours (see Molecular Genetics, below). Benign urothelial neoplasms, including exophytic and inverted urothelial papillomas, are much less frequent than carcinomas. Most urothelial carcinomas are preinvasive or early invasive papillary tumours. These tumours show a range of cytological and architectural atypia, which has been the basis for several proposed grading systems (see Microscopic Features, below). They are commonly multifocal, and may involve other sites along the urinary tract in addition to the urinary bladder. Papillary tumours tend to recur several times and eventually can invade the underlying lamina propria and muscularis propria. Less commonly, urinary bladder neoplasms start as nonpapillary or flat intraepithelial lesions, referred to as primary or *de novo* urothelial carcinoma *in situ*. These tumours also have the potential to invade the underlying wall. Up to 20% of urinary bladder neoplasms are diagnosed in an advanced stage, with invasion of the muscularis propria. When invasive, urothelial neoplasms have the potential to produce metastasis and death due to disease.

Epidemiology

It is estimated that 54 000 new cases of urinary bladder cancer will be diagnosed in the United States in the year 2001, constituting the fifth most common tumour in that country. Despite recent advances in treatment of both early and advanced urinary bladder tumours, an estimated 12 000 persons are expected to die of the disease in the United States in the year 2001. Urinary bladder cancer is more frequent in North America and Western Europe and uncommon in Japan. In the United States, it is twice as common in Caucasians as in African-Americans. It also has a 2.6-fold higher incidence in men than in women. Most tumours arise during late adulthood, with a median age at diagnosis of over 65 years.

Aetiology

Several aetiological factors have been associated with the development of urinary bladder cancer (Johansson and Cohen, 1997). This still growing list started in the nineteenth century when an increased incidence in urinary bladder tumours was noted among workers in the dye industry. Today, new data on the role of tumour-suppressor genes, oncogenes and other external influences have enhanced our knowledge of the pathogenesis of urinary bladder cancer (**Table 6**).

Tobacco

Approximately 33% of urinary bladder tumours (50–80% of tumours in men) are associated with smoking. The risk in smokers is increased 3–7-fold compared with nonsmokers, depending on the number of pack-years and smoking habits. Ex-smokers have a reduced risk, but the period of time necessary for them to return to the same risk level as nonsmokers is unknown.

Arylamines

Occupational exposure to arylamines, which are widely used in the aniline dye and rubber industry, has been associated with up to one-third of urinary bladder cancers. The risk is dependent on the intensity and duration of the exposure, and tumours usually appear 15–40 years after the first exposure. The insult to the urothelium results when the urinary enzyme glucuronidase splits the nontoxic glucuronate conjugate previously formed in the liver, releasing an electrophilic reactant. The process is highly specific to humans, as we are one of the few species to possess glucuronidase.

Radiation

There is an increased risk of developing urothelial carcinoma in women treated with radiation for cervical carcinoma. The tumours usually develop many years after exposure and are usually of high grade and high stage at presentation.

Cyclophosphamide

Cyclophosphamide, an alkylating agent used as an immunosuppressant, is associated with up to a 10.7% cumulative risk of developing urinary bladder cancer after 12 years of exposure.

Urinary Tract Infection

Chronic urinary tract infection also appears to confer a significant risk for urinary bladder neoplasia, especially in women. This mechanism also explains the higher incidence of urinary bladder neoplasms in paraplegic patients with indwelling catheters. Urinary bacteria increase the levels of volatile nitrosamines, believed to be mediators in the development of neoplasia.

Other Aetiological Factors

Phenacetin has been linked to the development of urothelial carcinoma of the renal pelvis. In urinary bladder tumours, however, only a weak association has been demonstrated, except when high cumulative doses have been used. Artificial sweeteners have been associated with urinary bladder tumours in animal studies. A single case-control study in humans reported a 1.6-fold increase in risk for men who use saccharin. These findings, however, have not been confirmed by other studies, and thus the aggregate data appear to show no significant risk of urinary bladder cancer for persons using artificial sweeteners. Similarly, some studies have suggested an increased risk in coffee and tea drinkers, but based on the aggregate data the risk level appears to be weak to nonexistent. Although still limited, the overall data also show little if any role for human papillomavirus in urinary bladder neoplasia.

Screening and Prevention

Screening for urothelial cancer is usually done in one of three settings: (1) screening individuals at high risk for urinary bladder cancer (i.e., exposure to known carcinogens), (2) screening individuals with microscopic haematuria or irritative voiding symptoms, and (3) monitoring patients following diagnosis and conservative local therapy of urothelial carcinoma (i.e., for detection of recurrences).

A variety of methods have recently been developed, most of which target molecules that are present more frequently in the urine of patients with urothelial carcinoma than in that of patients without urothelial carcinoma (**Table 9**). As yet, none of these methods has replaced urine cytology, which is inexpensive, noninvasive and still considered the 'gold standard' screening method (Brown, 2000).

Urine Cytology

Urine cytology is still the most common screening method for diagnosing urothelial carcinoma. Although it is primarily used for the follow-up of patients with a diagnosis of urothelial carcinoma, it has value in the screening of

Table 9 Screening and monitoring methods for urothelial carcinoma

Urine cytology
Bladder tumour antigen
Nuclear matrix proteins
Fibrin/fibrinogen degradation products
Telomerase activity
Hyaluronic acid/hyaluronidase
Cystoscopy and biopsy
Cystogram and excretory urography

high-risk populations. The main limitations of urine cytology are its low sensitivity for low-grade lesions and that its accuracy is dependent on the experience of the pathologist reviewing the specimen. Sensitivity for low-grade tumours has ranged from 0 to 100%, with most studies reporting less than 50%. For high-grade tumours, including carcinoma *in situ*, sensitivity and specificity approach 90 and 100%, respectively.

Bladder Tumour Antigen

Urothelial neoplasms secrete enzymes that lyse the urothelial basement membrane, exposing antigens that can be detected in the urine by a latex agglutination reaction. Initial studies have suggested that this method has a higher sensitivity than urine cytology, but other studies have found conflicting results. Further, specificity is significantly reduced in patients who have urinary tract infections or urolithiasis or have undergone prior instrumentation.

Nuclear Matrix Proteins

Nuclear matrix proteins, the structural framework of the nucleus, are present at low levels in the urine of normal individuals and in increased quantity in patients with urothelial carcinoma. An immunoassay that quantifies the amount of nuclear matrix proteins has been developed. Its sensitivity and specificity are dependent on the cut-off value. An overall sensitivity of 66% has been reported. Specificity is significantly diminished in the presence of gross haematuria and prostate cancer, but even in optimal cases the test does not appear superior to urine cytology.

Fibrin/Fibrinogen Degradation Products

Increased vascular permeability induced by a urinary bladder tumour results in leakage of fibrinogen and plasminogen. The fibrinogen is converted into an extravascular fibrin clot but is degraded by plasmin, which is activated by the urinary enzyme urokinase. Thus, the presence of fibrin or fibrinogen degradation products is indicative of urinary bladder carcinoma. The currently available assay for detecting these products has an overall sensitivity in the range 48–68% and a specificity of around 80%, which is also diminished in patients with gross haematuria and prostate cancer.

Telomerase

Degradation of telomeres after several cell cycles is a cell death mechanism. Telomerase's primary function is to repair the telomeres at the end of each cell cycle, delaying their degradation. Thus, over-activity of telomerases, prolongs cell life and may turn cells immortal. The method for measuring telomerase activity has a sensitivity of 70% and a specificity of up to 99% for detecting urinary bladder cancer. However, the assay is not yet widely available and requires laboratories with well-established expertise. False-positive results may occur in patients with inflammatory conditions, and false negatives may occur in patients with gross haematuria.

Hyaluronic Acid/Hyaluronidase

High levels of hyaluronic acid in the urine of patients with urinary bladder cancer can be detected with a sensitivity and specificity of 92%. Further, patients with high-grade neoplasms have elevated urinary hyaluronidase activity that can be detected with a 100% specificity. These still rather preliminary data suggest that this assay could be used for the diagnosis of low-grade lesions and the early detection of tumour progression.

Cystoscopy and Biopsy

Cystoscopy allows the direct visualization of the urinary bladder mucosa and biopsy of suspicious lesions. Urinary bladder washings obtained at the time of cystoscopic examination also provide useful cytological material. Carcinoma *in situ* is not always identified by cystoscopic examination, and thus random urinary bladder biopsies are usually performed to identify this lesion.

Cystograms and Excretory Urography

These methods provide useful information regarding the presence of multifocal disease involving the upper urinary tract and the localization of large urinary bladder masses that produce filling defects.

Gross Features

Noninvasive papillary tumours are recognized by cystoscopy as exophytic papillary fronds in the mucosa, usually of varying size. Up to 40% of tumours may be multifocal. Tumours occur most often in the lateral and posterior walls of the urinary bladder and least often in the dome. Urothelial carcinoma *in situ* usually appears as an erythaematous area in the mucosa of the urinary bladder, but it can also be grossly inapparent. When invasive, urothelial carcinomas are usually bulging, nodular exophytic tumours, or ulcerated and indurated. A papillary component may still be present, albeit focally. Some tumours display diffuse infiltration of the urinary bladder wall with relatively unremarkable mucosa, analogous to the so-called linitis plastica of the stomach. Large, fleshy, polypoid tumours that fill the urinary bladder cavity usually exhibit sarcomatoid features.

Microscopic Features

Papillary urothelial neoplasms range from small and benign to large and aggressive tumours. The distinction is usually based on the degree of cytological and architectural atypia, which is the basis of the grading schemes used for these neoplasms. Recently, a new grading system, the one used in this discussion, was proposed by a conjoint effort of the World Health Organisation (WHO) and the

Table 10 World Health Organisation/International Society of Urologic Pathology grading scheme for urothelial neoplasia

Papillary lesions	Flat lesions
Papillary hyperplasia	Flat urothelial hyperplasia
Urothelial papilloma	Reactive atypia
Inverted papilloma	Atypia of unknown significance
Papillary urothelial neoplasm of low malignant potential	Dysplasia (low-grade intraurothelial neoplasia)
Papillary urothelial carcinoma, low grade	Carcinoma *in situ* (high-grade intraurothelial neoplasia
Papillary urothelial carcinoma, high grade	

International Society of Urologic Pathology (ISUP) (**Table 10**) (Epstein *et al*., 1998). In this system, benign papillary tumours are referred to as papillomas. They are usually small, delicate, papillary epithelial proliferations attached to the mucosa by a thin fibrovascular stalk. The urothelial lining is identical with that of the normal urinary bladder mucosa. Papillomas represent less than 1% of tumours and are usually seen in younger patients. These lesions can recur, although probably at a lower rate than other higher grade lesions. Tumours with slightly more complex architecture, of larger size and lined by cytologically normal or minimally abnormal urothelium are referred to as papillary urothelial neoplasms of low malignant potential. They have a tendency to recur but not invade or cause death. These tumours were previously referred to as transitional cell carcinomas, grade 1 (using the old WHO classification system). Papillary tumours containing cytologically malignant cells are currently classified as low-grade and high-grade papillary urothelial carcinomas (previously referred to as grade 2 and grade 3 transitional cell carcinomas, respectively). These tumours have a higher degree of nuclear pleomorphism and disorganisation of the epithelial cells, progressively losing their resemblance to normal urothelium (**Figure 7; see colour plate section**). High-grade tumours usually have less well-developed, blunt and fused papillae, with occasional solid areas.

Flat urothelial lesions also display a range of cytological changes, lesions at the lower end of the spectrum being referred to as urothelial dysplasia and at the higher end as carcinoma *in situ* (**Figure 8; see colour plate section**). As opposed to papillary lesions, however, only the high-grade lesions (i.e. carcinoma *in situ*) are known clinically significant, worthy of therapeutic intervention. As stated before, carcinoma *in situ* may be primary (i.e. arising *de novo*), or more commonly secondary (i.e. arising in association with previous or concurrent urothelial carcinoma). In the latter setting, its presence confers increased risk of progression to invasive disease.

Invasion into the underlying wall is most common in high-grade papillary tumours or in patients with carcinoma *in situ*. The extent of invasion ranges from microscopic foci of tumours invading the lamina propria to bulky tumours extending through the urinary bladder wall into the perivesical adipose tissue. The invasive components usually display high nuclear grade and may grow in nests, cords or trabeculae of neoplastic cells, which infiltrate the muscle bundles and adipose tissue (**Figure 9; see colour plate section**).

Sarcomatoid urothelial carcinoma is a high-grade neoplasm that has partially or totally lost its carcinomatous morphological phenotype and shows differentiation into spindle-cell mesenchymal elements. A well-defined malignant heterologous mesenchymal component, (e.g. leiomyosarcoma, osteosarcoma, rhabdomyosarcoma) may be present, but most often the mesenchymal component is relatively undifferentiated and high grade. Diagnosis of sarcomatoid carcinoma requires the presence of a focal, unequivocal, invasive or *in situ* epithelial component or evidence of epithelial differentiation either by immunohistochemistry (i.e. cytokeratin expression) or electron microscopy (i.e. presence of desmosomes or other epithelial elements). Differentiation between sarcomatoid carcinoma and sarcoma of the urinary bladder is clinically relevant, as the prognosis is significantly worse for sarcomatoid carcinomas.

Immunohistochemistry

The urothelium is a complex epithelium that expresses a wide variety of cytokeratin intermediate filaments. The urothelium and its neoplasms express both low and high molecular weight cytokeratin, and cytokeratin 7 and cytokeratin 20; however, up to 40% of high-grade urothelial cancers are negative for cytokeratin 20. Carcinoembryonic antigen (CEA) is expressed in approximately 60% of urothelial carcinomas, most frequently in those that show glandular differentiation. Cytokeratin 10 and 14 are frequently expressed in areas of squamous differentiation. The urothelium usually lacks vimentin expression. Expression of other epithelial markers, such as epithelial membrane antigen and Leu-M1, has also been reported. Uroplakin and thrombomodulin are newer markers that are also useful for staining urothelial tumours. Immunohistochemistry has also been used to identify a wide range of molecular markers, mostly for prognostic purposes (see Prognostic Factors, below, for a brief review of these markers).

Molecular Genetic Findings

The data from cytogenetic and molecular studies of urinary bladder carcinoma support the two-pathway mechanism seen morphologically (Cordon-Cardo *et al*., 1997; Lee and Droller, 2000). Most low-grade papillary neoplasms that have been studied have been associated with deletions

of chromosome 9, including monosomy of chromosome 9 and 9p and/or 9q deletions. Chromosome 9 deletions are the only genetic change consistently present in noninvasive and early invasive papillary tumours and are thus believed to be an initial step in the development of these neoplasms. The 9p deletions involve the tumour-suppressor genes *p15* and *p16*, which encode inhibitors of cyclin-dependent kinase. The 9q deletions also likely include multiple tumour-suppressor genes, most of which are yet to be identified. On the other hand, urothelial carcinoma *in situ* is usually associated with deletions of 14q and deletions of 17p or mutations of the *p53* gene. These alterations are rarely seen in noninvasive, low-grade papillary tumours. Invasive tumours are similarly associated with mutations of p53 and deletions of 13q, the site of the *RB* gene. Other alterations observed in invasive tumours, but seldom in papillary tumours, include deletions of 11p, 3p, 8p, 6q, 4p, and 5q and increased expression of Ras, c-Myc and epidermal growth factor receptors.

Hence, these data support two distinct molecular pathways for urothelial carcinoma. On one side, alterations of chromosome 9 induce the formation of papillary tumours, which have low potential for progressing. If additional molecular alterations are acquired (*p53* mutations, 17p−, 3p−, 5q−), the tumours progress to high-grade lesions and invasion. On the other hand, *p53* mutations, 14q deletions and possibly other mechanisms result in urothelial carcinoma *in situ* with a propensity for progressing and developing invasive disease upon acquisition of additional alterations. Molecular events correlating with development of advanced disease and metastasis are also likely to be identified in the future.

Prognostic Factors

Numerous factors are used for evaluating the prognosis of urothelial carcinoma (**Table 11**). The extent of disease is the most valuable prognostic information used for the management of urothelial carcinoma. When dealing with noninvasive or at most lamina propria-invasive tumours, management is focused on local control of the disease and parameters in this setting are those that can predict the frequency of tumour recurrence and likelihood of progression to muscle-invasive disease. Once the tumour has invaded the muscularis propria layer (**Figure 9**), aggressive management (usually cystectomy or radiation) is usually indicated, and useful prognostic parameters are those that can help predict development of distant metastasis or local recurrence. Thus, the main prognostic indicator for urothelial carcinoma, and the one on which others are dependent, is tumour stage (Mazzucchelli *et al*., 1994).

Table 11 Prognostic factors in urothelial carcinoma

Pathological factors
Depth of invasion (stage)
Tumour grade
Histological type
Vascular invasion
Tumour size
Multicentricity
Associated carcinoma *in situ*
Biological/molecular factors
Blood group antigens
DNA ploidy
p53 mutations
Proliferation markers
Retinoblastoma gene mutations
p21^{WAF1}

Tumour Stage

Tumour stage is important not only because of the powerful prognostic information it conveys but also because it defines the management of the patient. The TNM system (**Table 12**) is the most commonly used staging system. Five-year survival for patients with T1 tumours is approximately 90% but drops to 70% for patients with T2 tumours and to 35% for those with T3 and T4 neoplasms (Mazzucchelli *et al*., 1994). Owing to the abrupt difference in clinical behaviour between tumours showing only lamina propria invasion and those showing muscularis propria invasion, there is a tendency to simplify the staging system into two categories: 'superficial' (i.e. pTa pTis, and pT1 tumours) and 'invasive' (i.e. pT2, pT3, and pT4 tumours). While this nomenclature identifies two subsets of urinary bladder cancer that are traditionally managed in substantially different fashion, it oversimplifies and diminishes the power of the staging system. Further, the inaccurate terminology is confusing and can lead to miscommunication between pathologists and clinicians.

Tumour grade

Prognostic significance of grading in urinary bladder tumours is mostly limited to noninvasive or early invasive papillary tumours. Carcinoma *in situ* is a high-grade lesion by definition, and flat lesions of lesser grade (i.e. dysplasia) imply urothelial instability, but do not warrant therapeutic intervention. In muscle-invasive disease, grade has not provided any more prognostic information than stage alone. However, grade has repeatedly been shown to predict tumour recurrence and progression in papillary neoplasia. Patients with papillary urothelial neoplasms of low malignant potential have a 98% 10-year survival, whereas high-grade tumours are associated with a 35% 10-year survival. High-grade tumours with invasion into the lamina propria account for 60% of tumours that progress to muscle invasion (Lapham *et al*., 1997).

Histological Tumour Type

The histological type of urinary bladder tumour has also been shown to have significant prognostic value.

Table 12 1997 TNM staging system for the urinary bladder

Primary tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Ta	Papillary noninvasive carcinoma
Tis	Carcinoma *in situ* 'flat tumour'
T1	Tumour invades subepithelial connective tissue
T2	Tumour invades muscle
T2a	Tumour invades superficial muscle (inner half)
T2b	Tumour invades deep muscle (outer half)
T3	Tumour invades perivesical tissue
T3a	Microscopically
T3b	Macroscpocially (extravesical mass)
T4	Tumour invades any of the following: prostate, uterus, vagina, pelvic wall, abdominal wall
T4a	Tumour invades prostate, uterus, vagina
T4b	Tumour invades pelvic wall, abdominal wall
Regional lymph nodes (N)[a]	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in a single lymph node, 2 cm or less in greatest dimension
N2	Metastasis in a single lymph node, more than 2 cm but not more than 5 cm in greatest dimension; or multiple lymph nodes, none more than 5 cm in greatest dimension
N3	Metastasis in a lymph node more than 5 cm in greatest dimension
Distant metastasis (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Stage groupings			
Stage 0a	Ta	N0	M0
Stage 0is	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T4a	N0	M0
Stage IV	T4b	N0	M0
	Any T	Any N	M0
	Any T	Any N	M1

[a]Regional lymph nodes are those within the true pelvis; all others are distant lymph nodes.

Squamous and glandular differentiation have been associated with poor responses to radiotherapy and to chemotherapy, respectively. Both are associated with high-stage urinary bladder cancer. Certain histological types of carcinoma, such as sarcomatoid carcinoma, small cell carcinoma and the variant of urothelial carcinoma known as micropapillary carcinoma, usually present at an advanced tumour stage (Lapham *et al*., 1997). A variant known as lymphoepithelioma-like carcinoma of the urinary bladder may be amenable to chemotherapy alone.

Vascular invasion

Vascular invasion in tumours with limited lamina propria invasion has been associated with worse prognosis. However, data on the impact of vascular invasion in urothelial carcinoma are conflicting and difficult to interpret, owing mostly to the fact that in the urinary bladder, vascular invasion is difficult to recognize histologically, is easily over-diagnosed and confirmation usually requires use of immunohistochemical stains for endothelium (Lapham *et al*., 1997).

Multicentric Tumours and Tumour Size

Tumour size, as assessed at time of cystoscopy, is correlated with muscularis propria-invasive disease. Similarly, multicentric tumours (i.e. presence of multiple tumours at the time of cystoscopy) are more likely to both recur and progress (Lapham *et al*., 1997).

Adjacent Carcinoma In Situ

In cases of papillary bladder tumours, the presence of carcinoma *in situ* in the adjacent mucosa has been associated with a higher risk of recurrence and progression than in tumours without carcinoma *in situ*. Thus, it is common practice to randomly biopsy urothelium that appears normal at the time of resection of a papillary tumor (Lapham *et al*., 1997). Studies have also shown that dysplasia in the adjacent urothelium is associated with higher risk of recurrence and progression.

Biological Prognostic Factors

A growing list of biological/molecular prognostic markers has been studied in the literature (**Table 11**) (Stein *et al*., 1998). Most of them, however, have not provided significant or consistent predictive power beyond traditional pathological parameters. Several of these markers are currently being evaluated in a prospective and/or randomized manner, in order to determine their true clinical significance.

Lack of expression of blood group antigens, which are commonly expressed in normal urothelium, has been associated with tumour progression in patients with urinary bladder neoplasms. DNA ploidy has been shown to predict recurrence and progression in stage Ta and T1 tumours, and aneuploidy has been associated with a better response to radiation therapy. Proliferation indexes, whether derived by DNA flow cytometry, thymidine labelling, Ki-67 immunohistochemical staining or other methods, are correlated with tumour grade, disease progression and poor survival. As stated before, p53 mutations have also been associated with progression of urinary bladder tumours and possibly with initiation of cancer via the carcinoma *in situ* pathway. Current data suggest that p53 status is most useful in the management of locally advanced urinary bladder neoplasms, by helping to select patients who would benefit most from urinary bladder preservation or chemotherapy. Other molecular/biological prognostic indicators suggested

in urinary bladder cancer include mutations in oncogenes, epidermal growth factor receptors, growth factors, adhesion molecules, angiogenesis inhibitors, and other cell-cycle regulatory proteins, such as the gene products of the retinoblastoma tumour-suppressor gene and $p21^{WAF1}$. Analysis of these markers is beyond the scope of this discussion, but a recent comprehensive review is available (Stein *et al.*, 1998).

Overview of Present Clinical Management

As stated above, clinical management is highly dependent on the stage of disease and includes local resection of a surface tumour, intravesical therapy, radical resection, radiation therapy and systemic therapy.

Stage Ta tumours are treated mainly with transurethral resection and fulguration of visible tumours. Patients at high risk for recurrence or progression (i.e. those who have high-grade tumours, large tumours, associated carcinoma *in situ* or multiple tumours) also benefit from intravesical therapy, including immunotherapy (*Bacillus Calmette-Guérin*, interferon) and chemotherapy (thiotepa, mitomycin C, doxorubicin). These patients require close follow-up with repeat cystoscopy and urine or bladder-wash cytology every 3 months for 2 years, every 6 months for a further 2 years and subsequently every year for the rest of their lives. Stage Tis lesions are treated mainly with intravesical therapy and require close surveillance owing to the high rate of progression to invasive disease. Stage T1 tumours are most often treated conservatively with local resection and intravesical therapy, although some physicians advocate cystectomy for high-grade T1 tumours. Standard treatment for muscularis propria invasive tumours (i.e. stages T2, T3 and T4) is radical cystectomy or cystoprostatectomy. Radiation therapy is also an acceptable treatment and is standard therapy in some parts of the world. Local or distant metastatic disease is treated with adjuvant systemic chemotherapy; doxorubicin, cisplatin and methotrexate being the most commonly used agents.

Other Carcinomas of the Urinary Bladder

Squamous cell carcinoma accounts for up to 75% of urinary bladder cancers in areas where schistosomiasis is endemic. In the developed world they constitute less than 10% of all urinary bladder cancers. The mean age of presentation for these tumours is 46 years, almost 20 years younger than that for urothelial carcinomas in the Western world (Johansson and Cohen, 1997). Men are affected more often than women. Patients usually present with haematuria or irritative symptoms and most tumours are usually high stage at the time of presentation. Infection with *Schistosoma haematobium* is endemic to the Nile river valley in Egypt and other parts of Africa. Its eggs are deposited in the urinary bladder wall, eliciting a chronic granulomatous inflammatory response, fibrosis, calcification and squamous or glandular metaplasia of the urothelium. Tumours associated with schistosomiasis are mostly squamous cell carcinomas (75%), with a smaller proportion of adenocarcinomas (6%). Schistosomiasis appears to induce urinary bladder neoplasia by promoting increased cell proliferation as a result of inflammation, which in turn provides an increased risk of spontaneous genetic mutations. High levels of nitrosamines have also been found in patients with schistosomiasis, as in patients with chronic urinary tract infections. In countries where schistosomiasis is not endemic, squamous cell carcinoma usually arises secondary to chronic irritation caused by factors such as calculi, urinary retention, indwelling catheters, etc. These tumours tend to be large and bulky, often necrotic, sometimes filling most of the lumen of the urinary bladder. Squamous cell carcinomas display similar histology to those occurring at other sites and include tumour islands of squamous cells showing intercellular bridges and keratinization. Typical areas of urothelial carcinoma are not present. Squamous metaplasia of the adjacent urothelium is usually present. Verrucous carcinoma is a variant of squamous cell carcinoma that rarely affects the urinary bladder. They are graded according to the amount of keratinization and degree of nuclear pleomorphism. The TNM staging system for urinary bladder tumours (**Table 12**) is also used for these tumours and is the most useful prognostic indicator. Radical cystectomy or cystoprostatectomy with or without radiation therapy is the mainstay of therapy for this tumour.

Adenocarcinomas of the urinary bladder are rare, forming less than 1% of urinary bladder cancers. There is a significant association between urinary bladder exstrophy and the development of urinary bladder adenocarcinoma, although the mechanism by which this occurs remains unknown. Adenocarcinomas may be classified according to their location into urachal and non-urachal types. Both types show a variety of histological patterns, including enteric (i.e. mimicking colonic adenocarcinoma), signet ring, mucinous (colloid), clear cell, mixed and not otherwise specified (NOS).

URETHRA

Tumours of the urethra are distinctly rare. Owing to their significant anatomical and pathological differences, tumours arising in the female urethra are discussed separately from those arising in the male urethra (Grabstald, 1973; Ray *et al.*, 1977; Amin and Young, 1997). In both settings, most tumours are epithelial neoplasms (**Table 13**), with only leiomyoma and melanomas of the urethra being common enough to deserve mention.

Carcinomas of the Female Urethra

Urethral carcinomas typically occur in the postmenopausal period. Clinically patients present with one or more of

the following symptoms: vaginal or urethral bleeding, dysuria, urinary frequency, incontinence, urinary tract infection, perineal or introital mass or urinary obstruction. Associated aetiological factors include trauma and infection, human papillomavirus for squamous cell and urothelial carcinomas and urethral diverticula for clear cell adenocarcinomas.

The histological spectrum of urethral carcinomas is similar to that of the urinary bladder, although the histological type is highly dependent on the sites where they arise. Tumours involving the distal urethra and meatus (approximately 70% of tumours) are usually squamous cell carcinomas (**Figure 10; see colour plate section**). They are aggressive neoplasms that present at an advanced stage. Carcinomas of the proximal urethra are either urothelial carcinomas or adenocarcinomas. Urothelial carcinomas account for approximately 20% of tumours and exhibit the same diversity of histology as that seen in the urinary bladder, including papillary, flat, noninvasive and invasive tumours. Frequently, they are preceded by, followed by or occur concurrently with a urinary bladder urothelial carcinoma. Adenocarcinomas account for approximately 10% of female urethral carcinomas. Approximately 40% of adenocarcinomas are clear cell adenocarcinomas and the remaining show colloid, signet ring or NOS histology. Clear cell adenocarcinomas are pathologically and clinically distinctive. Histologically, they display papillary, tubulocystic, tubular or solid patterns. The tubules are usually hollow or contain eosinophilic secretions. The lining cells show characteristic hobnail shape and typically display clear to eosinophilic cytoplasm. Patients with clear cell adenocarcinomas appear to have a slightly better prognosis than those with nonclear cell adenocarcinomas, with between 30% and 40% of patients dying within 24 months; the survival is significantly lower for other types of adenocarcinomas.

Table 13 Histological classification of carcinomas of the urethra

1. Primary Tumors
- a. Transitional cell carcinoma
- b. Squamous cell carcinoma
- c. Adenocarcinoma
  - i. Clear cell carcinoma
  - ii. Colloid (mucinous) carcinoma
  - iii. Signet-ring cell carcinoma
  - iv. Adenocarcinoma, not otherwise specified
- d. Adenosquamous carcinoma
- e. Undifferentiated carcinoma

2. Secondary Tumors
- a. Direct extension from adjacent organs
- b. Metastatic

The prognosis of female urethral carcinomas is relatively poor. Location of the tumour and stage of the disease are the most important prognostic parameters. Between 18% and 50% of women will have metastasis at presentation. Tumours restricted to the anterior urethra are more amenable to local surgical excision and have a better outcome than those involving the posterior or entire urethra. In 1973, Grabstald proposed a staging system specific for female urethral carcinomas (Grabstald, 1973) (**Table 14**). Five-year survival when this system is used is roughly 45% for stage A, 40% for stage B, 25% for stage C, and 18% for stage D. The 1997 TNM staging system (**Table 15**) is applicable to both male and female urethral tumours.

Carcinomas of the Male Urethra

Tumours of the male urethra occur in the sixth and seventh decades and usually present with obstructive symptoms or

Table 14 Staging systems in urethral carcinoma

Female urethral tumours: Grabstald system	1997 UICC/AJCC TNM system	Male urethral tumours: Ray system
	Ta Noninvasive papillary, polypoid or verrucous carcinoma	
	Tis Carcinoma *in situ*	
A Tumour invades submucosa	T1 Tumour invades subepithelial connective tissue	A Tumour invades lamina propria
B Tumour invades muscle	T2 Tumour invades corpus spongiosum, prostate or periurethral muscle	B Tumour invades substance of corpus spongiosum or prostate
C Tumour invades vagina, bladder, labia or clitoris	T3 Tumour invades corpus cavernosum, beyond prostatic capsule, anterior vagina or bladder neck	C Tumour invades into corpus cavernosum, fat or beyond the prostatic capsule
	T4 Tumour invades other adjacent organs	
D Metastatic disease	N1–3 Regional metastasis	D1 Regional metastasis
	M1 Distant metastasis	D2 Distant metastasis

Table 15 1997 TNM staging system for the urethra

Primary tumour (T)
TX Primary tumour cannot be assessed
T0 No evidence of primary tumour
Ta Noninvasive papillary, polypoid or verrucous carcinoma
Tis Carcinoma *in situ*
T1 Tumour invades subepithelial connective tissue
T2 Tumour invades any of the following: corpus spongiosum, prostate, periurethral muscle
T3 Tumour invades any of the following: corpus cavernosum, beyond prostatic capsule, anterior vagina, bladder neck
T4 Tumour invades other adjacent organs

Regional lymph nodes (N)
NX Regional lymph nodes cannot be assessed
N0 No regional lymph node metastasis
N1 Metastasis in a single lymph node, 2 cm or less in greatest dimension
N2 Metastasis in a single lymph node more than 2 cm greatest dimension, or in multiple lymph nodes

Distant metastasis (M)
MX Distant metastasis cannot be assessed
M0 No distant metastasis
M1 Distant metastasis

Stage groupings

Stage 0a	Ta	N0	M0
Stage 0is	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T1	N1	M0
	T2	N1	M0
	T3	N0	M0
	T3	N1	M0
Stage IV	T4	N0	M0
	T4	N1	M0
	Any T	N2	M0
	Any T	Any N	M1

infection. Tumours of the posterior urethra are associated with haematuria, purulent discharge or urinary obstruction. Some patients present with a palpable mass. Other symptoms include painful priapism, penile erosion and impotence. Suggested aetiological factors include chronic irritation from strictures or infectious processes. Human papillomavirus has been documented in squamous and urothelial carcinomas.

Grossly, tumours appear as ulcerative, nodular, papillary, cauliflower-like or firm solid masses. About 75% of tumours are squamous cell carcinomas, with the remainder being urothelial carcinomas, adenocarcinomas or undifferentiated carcinomas. Squamous cell carcinomas arise in a background of metaplasia. Urothelial carcinomas are seen most frequently in the prostatic urethra but may also occur in the membranous or penile urethra. They display the same histological spectrum as those arising in the urinary bladder. As in the female urethra, there may be a synchronous or metachronous urinary bladder tumour. Adenocarcinomas are less common in the male than in the female urethra, and clear cell adenocarcinoma is an oddity in the male urethra.

Prognosis is similarly dependent on location and extent of the disease. A staging system specific for male urethral cancer was proposed in 1977 by Ray *et al*. (Ray *et al*., 1977) (**Table 14**), but the TNM system (**Table 15**) is also applicable. Anterior tumours have a significantly better outcome than posterior tumours, as they can be managed successfully by surgery. Tumours of the anterior urethra generally metastasize to inguinal and external iliac lymph nodes, whereas posterior urethral tumours spread to internal iliac and hypogastric lymph nodes. Most often, the sign of treatment failure is patient's local recurrence, rather than distant metastasis.

Both male and female urethral tumours are initially treated for local excision if possible, including lymphadenectomy. Radiation therapy is indicated in cases that recur or in those not amenable to surgical excision. Adjuvant systemic therapy is indicated in metastatic disease.

REFERENCES

Amin, M. B. and Young, R. H. (1997). Primary carcinomas of the urethra. *Seminars in Diagnostic Pathology*, **14**, 147–160.

Boccon-Gibod, L. A. (1998). Pathological evaluation of renal tumors in children: international society of pediatric oncology approach. *Pediatric and Developmental Pathology*, **1**, 243–248.

Brown, F. M. (2000). Urine cytology. Is it still the gold standard for screening? *The Urologic Clinics of North America*, **27**, 25–37.

Cordon-Cardo, C., *et al*. (1997). Genetic studies and molecular markers of bladder cancer. *Seminars in Surgical Oncology*, **13**, 319–327.

Epstein, J. I., *et al*. (1998). The World Health Organization/International Society of Urological Pathology consensus classification of urothelial (transitional cell) neoplasms of the urinary bladder. Bladder Consensus Conference Committee. *American Journal of Surgical Pathology*, **22**, 1435–1448.

Faria, P., *et al*. (1996). Focal versus diffuse anaplasia in Wilms tumor–new definitions with prognostic significance: a report from the National Wilms Tumor Study Group. *American Journal of Surgical Pathology*, **20**, 909–920.

Grabstald, H. (1973). Proceedings: tumors of the urethra in men and women. *Cancer*, **32**, 1236–1255.

Johansson, S. L. and Cohen, S. M. (1997). Epidemiology and etiology of bladder cancer. *Seminars in Surgical Oncology*, **13**, 291–298.

Kovacs, G., *et al*. (1997). The Heidelberg classification of renal cell tumours. *Journal of Pathology*, **183**, 131–133.

Lapham, R. L., *et al*. (1997). Pathologic prognostic parameters in bladder urothelial biopsy, transurethral resection, and cystectomy specimens. *Seminars in Diagnostic Pathology*, **14**, 109–122.

Lee, R. and Droller, M. J. (2000). The natural history of bladder cancer. Implications for therapy. *The Urologic Clinics of North America*, **27**, 1–13.

Mazzucchelli, L., *et al*. (1994). Invasion depth is the most important prognostic factor for transitional-cell carcinoma in a prospective trial of radical cystectomy and adjuvant chemotherapy. *International Journal of Cancer*, **57**, 15–20.

McLaughlin, J. K. and Lipworth, L. (2000). Epidemiologic aspects of renal cell cancer. *Seminars in Oncology*, **27**, 115–123.

Neville, H. L. and Ritchey, M. L. (2000). Wilms tumor. Overview of National Wilms Tumor Study Group results. *The Urologic Clinics of North America*, **27**, 435–442.

Ray, B., *et al*. (1977). Experience with primary carcinoma of the male urethra. *Journal of Urology*, **117**, 591–594.

Rini, B. I. and Vogelzang, N. J. (2000). Prognostic factors in renal carcinoma. *Seminars in Oncology*, **27**, 213–220.

Schmidt, D. and Beckwith, J. B. (1995). Histopathology of childhood renal tumors. *Hematology and Oncology Clinics of North America*, **9**, 1179–1200.

Steffens, J. and Nagel, R. (1988). Tumours of the renal pelvis and ureter. Observations in 170 patients. *British Journal of Urology*, **61**, 277–283.

Stein, J. P., *et al*. (1998). Prognostic markers in bladder cancer: a contemporary review of the literature. *Journal of Urology*, **160**, 645–659.

Störkel, S., *et al*. (1997). Classification of renal cell carcinoma: Workgroup No. 1. Union Internationale Contre le Cancer (UICC) and the American Joint Committee on Cancer (AJCC). *Cancer*, **80**, 987–989.

Wunderlich, H., *et al*. (1998). Nephron sparing surgery for renal cell carcinoma 4 cm. or less in diameter: indicated or under treated? *Journal of Urology*, **159**, 1465–1469.

FURTHER READING

Bostwick, D. G. and Eble, J. N. (eds) (1997). *Urologic Surgical Pathology*, 1st edn (Mosby, Philadelphia, PA).

Murphy, W. M., *et al*. (eds) (1994). Tumors of the kidney, bladder, and related urinary structures. *Atlas of Tumor Pathology*, third series, fascicle 11 (Armed Forces Institute of Pathology, Washington, D.C.)

Vogelzang, N. J., *et al*. (eds) (2000). *Comprehensive Textbook of Genitourinary Oncology*, 2nd edn (Lippincott Williams & Wilkins, Philadelphia, PA).

Walsh, P. C., *et al*. (eds) (1998). *Campbell's Urology*, 7th edn (W. B. Saunders, Philadelphia, PA).

Raghavan, D., *et al*. (eds) (1997) *Principles and Practice of Genitourinary Oncology*, 1st edn (1997). (Lippincott-Raven, Philadelphia, PA).

Chapter 45

Male Genital Tract

Thomas M. Ulbright
Indiana University School of Medicine, Indianapolis, IN, USA

CONTENTS

• Prostate
• Penis
• Seminal Vesicle
• Testis

PROSTATE

Normal Development and Anatomy

The prostate gland develops as outgrowths of epithelium from the urogenital sinus that are surrounded by primitive mesenchyme. This epithelial component forms the ducts and acini of the peripheral and transition zones of the prostate, and the mesenchyme differentiates into its fibromuscular stroma. The distal prostatic urethra, formed from the urogenital sinus, is thus in continuity with the prostatic ducts. The central zone glands and ducts may derive from mesonephric duct remnants, similar to the development of the proximal prostatic urethra.

The prostate gland in young men weighs about 20 g and is shaped like a truncated cone with its base along the neck of the urinary bladder and its blunt apex at the urogenital diaphragm. The prostatic urethra courses through the gland with an angulation occurring at its approximate mid-portion at the level of the verumontanum, a mound-like elevation on the posterior urethral surface. The two ejaculatory ducts penetrate the base of the prostate and run in an anteroinferior direction through the gland to empty into the prostatic urethra, just lateral to either side of the verumontanum.

The prostate gland is currently described in terms of its zonal anatomy according to the McNeal model (McNeal, 1981). In this model, there are three major prostatic zones (**Figure 1**). The central zone accounts for about 25% of the prostatic volume and occupies a pyramidal-shaped region with its base at the bladder neck and apex at the point of angulation of the prostatic urethra. The transition zone represents only about 5% of the prostatic volume and consists of two small lobes just lateral to the proximal prostatic urethra. The remainder of the glandular prostate (approximately 70%) is represented by the peripheral zone, which surrounds both the central and transition zones in the basal aspect of the gland and constitutes essentially all of the glandular prostatic tissue distal to the urethral angulation. The prostate is surrounded by a condensation of stroma along its peripheral and lateral aspects that blends, anteriorly, with the anterior fibromuscular stroma. No true capsule is present at the apex and bladder neck aspects of the gland.

The major blood supply to the prostate derives from the branches of the internal iliac arteries. These branches enter the gland with the major neurovascular pedicles at either superolateral aspect of the gland (**Figure 2**). Venous drainage occurs through a plexus of veins in the capsular region and into the internal iliac veins. The prostatic lymphatics mainly drain into the lymph nodes along the internal iliac vessels, but direct drainage into the external iliac nodes may also occur.

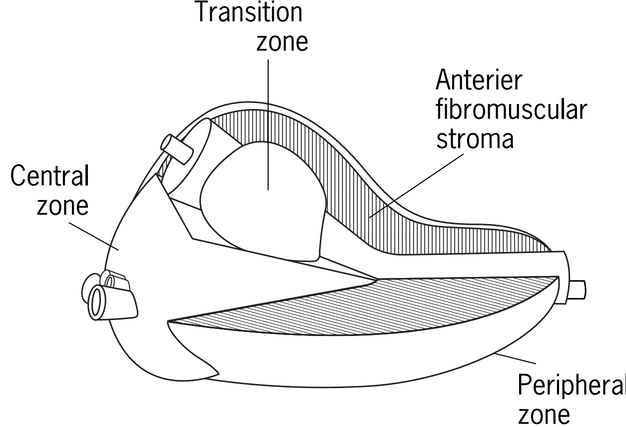


Figure 1 The McNeal model of prostatic anatomy. (Adapted from Lee *et al.*, 1989.)

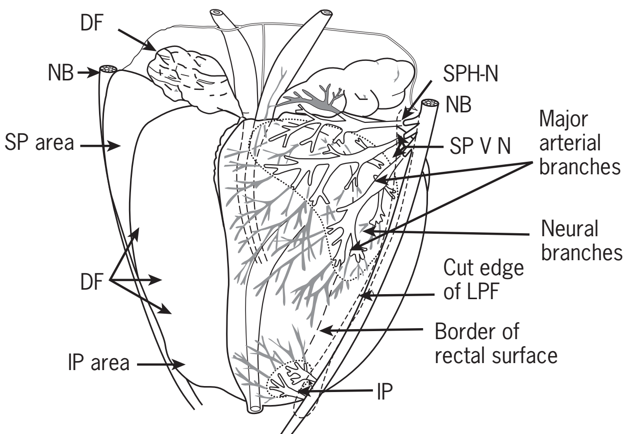


Figure 2 Posterolateral view of the prostate showing the major arterial and neural branches. (Adapted from McNeal, 1997.)

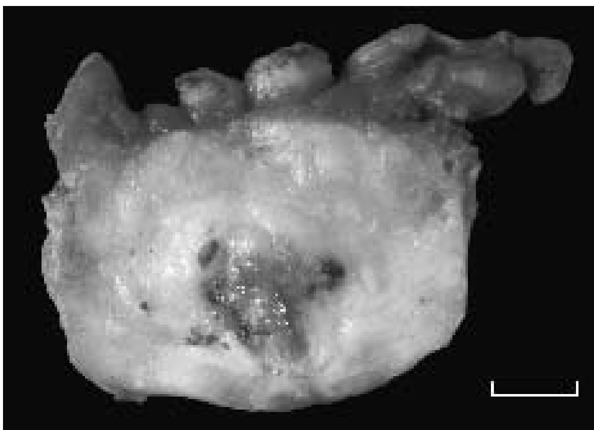


Figure 3 Cut surface of a prostate gland in the anterior–posterior plane shows yellow nodules of adenocarcinoma in the peripheral zone at bottom right. Bar, 1 cm.

Tumour Pathology

A diverse number of different types of neoplasm may occur in the prostate gland, but the overwhelming majority (>95%) of malignant tumours of the prostate are adenocarcinomas. With the elaboration of the McNeal model of prostatic anatomy (see above), it has become clear that different zones of the prostate have different propensities to develop adenocarcinoma. Accordingly, about 75% of adenocarcinomas develop in the peripheral zone, about 15% in the transition zone and the residual cases in the central zone (McNeal *et al*., 1988). Because much of the peripheral zone abuts the anterior wall of the rectum, some prostate carcinomas, when sufficiently large, may present palpable abnormalities on digital rectal examination (DRE). However, DRE is not an especially sensitive or specific method for detection; it identified only 56% of patients who were ultimately diagnosed with prostate cancer in a screening programme and had a positive predictive value of 21% (Catalona *et al*., 1994).

Prostatic adenocarcinoma may not be grossly apparent. In one series that examined radical prostatectomy specimens, 63% of tumours were correctly recognized on gross examination; no tumour was apparent in 22% of cases, and tumour was falsely identified in 19% of the cases (Renshaw, 1998). When they are recognizable, the tumours often have a yellow to grey solid character that contrasts with the frequently spongy appearance of the adjacent non-neoplastic prostatic parenchyma (**Figure 3**). Those tumours that develop in the transition zone may sometimes be grossly recognizable in prostate chips that have a tan to yellow discoloration.

Prostatic adenocarcinoma is recognized microscopically through a combination of architectural and cytological features. Typically, the malignant glands are smaller in size than the associated benign acini, and often, but not always, tend to be arranged in haphazard or infiltrative patterns that contrast with the more organized, lobular arrangement of benign glands and patterns that tend to be produced by benign pathological processes.

The grading of prostatic adenocarcinomas may be performed by utilization of a variety of schemes, but the one in most widespread usage is the Gleason method (Gleason, 1966; Gleason and Mellinger, 1974) that has been shown to be of significant prognostic value in a variety of studies (Thomas *et al*., 1982; Gleason, 1992; Albertsen *et al*., 1999). This system relies on the architectural assessment of tumour growth according to five histological grades (**Figure 4**). Those tumours whose growth is a well-circumscribed collection of uniform small glands are assigned a grade of 1 (**Figure 5**) and, on the other end of the scale, grade 5 tumours are those having a diffuse sheet-like arrangement of malignant cells, isolated single cells infiltrating the stroma (**Figure 6**), and/or nests with central necrosis (**Figure 7**). Tumours with features intermediate between these two extremes receive an intermediate grade. Since many prostatic adenocarcinomas are heterogeneous with respect to their growth patterns, the Gleason method furthermore sums the grades of the two most prominent patterns to arrive at a 'pattern score' that varies from 2 to 10. When there are two patterns of close to equal proportions that are in competition for the second grade, the one with the higher grade is chosen. If only one grade is present, its value is doubled to determine the pattern score. Recent research also demonstrates that even small amounts of grade 4 or grade 5 tumours are prognostically important, and there may be a role for the assessment of so-called 'tertiary grades' in prostate cancer (Pan *et al*., 2000).

The Gleason method, in prostatectomy specimens, correlates with disease-free survival and overall survival

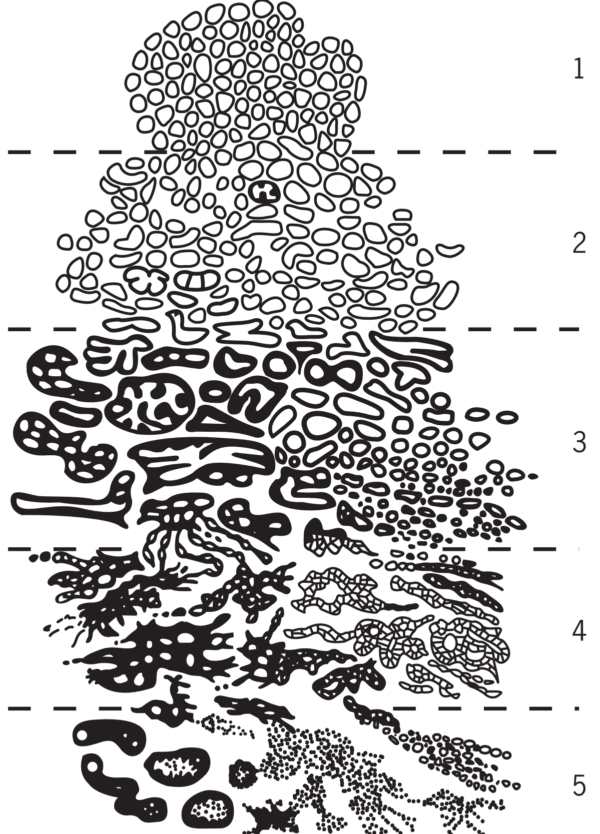


Figure 4 The Gleason grading system of prostate carcinoma. There are five grades based on the architectural pattern of tumour growth. (Adapted from Gleason, 1992.)

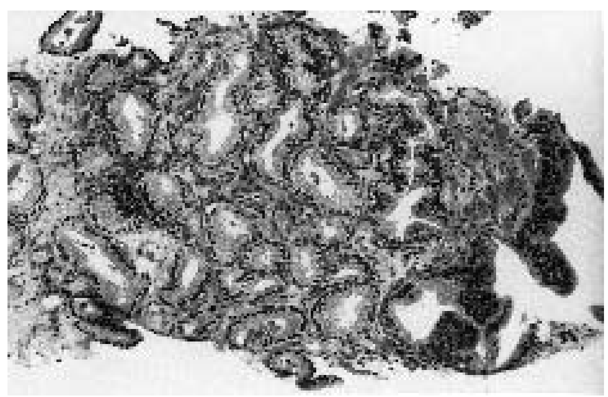


Figure 5 This well-circumscribed collection of relatively uniform-sized glands represents a Gleason grade 1 or grade 2 adenocarcinoma of the prostate.

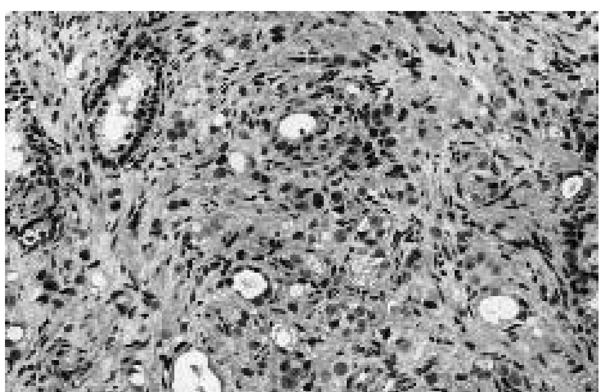


Figure 6 Atypical cells that infiltrate the stroma as irregular cords and single cells represent a grade 5 adenocarcinoma of the prostate. Note invasion between residual benign glands (left).

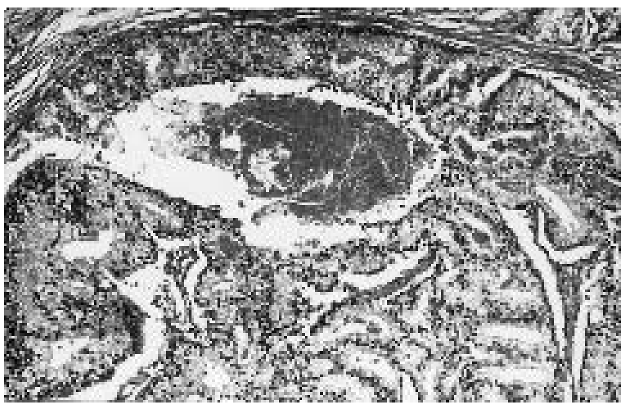


Figure 7 A cribriform pattern of adenocarcinoma of the prostate with zones of necrosis, a Gleason grade 5 tumour.

(**Figure 8**). Because of sampling problems, needle biopsy specimens are subject to grading disparity with respect to prostatectomy specimens, most commonly undergrading secondary to failure to sample a high-grade component. Such undergrading occurs in approximately 40% of the cases and most typically is a problem with small amounts of low-grade tumour on needle biopsy. Nevertheless, Gleason grading of needle biopsies and transurethral resectates does correlate with the final Gleason grade in prostatectomy specimens and also correlates with disease-free and overall survival.

In addition to the architectural patterns that are the initial basis for concern and that provide the Gleason pattern score, the diagnosis of cancer must be confirmed on cytological grounds. This is necessary because of several benign entities, such as atypical adenomatous hyperplasia (adenosis), sclerosing adenosis, prostatic extension of nephrogenic adenoma (nephrogenic metaplasia), mesonephric remnant hyperplasia and verumontanum mucosal gland hyperplasia, that may mimic prostatic adenocarcinoma on architectural grounds, particularly in small tissue specimens. There are several cytological features that, together, provide strong reassurance concerning the diagnosis of carcinoma (**Figure 9**). These include nucleolar prominence (usually considered in excess of 1.5 μm in

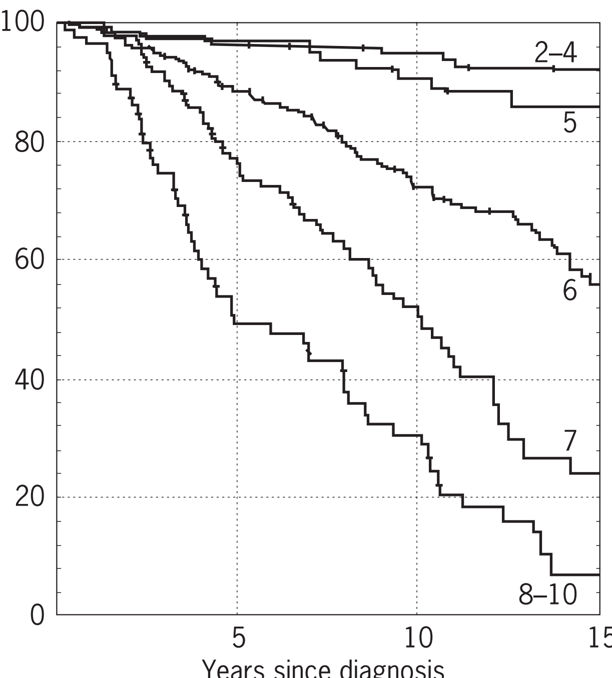


Figure 8 The relationship of Gleason pattern score to disease-specific survival in prostate adenocarcinoma. (Adapted from Albertsen *et al.*, 1999.)

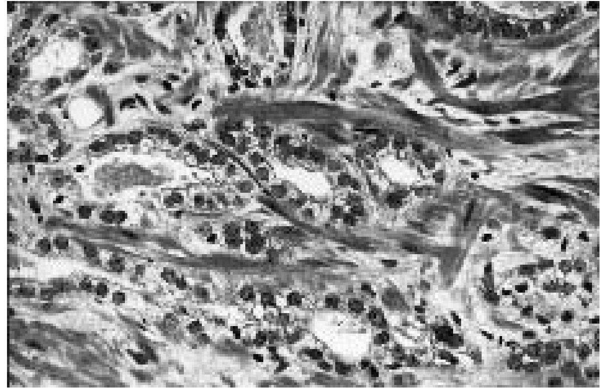


Figure 9 This adenocarcinoma of the prostate shows nuclear enlargement, hyperchromasia, prominent nucleoli and absence of a basal cell layer. The neoplastic glands intercalate between bundles of stromal collagen.

diameter), nuclear enlargement (compared with adjacent benign glands), nuclear hyperchromasia and irregularity, glands with an absence of basal cells and amphophilic cytoplasm (in contrast to the more pale cytoplasm of benign glands). Ancillary features that support a diagnosis of adenocarcinoma, but that are by no means specific for it,

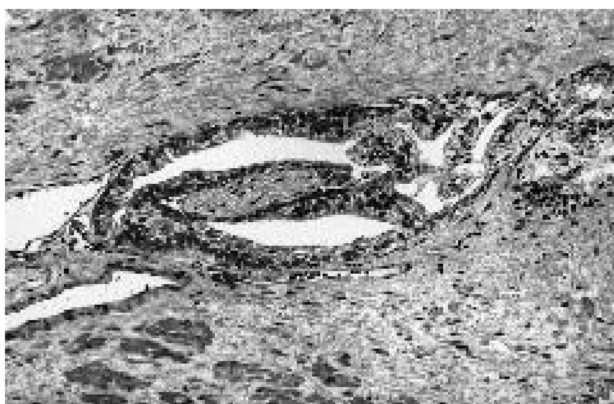


Figure 10 Circumferential perineural invasion in prostate adenocarcinoma, a finding considered pathognomonic of cancer.

include the presence of intraluminal acid mucins (usually visible as wispy, slightly basophilic intraluminal material on haematoxylin and eosin-stained sections) and the occurrence of intraluminal, eosinophilic, rod-shaped or polygonal crystalloids.

Only a few features are considered pathognomonic of carcinoma; these include huge nucleoli ($>3\,\mu m$), mucinous fibroplasia (also termed 'collagenous micronodules'), the occurrence of sieve-like, cribriform architectural patterns (**Figure 7**), and circumferential perineural invasion (**Figure 10**). The occurrence of any of these 'pathognomonic' features in biopsy material is rare, so the diagnosis of adenocarcinoma usually rests upon a constellation of features that vary in significance but that together permit its accurate and reproducible recognition.

There are two putative precursors for prostatic adenocarcinoma, one well established and the second of doubtful premalignant potential. The well-established precursor is now termed 'high-grade prostatic intraepithelial neoplasia' (PIN) (Bostwick and Brawer, 1987), although it has also been termed 'atypical hyperplasia' and 'intraductal dysplasia.' The poorly established possible precursor is known most commonly as either 'atypical adenomatous hyperplasia' or 'adenosis'.

PIN is considered a spectrum of atypical prostatic epithelial change that develops in pre-existing ducts and acini (**Figure 11**). Changes that include mild nuclear enlargement and stratification without nucleolar prominence characterize 'low-grade PIN,' a purely descriptive term that is not considered a diagnostic entity because of its unclear clinical significance and its poorly reproducible recognition. Cells resembling those of adenocarcinoma characterize high-grade PIN. Thus, on scanning magnification, one can appreciate the presence of a greater degree of basophilia in glands affected by high grade PIN than in normal glands (**Figure 12**). This reflects the cellular crowding, stratification, nuclear enlargement and nuclear hyperchromasia in these glands (**Figure 13**). As in

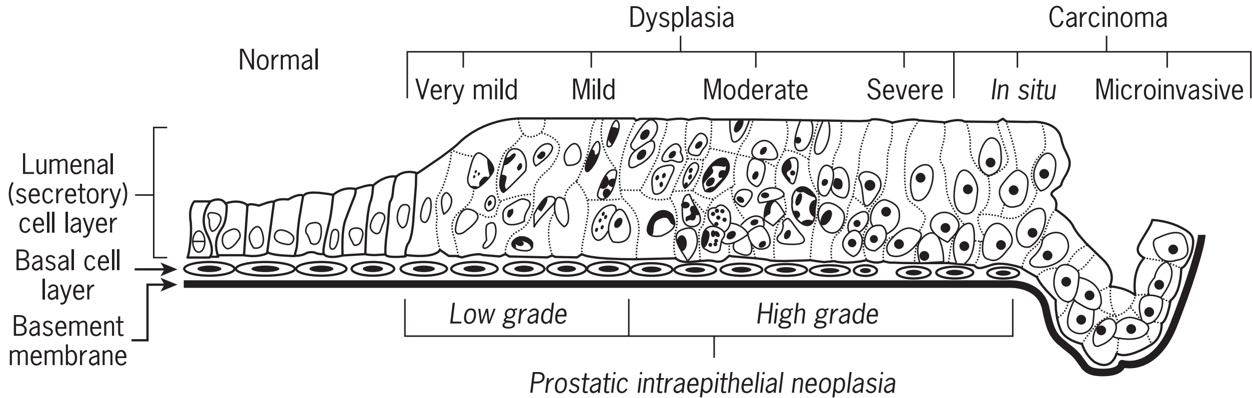


Figure 11 Drawing showing the spectrum of atypical epithelial changes in prostatic intraepithelial neoplasia (PIN), with final progression to invasive carcinoma. Cellular stratification and nucleolar prominence characterize high grade PIN. (Adapted from Bostwick, 1997.)

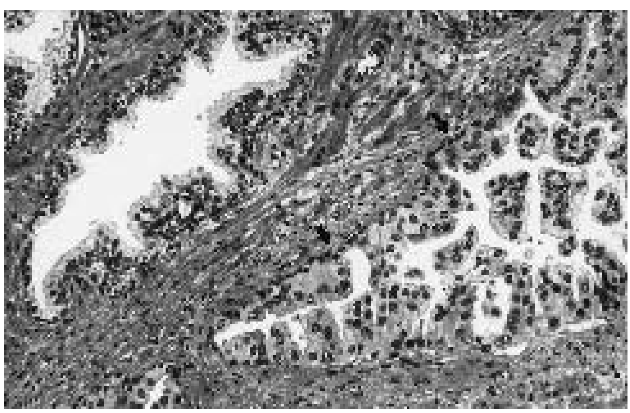


Figure 12 A normal prostate gland (left) contrasts with one having a papillary configuration of high grade PIN (right). The latter shows nuclear enlargement and cellular stratification, but there remains a basal cell layer (arrows).

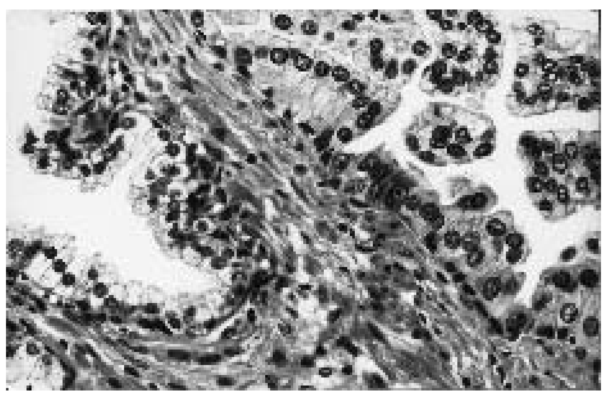


Figure 13 The nuclei in a gland with high-grade PIN (right) are focally stratified, have prominent nucleoli and are larger than those in an adjacent, normal gland (left). The gland with PIN still has an identifiable basal cell layer.

prostatic adenocarcinoma, nucleolar prominence is a key finding in high-grade PIN. In contrast to the acini of prostatic adenocarcinoma, the glandular structures of high-grade PIN have a residual basal cell layer, although it may be attenuated or focally absent. In the differential diagnosis of adenocarcinoma versus high-grade PIN, the absence or presence of a residual basal cell layer is thus a key finding. Several architectural patterns of high-grade PIN are common, caused by the piling up of increased numbers of epithelial cells in pre-existing duct/acinar units. Thus, tufted, papillary, cribriform and flat patterns occur, usually two or more together. Some authors have argued that the cribriform pattern of PIN is not a precursor to invasive adenocarcinoma, but reflects the growth of invasive adenocarcinoma into the duct/acinar system of the prostate, even though the invasive component may not be apparent in the available specimen (McNeal and Yemoto, 1996). This conclusion is based on the rare finding of cribriform 'PIN' in cancer-bearing prostates at sites remote from the tumour, whereas other patterns of 'PIN' are common at such foci.

There are numerous pieces of evidence that link high grade PIN with adenocarcinoma (**Table 1**). These include similar ploidy values, genetic and cytogenetic changes in high-grade PIN and the associated adenocarcinoma, the similar cytological appearance of both processes (as mentioned above), similar ultrastructure, morphometric values, topographical distribution in the prostate (i.e. almost always in the peripheral zone), lectin binding patterns and the occurrence of so-called transitive glands. The latter are adenocarcinomatous glands that appear to 'bud' from foci of high-grade PIN. There is a much higher frequency of high-grade PIN in prostates with cancer than in those without cancer (McNeal and Bostwick, 1986; Bostwick and Brawer, 1987). Most importantly, there is a significant risk for the subsequent identification of adenocarcinoma of the prostate in a patient who has

Table 1 Evidence for the association of high-grade prostatic intraepithelial neoplasia (PIN) and prostatic carcinoma. (Adapted from Bostwick, 1997.)

Histology	Immunophenotype (continued)
Similar architectural and cytological features	For some biomarkers there is progressive increase in expression with increasing grades of PIN and cancer, including type IV collagenase, TGF-alpha, EGF, EGFR, Lewis Y antigen and *c-erb-2* oncogene
Location	Morphometery
Both are located chiefly in the peripheral zone and are multicentric	High-grade PIN and cancer have similar nuclear area, chromatin content and distribution, nuclear perimeter, nuclear diameter and nuclear roundness.
Close spatial association of PIN and cancer	High-grade PIN and cancer have similar nucleolar number, size and location
Correlation with cell proliferation and death (apoptosis)	DNA content
Growth fraction of PIN is similar to cancer	High-grade PIN and cancer have similar frequency of aneuploidy
Number of apoptotic bodies in PIN is similar to cancer	Genetic instability
Apoptosis-suppressing oncoprotein bcl-2 expression is increased in PIN and cancer	High-grade PIN and cancer have similar frequency of allelic loss
Loss of basal cell layer	High-grade PIN and cancer have similar foci of allelic loss
The highest grade of PIN has loss of basal cell layer, similar to cancer	Microvessel density
Increased frequency of PIN in the presence of cancer	Progressive increase in microvessel density from PIN to cancer
Increased extent of PIN in the presence of cancer	Origin
Increased severity of PIN in the presence of cancer	Cancer found to arise in foci of PIN
Immunophenotype	Age
PIN is more closely related to cancer than benign epithelium	Age incidence peak of PIN precedes cancer
For some biomarkers there is progressive loss of expression with increasing grade of PIN and cancer, including prostrate-specific antigen, neuroendocrine cells, cytoskeletal proteins and secretory proteins	Predictive value of high-grade PIN
	PIN on biopsy has high predictive value for cancer on subsequent biopsy

high-grade PIN identified on needle biopsy. Several studies have shown up to a 50% frequency of invasive adenocarcinoma in second biopsy procedures of patients whose initial biopsy showed high-grade PIN. Whereas high-grade PIN and conventional acinar adenocarcinoma both occur disproportionately in the peripheral zone of the prostate, the other putative precursor for adenocarcinoma, atypical adenomatous hyperplasia (AAH), is usually seen in the transition zone. Hence AAH is sampled most commonly in transurethral resections specimens and is uncommon in needle biopsy material.

AAH is an architecturally worrisome lesion that consists of a circumscribed proliferation of tightly packed, relatively small glands, usually having abundant pale cytoplasm (**Figure 14**). Hence it is most likely to be confused with a Gleason grade 1 or grade 2 adenocarcinoma. Careful inspection, however, usually discloses that there is a greater variation in gland size and shape than in the typical low-grade adenocarcinoma, and many of the glands of AAH have an undulating border that contrasts with the straight luminal edges of adenocarcinoma. Most importantly, AAH lacks the cytological features of carcinoma, and basal cells are present, although frequently inconspicuous, often requiring high molecular weight cytokeratin immunostains for their demonstration.

Unlike high-grade PIN, the evidence linking AAH to adenocarcinoma is scant. One study identified a higher rate of subsequent carcinoma in patients who were initially diagnosed with 'adenosis' (Brawn, 1982), but that study has been criticized for misinterpreting some cases of low-grade adenocarcinoma for 'adenosis.' There

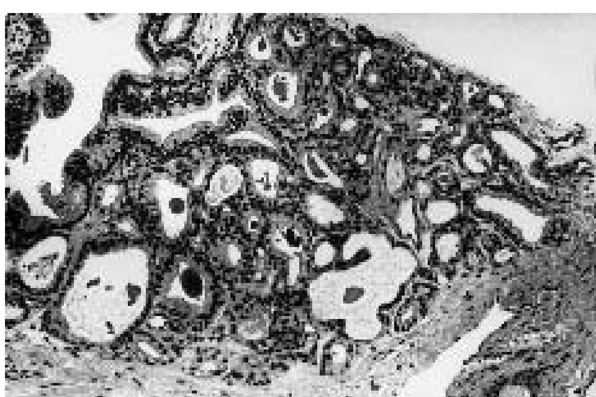


Figure 14 Atypical adenomatous hyperplasia consists of a collection of small glands that mimics the architectural features of a Gleason grade 1 or grade 2 adenocarcinoma. The close relationship to a 'parent' duct is a helpful distinguishing feature. At high magnification, these glands lacked the cytological features of adenocarcinoma.

is unquestionably a degree of morphological similarity between AAH and the typically low-grade adenocarcinomas that tend to occur in the transition zone of the prostate. Additionally, AAH, on average, tends to occur in patients who are 5–10 years younger than those with adenocarcinoma (Srigley, 1988; Amin *et al.*, 1993), supporting the evolution of the latter from the former. In addition, AAH has proliferation rates intermediate between those of normal prostate and adenocarcinoma. Disparate results have been obtained concerning cytogenetic abnormalities in AAH. One study found loss of heterozygosity (LOH) in 47% of cases by utilizing polymorphic microsatellite probes to chromosomes 7q, 8p, 8q and 18q (Cheng *et al.*, 1998). Another study, however, using probes to 1q, 6q, 7q, 8p, 10q, 13q, 16q, 17p, 17q and 18q, identified LOH that was confined to chromosome 8p in only 12% of AAH cases (Doll *et al.*, 1999). The former provided support for AAH as a neoplastic precursor but the latter questioned its neoplastic potential.

Immunohistochemistry may assist in the evaluation of prostatic adenocarcinoma. The secretory cells of the non-neoplastic prostate stain positively for prostate-specific antigen (PSA) and prostatic acid phosphatase (PAP), and these substances are also identified in all but a very small percentage of poorly differentiated adenocarcinomas of the prostate, although the staining can be focal in many tumours. Immunostains for PSA and PAP may therefore be of help in distinguishing prostatic carcinoma from transitional cell carcinoma that has extended from the bladder. In addition, certain mimics of adenocarcinoma, such as nephrogenic adenoma (nephrogenic metaplasia), hyperplasia of mesonephric remnants, adenotic seminal vesicle and prostatic xanthoma, are negative for PSA and PAP, so immunostaining can help in their differentiation from prostatic adenocarcinoma. Also helpful are immunostains using antibodies directed against high molecular weight cytokeratins, including the most commonly used one, clone 34βE12. Such stains highlight the basal cells that are present in non-neoplastic glands but that are absent in adenocarcinoma. In a worrisome lesion that is suspicious for adenocarcinoma by routine light microscopy, the absence of high molecular weight cytokeratin in the worrisome focus may provide sufficient additional evidence to permit the diagnosis of adenocarcinoma (**Figure 15**). Conversely, the presence of high molecular weight positive basal cells in the worrisome lesion is strong evidence that it does not represent invasive adenocarcinoma.

In addition to the usual form of adenocarcinoma, a number of specialized forms of prostatic carcinoma exist. One such type is the small cell carcinoma that, pathologically, resembles the much more common form of lung cancer. Unlike conventional adenocarcinoma, small cell carcinoma may not be associated with PSA elevations. For patients who have *de novo* small cell

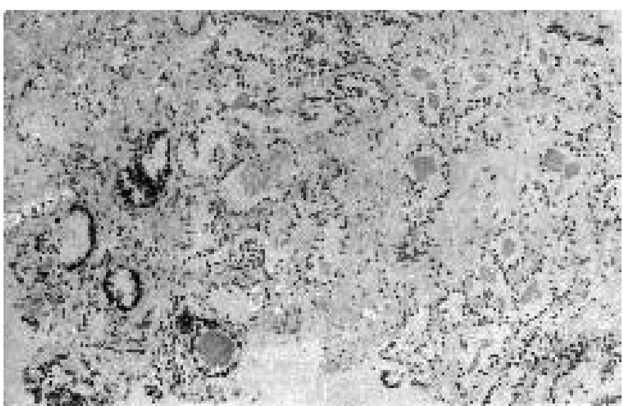


Figure 15 A high molecular weight cytokeratin immunostain is negative in a small acinar proliferation in the prostate, supporting that it is adenocarcinoma. A few non-neoplastic glands (left) show positively staining basal cells.

carcinoma, most have advanced stage disease at presentation. About half of small cell carcinomas are seen in patients who had a conventional acinar adenocarcinoma that failed hormonal or chemotherapeutic management, suggesting that many cases derive from therapy-resistant clones as a treatment selection phenomenon. The neuroendocrine nature of this tumour can be demonstrated with immunohistochemical markers, such as synaptophysin and chromogranin. Other types of prostatic carcinoma include prostatic duct, mucinous, signet-ring cell, adenosquamous, squamous, basaloid/adenoid cystic and transitional cell (i.e. of prostatic urethral or periurethral duct origin) types. It is controversial if the prostatic duct type of carcinoma merits separate classification since it is considered by some authorities to represent a conventional acinar adenocarcinoma with an unusual growth pattern caused by its extension into large periurethral ducts. Most continue to classify it separately but grade it using the Gleason method. In support of this approach, the clinical presentation is sometimes completely different from conventional adenocarcinoma. Patients with prostatic duct carcinoma may present with haematuria and, on cystoscopic examination, a mass may be seen protruding into the prostatic urethra. These features would not be expected with a conventional adenocarcinoma. The prior notion that prostatic duct carcinoma is derived from Mullerian tissues of the prostatic utricle (hence 'endometrioid' carcinoma of the prostate) is largely discredited on the basis of its positive staining with both of the usual markers of prostatic differentiation, PSA and PAP.

In addition to carcinomas, the prostate is also subject to much rarer mesenchymal tumours. Although virtually any form of sarcoma may be encountered in the prostate on occasion, only a select few merit special mention. Embryonal rhabdomyosarcoma of the prostate has an appearance similar to that seen in the soft tissues and

tends to occur in children and young adult patients, but it may rarely be seen in older adults. There also occur mixed epithelial–stromal tumours of prostatic origin that have a histological appearance and spectrum similar to the much more common phyllodes tumours of the breast (Gaudin *et al*., 1998) and the very rare, similarly named tumour of the seminal vesicle (see **Figure 23**). At the benign end of the spectrum, glands lined by bland epithelium are embedded in a cellular stroma, often with abundant stromal mucopolysaccharide. Cellularity, cytologic atypia and mitotic rate remain low. Malignant epithelial–stromal tumours of the prostate are characterized by increased stromal cellularity, stromal atypia and mitotic figures. Overgrowth of atypical stroma, to the exclusion of glands, remains an important malignant criterion. Between these two ends of the spectrum, there are tumours having intermediate features that are placed in a category of uncertain malignant potential. Benign tumours may recur locally, but do not metastasize, and this seems also to be the case with tumours thus far categorized as 'uncertain malignant potential,' although recurrence is more common than with unequivocally benign lesions. Malignant tumours are reported to metastasize.

Epidemiology and Aetiology

Apart from carcinomas of skin, prostate cancer is the most common malignant tumour of men. There is an estimated overall prevalence of 30% in men over 50 years of age, but the prevalence increases with age so that about 80% of men over 80 years old have prostate cancer based on autopsy studies. The apparent incidence of prostate cancer in the USA underwent a twofold increase from 1976 to 1994, but this observation may largely be attributable to better detection methods, including measurement of serum PSA levels and imaging the prostate with transrectal ultrasound (TRUS) (see chapter on *Ultrasound*). This viewpoint is supported by a decline in prostate cancer mortality rates by 0.5% per year between 1990 and 1994 that followed an increasing mortality rate from 1976 to 90. Nonetheless, prostate cancer currently ranks second only to lung cancer as a cause of cancer mortality in men in the USA. It therefore represents a major public health problem.

Epidemiological studies (see chapter on *Identifying Cancer Causes through Epidemiology*) demonstrate wide variations in the clinical incidence and mortality rates from prostate cancer in different countries. Low rates are seen in the Far East and many third-world countries, whereas there is a high incidence in Northern Europe, the USA, and Canada. Since migrated populations tend to acquire the prostate cancer rate of the new country of residence, environmental factors appear to have a major impact on prostate cancer incidence. Despite marked differences in the incidence of clinical carcinoma of the prostate in different geographic locations, the frequency of detectable prostate cancer at autopsy in these different areas is similar. These observations suggest that in countries with a high rate of clinical carcinoma, there is an increased rate of conversion of 'latent' carcinoma to clinical carcinoma.

Ethnic factors also appear important in determining prostate cancer risk. In the USA, there is a significantly higher mortality rate among African Americans than in Caucasians, and Jewish populations have a lower mortality rate compared with non-Jewish men. It is often not clear if genetic factors account for such differences or if differences in cultural habits are largely responsible. Some have suggested that the higher incidence of clinical prostate cancer in African Americans can be at least partially attributed to their higher levels of serum testosterone compared with other ethnic groups. It is clear that androgen exposure is necessary for the development of prostate cancer since prepubertal castrates do not develop prostatic carcinoma.

Among the environmental factors that have been linked to prostatic carcinoma, high fat intake has one of the more consistent associations. Intake of saturated fat, rather than monounsaturated or polyunsaturated forms, largely correlates with this increased risk and may account for at least some of the difference in prostate cancer incidence between Western and Far Eastern countries. Other factors that have been linked to an increased risk of prostatic carcinoma include vitamin D deficiency (partially based on the higher incidence in northern latitudes and among darkly pigmented persons), exposure to cadmium, employment in the textile, rubber, drug, chemical and atomic energy industries, large body mass and high plasma levels of insulin-like growth factor 1. Decreased risk has been associated with high soy protein intake (perhaps related to the phyto-oestrogens that are present in soy), high selenium intake, increased serum levels of lycopene (a carotenoid found in tomatoes) and increased intake of α-tocopherol.

It is now clear that prostate cancer has a genetic component (see chapter on *Inherited Predispositions to Cancer*), but cases with a clear familial component represent a small minority of prostate cancers. In men who have a single first-degree relative (father or brother) with prostate cancer, there is a calculated twofold increased risk, whereas having two or more affected first-degree relatives has been associated with an increased risk of 5–11-fold. Younger age (<50 years) of onset in a first-degree relative also appears to convey an increased risk beyond that seen with an older first-degree relative having had prostate cancer. A putative prostate cancer susceptibility gene has been identified on the long arm of chromosome 1.

Despite these numerous observations concerning factors conveying increased or decreased risk of prostate cancer, the cause or causes of carcinoma of the prostate

remain unknown. It is likely that environmental factors make certain important genetic mutations either more or less likely to occur in prostatic epithelial stem cells. One or more such mutations must be required for malignant transformation. It also seems likely that cases with a familial component represent tumours developing in patients with a germline mutation in one of the important genes for at least one of the pathways to malignant transformation. More specific information, however, is lacking at this time.

Screening and Prevention

The early detection of prostate cancer provides the best hope for its cure. The current screening recommendations are for men 50 years of age or older to have annual DRE and serum PSA measurements. Abnormal findings are typically followed by TRUS and, in some cases, biopsy. For men with a positive family history, it is recommended that screening begin earlier, at age 40 years. Unfortunately, neither PSA measurement nor DRE is entirely specific or sensitive. Benign prostatic hyperplasia and prostatitis can cause PSA elevations and abnormal DRE. The 'normal' PSA value is usually considered $<4\,\mathrm{ng\,mL^{-1}}$, and the positive predictive value of a PSA level exceeding $10\,\mathrm{ng\,mL^{-1}}$, is about 66%. A level of $4{-}10\,\mathrm{ng\,mL^{-1}}$, however, has one of only 22–35%. Furthermore, 22% of men with PSA values of $2.5{-}4.0\,\mathrm{ng\,mL^{-1}}$, generally considered in the 'normal' range, were found to have adenocarcinoma on prostate biopsy in one study (Catalona *et al*., 1997). A positive DRE has a positive predictive value of 21–39%.

In an effort to improve the specificity and sensitivity of PSA measurements, additional PSA measurements have been employed. Thus, the PSA density (total serum PSA per cubic centimetre of prostatic volume as determined by TRUS), PSA velocity (change in PSA value over time) and percentage of free (non-protein-bound) PSA improve the sensitivity and specificity for the detection of carcinoma and permit the avoidance of biopsy in some cases.

There have been no controlled trials to date that have documented a successful method of prostate cancer prevention. Potentially successful means might include drugs to decrease serum androgens or inhibit their tissue effects, a reduction in saturated fats in the diet and dietary supplements to provide increased amounts of phyto-oestrogens, vitamin E, selenium and/or lycopene. In one ongoing study looking at prostate cancer prevention with the 5-α-reductase inhibitor finasteride (which prevents the conversion of testosterone to its active metabolite, dihydrotestosterone), preliminary results showed no decrease in high-grade PIN among those receiving the drug, casting some doubt on its ultimate efficacy as a preventive measure.

Molecular Genetic Findings

A heterogeneous mixture of genetic alterations characterizes prostate cancers. Some of the most common abnormalities include loss of heterozygosity corresponding to loci on 7q, 8p, 10q, 13q, 16q and 18q. Loss of genetic material at these locations may well correspond to deletions of tumour suppressor genes that normally restrict cellular growth.

The association of prostate cancer with vitamin D deficiency has prompted the investigation of the role of genetic variation of the vitamin D receptor gene in prostate cancer. In one study of Japanese men with and without prostate cancer, one of three polymorphisms of the vitamin D receptor gene was associated with one-third the risk of prostatic cancer (Habuchi *et al*., 2000).

Because of the relationship of prostate cancer to androgen, studies of a gene (*CYP3A4*) that is normally involved in the deactivation of testosterone have been conducted and shown interesting results. A CYP3A4 variant has been linked to a higher frequency of prostate cancer and, more recently, this variant was found in a much higher proportion of African Americans, a group known to be at relatively high risk for prostatic carcinoma. Furthermore, African Americans with prostate cancer were homozygous for the variant gene significantly more frequently (46%) compared with a control group (28%) who did not have prostate cancer (Paris *et al*., 1999). Along similar lines, a missense mutation in the 5-α-reductase gene that is responsible for coding the protein that converts testosterone to its active metabolite was found to be associated with a much higher enzyme activity and was identified seven times more frequently in African Americans with prostate cancer compared with a healthy control population of African Americans (Makridakis *et al*., 1999).

A novel DNA transcript, designated *DD3* and mapping to chromosome 9q, was recently shown to be overexpressed in 53 of 56 prostate tumours, but it was not identified in other tumours or non-neoplastic tissues other than prostate (Bussemakers *et al*., 1999). Another gene that appears important in prostate cancer progression is the androgen receptor gene located on Xq-12, which is amplified in one-third of hormone refractory prostate cancers (Nupponen and Visakorpi, 1999).

Prognostic Factors

A number of features impact on the prognosis of prostate cancer. Initial stage remains of paramount importance. Clinical stage correlates with overall survival (**Figure 16**) and pathological stage, as determined from the examination of the radical prostatectomy specimen, provides a stronger assessment of the prognosis (**Figure 17**). Patients with incidentally identified

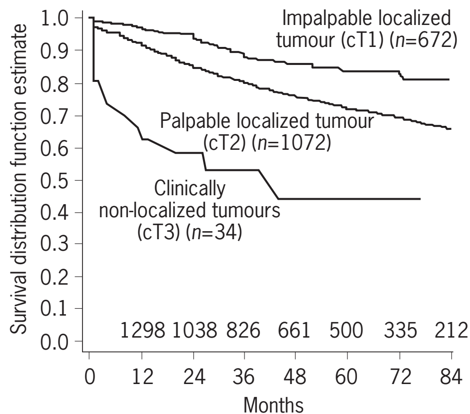


Figure 16 There is a decline in the survival of patients with prostate cancer treated by prostatectomy with advancing tumour stage as determined by clinical assessment. (Adapted from Catalona and Smith, 1998.)

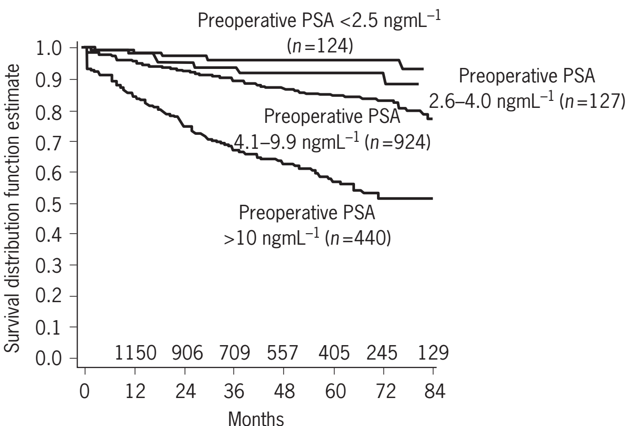


Figure 18 There is a decline in the survival of patients with prostate cancer treated by prostatectomy with increasing levels of serum PSA determined preoperatively. (Adapted from Catalona and Smith, 1998.)

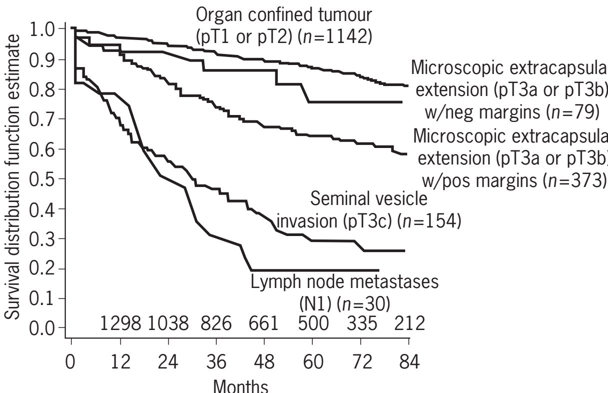


Figure 17 There is a decline in the survival of patients with prostate cancer treated by prostatectomy with advancing tumour stage as determined by pathological examination of the prostatectomy specimen and regional lymph nodes. (Adapted from Catalona and Smith, 1998.)

tumours in transurethral resectates (stage T1a) have a 95% 10-year survival, but there is a progressive decline in 10-year cancer-specific survival with advancing tumour stage; patients with T1b–T2 tumours have about an 80% 10-year cancer-specific survival, and those with T3–T4 tumours have a 60% 10-year cancer-specific survival. Patients who have nodal involvement at operation are rarely cured, but may have a prolonged survival, with a ten-year cancer-specific survival of about 40%.

In addition to stage, significant prognostic factors include pretreatment PSA levels (**Figure 18**), Gleason score (**Figure 8**) and, in patients treated by prostatectomy, tumour volume, margin status, the presence of perineural invasion and tumour ploidy in the prostatectomy specimen. These factors are not necessarily independent; some investigators, for instance, maintain that tumour volume provides no further prognostic information once Gleason score and margin status are determined.

Overview of Present Clinical Management

There are several treatment options for patients with carcinoma of the prostate. These include radical prostatectomy, radiation treatment by either external beam or radioactive seed implants (brachytherapy), anti-androgen treatment (orchiectomy or drug-induced androgen blockade) and so-called 'watchful waiting.' Patients with localized disease (T1–T2) are the best candidates for radical prostatectomy. Advanced prostate cancers (T3 and T4) are usually treated locally by radiation, whereas the primary approach for patients with metastatic disease is with androgen deprivation. Because many prostate cancers have an indolent course and do not affect the longevity of the patient, some have advocated no specific treatment for those with localized tumours. These patients receive periodic follow-up and may be treated if there is evidence of disease progression beyond a certain point. One problem with this approach is the reliable identification of those patients who will have a slowly progressive course versus those who are likely to experience rapid progression. For older patients with low-grade tumours that are predicted to be of small volume (based on TRUS and needle biopsy estimates), this may be a rational approach. T1a tumours (incidentally discovered in transurethral resectates) may be particularly amenable to this approach since they are often low-grade, low-volume lesions. Nonetheless, progression of the T1a tumours to high-stage lesions has been documented in up to 37% of patients who are followed beyond 10 years (Epstein *et al*., 1986). In one Swedish study there was no difference in cancer-specific survival comparing

treated versus untreated patients with localized tumours who had been followed for up to 15 years (**Figure 19**) (Johansson *et al*., 1997).

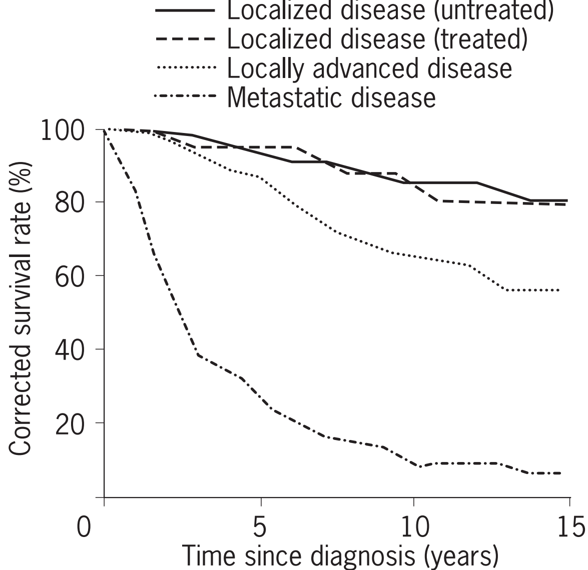


Figure 19 In this study there was no difference in the survival of patients with localized prostate cancer who received no treatment as compared to those who were treated. (Adapted from Johansson *et al*., 1997.)

PENIS

Normal Development and Structure

The penis develops from the genital tubercle, a swelling at the superior aspect of the cloacal folds. Such development is dependent on stimulation by androgenic hormones of testicular origin *in utero*. The superior portion of the cloacal folds is segregated from the inferior and become the urethral folds, which become incorporated into the penis to form the urethra. The genital tubercle, with androgenic stimulation, enlarges and elongates to form the phallus. A plaque of ectoderm at the distal aspect of the phallus forms the glans and distal most portion of the urethra.

The distal penis consists of the glans, the coronal sulcus and foreskin (prepuce) (**Figure 20**). The glans consists of a layer of stratified squamous epithelium, a thin lamina propria and the erectile tissue of the corpus spongiosum that surrounds the distal most portion of the urethra. The body or shaft represents the majority of the penis. The shaft is composed of three cylinders of erectile tissue, the paired corpora cavernosa and the ventralmost corpus spongiosum that contains the urethra. Each of these structures is covered by the tunica albuginea, a dense layer of connective tissue, and is encased in a layer of fascia (Buck's fascia). The most proximal portion of the penis, the root, is located in the perineum. Here, the paired corpora cavernosa form the penile crura, which have ligamentous insertions on the ischial bones. The central corpus spongiosum is expanded into the penile bulb, which is penetrated by the membranous urethra.

The arterial supply to the penis is derived from the pudendal branches of the internal iliac arteries. These give

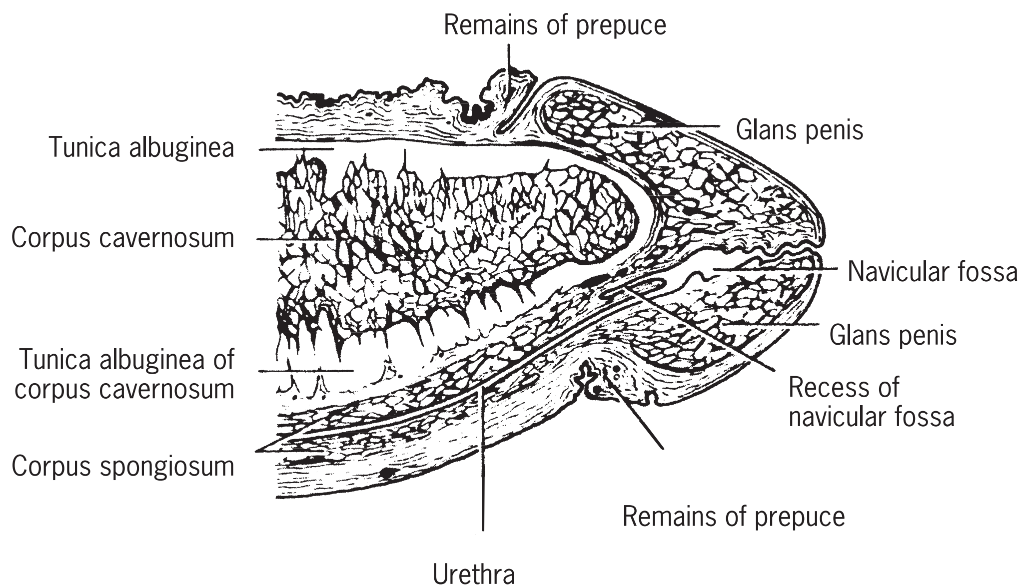


Figure 20 Sagittal section of distal penis, with foreskin removed. (Adapted from Romanes, 1986.)

rise to branches that run on the dorsal aspect, deep aspect, and bulb of the penis. The dorsal vein of the penis splits into two branches and drains into the prostatic plexus. The lymphatics that drain the glans and corpora cavernosa empty into superficial and deep inguinal nodes, whereas the skin of the shaft and foreskin empty into superficial inguinal lymph nodes.

Tumour Pathology

The overwhelming majority of penile cancers are squamous cell carcinomas. Most cases occur in older men (mean age, 58 years), but occasional younger patients may also be affected. The portions of penis that are most commonly affected are the glans (80%), foreskin mucosa (15%) and coronal sulcus (5%), but spread of tumours that originate in one of these locations to other penile components is common (Young *et al*., 2000).

Several subtypes of penile squamous cell carcinoma are described based on both the gross and histological appearances of these tumours. Young *et al*. (2000) have divided them into three major categories based on their grossly recognizable growth patterns: superficial spreading (35%), verruciform (25%) and vertical growth (20%), with the remainder having mixed patterns. The superficial spreading variant usually shows a plaque-like growth of white granular tissue that often spreads over large areas of the glans and foreskin. Verruciform tumours have an exophytic growth of white to grey tumour, typically with well-defined, circumscribed margins on cut surface. The vertical growth tumours are often ulcerated, deeply penetrating lesions with foci of haemorrhage, necrosis and cystic change.

Histologically, most tumours fall into the 'usual-type' squamous cell carcinoma category and are typically well differentiated and rarely poorly differentiated or sarcomatoid in type, with a spindle cell growth pattern. Other variants of squamous cell carcinoma include basaloid carcinoma (10%) (strongly associated with human papilloma virus (HPV; see chapter on *Infectious Agents and Cancer*)), papillary carcinoma (8%), which has a verruciform gross pattern but tends to have a jagged base composed of infiltrating nests of squamous cell carcinoma, warty carcinoma (6%) (**Figure 21**), which usually has a verruciform gross appearance and shows features related to HPV, and verrucous carcinoma (3%), which also has a verruciform appearance and pushing borders but lacks HPV-related features.

Preinvasive lesions of penile squamous carcinoma demonstrate the same spectrum of squamous cell dysplasia as is more commonly seen in the uterine cervix. For the high-grade dysplastic lesions, those affecting the glans may produce a diffusely erythematous clinical appearance, so-called 'erythroplasia of Queyrat.' 'Bowen disease' is the clinical term applied to high-grade squamous dysplasia that affects the penile shaft, producing a red or white elevated lesion. A third lesion, Bowenoid papulosis, typically affects the shaft as small, multiple papules in younger patients (mean age 30 years) than those with erythroplasia of Queyrat or Bowen disease (mean age in the sixth decade). Most commonly, the lesions of Bowenoid papulosis undergo spontaneous regression, whereas those of erythroplasia of Queyrat or Bowen disease persist and occasionally progress to invasive carcinoma.

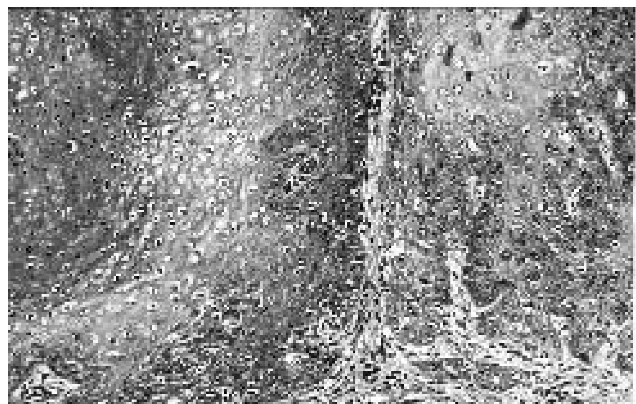


Figure 21 A portion of a 'warty' squamous cell carcinoma shows the nuclear enlargement and perinuclear clearing associated with the human papillomavirus (left) and an invasive growth of malignant-appearing squamous cells (right).

In addition to these variants of squamous cell carcinoma, some much more rare primary malignant tumours of the penis may also be seen. These include adenosquamous carcinoma, basal cell carcinoma, malignant melanoma, Paget disease, and various sarcomas. Of the sarcomas, those showing vascular differentiation are most common and include angiosarcoma, epithelioid hemangioendothelioma and Kaposi's sarcoma, the latter often occurring in patients with the acquired immunodeficiency syndrome. Other sarcomas include leiomyosarcoma, epithelioid sarcoma, malignant fibrous histiocytoma and embryonal rhabdomyosarcoma. Lymphoma may also rarely occur. Occasionally, metastatic tumours involve the penis, usually in patients with known advanced stage disease. The urinary bladder, prostate and kidney are the most common sources.

Epidemiology and Aetiology

The incidence of squamous cell carcinoma of the penis varies markedly around the world. It is common in Africa, Latin America and much of Asia, but rare in Europe and North America. Poor hygiene and phimosis are strongly associated with its development. Other features that are significantly associated with penile cancer include a history of HPV-related lesions, such as condylomata, and the absence of circumcision. It is likely that some forms are, like carcinoma of the uterine cervix, related to HPV infection given that HPV types can be identified in some

tumours and there is an increased frequency of penile cancer in the sexual partners of women with cervical dysplasia and carcinoma. It seems likely that absence of circumcision may permit carcinogenic substances or agents to accumulate in smegma to initiate tumour development. This hypothesis is also consistent with the anatomical distribution of penile cancer, with a marked predominance on the glans, mucosa of foreskin and coronal sulcus. Perhaps, in addition, the development of a thickened keratin layer that follows circumcision is a protective factor.

Screening and Prevention

Because of the rarity of penile carcinoma in Europe and North America, no experience with screening programmes is available. It is likely that simple inspection of high-risk patients would suffice to identify early lesions at a curable stage. In theory, precursor lesions could likely be detected by cytological preparations similar to those employed for the uterine cervix.

Circumcision in the neonatal period is considered a simple and effective method for the prevention of penile cancer. Schoen (1997) reported that the incidence of penile cancer among circumcised men in the USA is essentially zero, whereas it is 2.2/100 000 among uncircumcised men. Among a historical series of more than 1600 penile cancers that occurred in the last 60 years in the USA, none of the patients had been circumcised in infancy (Schoen, 1991).

Molecular Genetic Findings

Because of the rarity of penile cancer in Europe and North America, relatively little information is available concerning the molecular genetic events in these tumours. In one study of 64 penile cancers, HPV DNA was found in 36 cases (56%) (most commonly HPV 16), and 26% had evidence of *p53* mutations (Levi *et al*., 1998). An additional study identified c-Ha-*RAS* mutations in a metastasis of a penile squamous cell carcinoma that contained HPV 18 DNA. This mutation was not identified in the primary tumour or in an earlier metastasis, suggesting it was a late event involved in tumour progression (Leis *et al*., 1998).

Prognostic Factors

A number of features of penile carcinoma have prognostic significance regarding the occurrence of metastasis. Tumours with a verruciform growth pattern (Cubilla *et al*., 2000) have a better prognosis than those with a superficial spreading pattern which are, in turn, associated with a better outcome that those with vertical growth. Among the histological types, verrucous carcinoma has the best prognosis, warty and papillary squamous carcinomas have a good prognosis, but somewhat worse than the pure verrucous carcinomas, squamous cell carcinomas of the usual type have an intermediate prognosis and the most aggressive tumours are the basaloid (Cubilla *et al*., 1998) and sarcomatoid variants. A further refinement in outcome can be obtained by defining the degree of tumour differentiation, although this seems to have been applied mostly for the 'usual' type of squamous cell carcinoma rather than the less common variants (Horenblas and van Tinteren, 1994). Maiche *et al*. (1991) have further refined the grading methodology by developing a scoring system that factors the degree of keratinization, mitotic rate, cytological atypia and inflammatory response. Young *et al*. (2000) found a good correlation between a prognostic index (based on the tumour grade and anatomical level of invasion) and subsequent metastasis and tumour-related death. This is in line with the work of McDougal (1995) regarding the importance of the depth of tumour invasion and metastasis.

Overview of Present Clinical Management

The mainstay of treatment for penile carcinoma is surgical excision with lymph node dissection. For tumours of the foreskin, circumcision without penectomy may be adequate if frozen section evaluation of the margins is negative. For cases involving the glans, partial penectomy is required. In some cases with extensive areas of involvement, total penectomy may be necessary. In some cases, sentinel lymph node sampling, if negative, may permit the avoidance of more extensive inguinal lymph node dissection, but occasional cases of 'skip' metastases do occur.

SEMINAL VESICLE

Normal Development and Structure

The seminal vesicles develop from outpouchings of the mesonephric ducts, which also form the vasa deferentia. They are normally located along the posterolateral aspect of the urinary bladder and superior to the prostate, although they are occasionally embedded within the prostatic capsule. The excretory ducts of the seminal vesicles empty into the vasa deferentia, with the resultant conjoined structure known as the ejaculatory duct. Each seminal vesicle consists of numerous branching ducts embedded in a smooth muscle stroma.

Tumour Pathology

Primary cancers of the seminal vesicle are extraordinarily rare and hence have no known epidemiological features. Most have occurred in older men, with a mean age of 62 years, but occurrence in the third decade is also reported. Patients typically present with symptoms of urinary

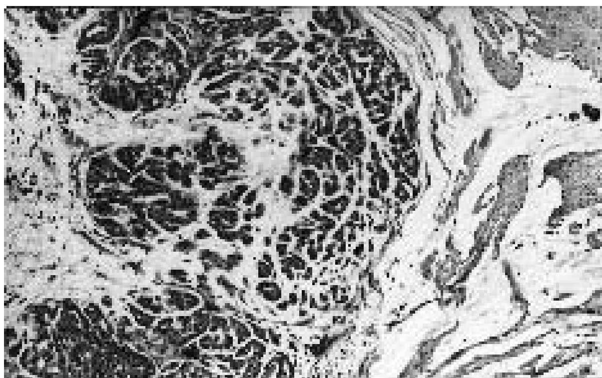


Figure 22 A mucinous carcinoma of the seminal vesicle has nests and cords of malignant epithelial cells in pools of extracellular mucinous secretion.

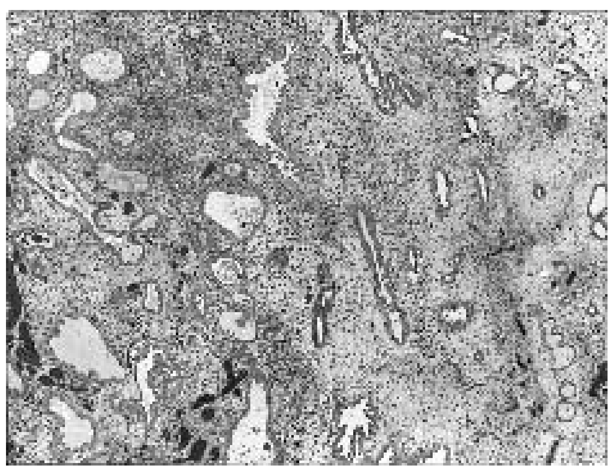


Figure 23 Epithelial–stromal tumour of the seminal vesicle. A cellular, neoplastic stroma surrounds numerous glands, some having a tufted epithelial lining. [From Young *et al.*, 2000, *Tumours of the Prostate, Seminal Vesicles, Urethra, and Penis*. 354 (Armed Forces Institute of Pathology, Washington, DC), by permission.]

obstruction, although haematuria and haemospermia also may occur. A palpable mass is present on digital rectal examination.

The tumours are usually multinodular, solid masses of grey–tan tissue. Microscopically, they are typically poorly differentiated adenocarcinomas. Papillary patterns are common, and clear cell carcinoma, similar to the tumour seen in the female genital tract, and mucinous carcinoma may also occur (**Figure 22**).

On immunohistochemical study, the primary carcinomas of the seminal vesicle are negative for PSA and PAP, assisting with the differential diagnosis of secondary involvement of the seminal vesicle by prostatic carcinoma, a much more common occurrence. In addition, seminal vesicle carcinomas may stain for carcinoembryonic antigen, CA-125 and cytokeratin 7, but are negative for cytokeratin 20 (Ormsby *et al.*, 2000).

In addition to the very rare primary carcinomas of the seminal vesicle, even more rare epithelial–stromal tumours also occur that are similar in appearance to those of prostatic origin. They typically occur in middle-aged to older patients who present with urinary retention. On gross examination, they are multicystic and may have solid foci. On microscopic examination, glandular structures, sometimes cystically dilated, are lined by cytologically benign cuboidal to columnar epithelium and surrounded by a variably cellular and atypical stroma that may show mitotic activity (**Figure 23**). Classification of the neoplasm as low or high grade depends on the stromal cellularity, atypia, mitotic activity and extent of stromal overgrowth. Immunostains for PSA and PAP are negative in the epithelial component, unlike the epithelial–stromal tumours of prostatic origin.

Rare sarcomas of the seminal vesicle origin are reported and include leiomyosarcoma, fibrosarcoma, malignant fibrous histiocytoma, liposarcoma and angiosarcoma.

Secondary spread of malignant tumours to the seminal vesicle is common. Overwhelmingly, these represent prostatic carcinoma or, less commonly, urothelial carcinoma from the bladder. Others include colorectal carcinoma, germ cell tumour and carcinoid. There is a significant decline in prognosis with secondary spread of prostatic carcinoma to the seminal vesicle as compared with otherwise similar tumours lacking such spread (**Figure 17**).

Prognostic Factors

Seminal vesicle carcinoma has a poor prognosis. Most of the reported patients have metastatic disease at the time of diagnosis; 95% of the patients with follow-up have survived less than 3 years. Of the few patients with survival of more than 1.5 years, all were treated with surgical excision (often requiring cystoprostatectomy) and anti-androgenic therapy, either with oestrogen, orchiectomy, or androgen inhibitors.

TESTIS

Normal Development and Structure

The testes develop initially as undifferentiated gonads represented by bilateral thickenings of mesenchyme between the mesenteric root and the mesonephros. The coelomic epithelium overlying these genital ridges proliferates and subsequently grows into the mesenchyme to form the primitive sex cords (**Figure 24**). The germ cells develop in the yolk sac but migrate along the midline to take residence in the gonads. The sex cord cells form the Sertoli cell component of the seminiferous tubules, which are also populated by the migrated germ cells. The Leydig cells differentiate from the primitive mesenchyme of the

interstitium. Gradually the testis assumes its adult shape as a distinct ovoid structure and migrates caudally, eventually to descend into the scrotal sacs in association with a small tongue of pelvic peritoneum, the tunica vaginalis. The seminiferous tubules anastomose with mesonephric tubules that form the ductular system of the testis and which empty into the mesonephric duct, the precursor of the vas deferens.

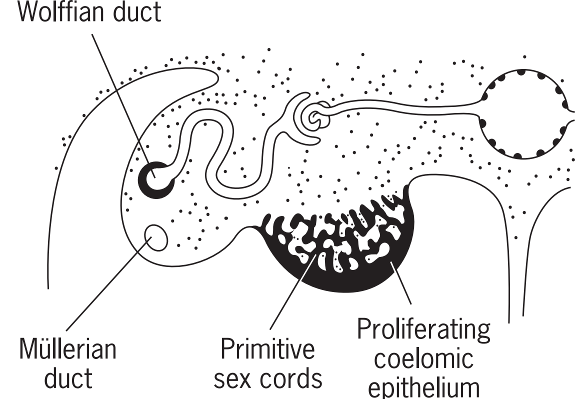


Figure 24 The primitive sex cords of the testis form from the proliferation of the coelomic epithelium that overlies the embryonic gonadal ridge. (Adapted from Langman, 1969.)

The ovoid testis is composed of numerous seminiferous tubules that anastomose with those of the rete testis and that, in turn, empty into the efferent ductules, which form the head of the epididymis, a structure closely applied to the external testicular surface (**Figure 25**). The testis is surrounded by a thick fibrous coating, the tunica albuginea, with a layer of mesothelium on its external aspect derived from the visceral layer of the tunica vaginalis.

The arterial supply to the testis and epididymis is from the testicular artery, usually a branch of the aorta, and from the artery of the vas deferens that derives from the superior vesical artery. The venous drainage exits the testis as a group of four to eight small veins at the hilum that invest the testicular artery in the spermatic cord as the pampiniform plexus. The lymphatics empty into the retroperitoneal lymph nodes.

Tumour Pathology

Although testicular germ cell tumours can be divided into two main categories that are often treated differently – seminomas and non-seminomatous germ cell tumours – both tumour types are derived from a common precursor malignant germ cell. This cell resembles those of

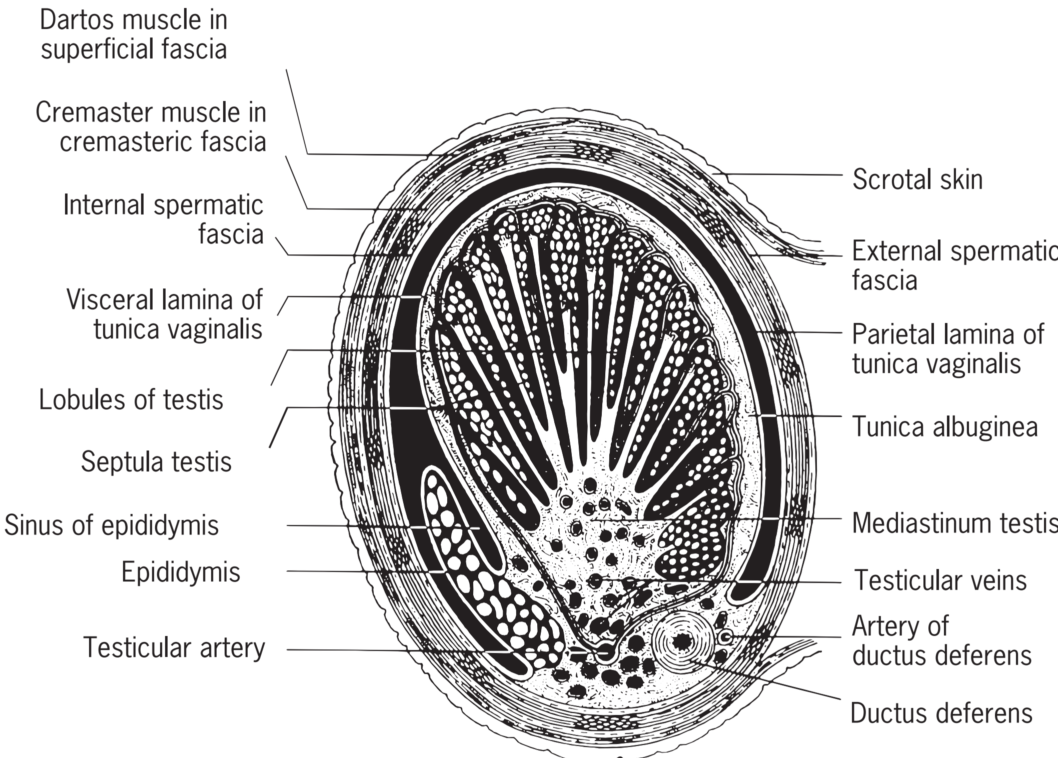


Figure 25 Cross-section of the testis and surrounding scrotum. (Adapted from Romanes, 1981.)

seminoma but also has the capacity to form non-seminomatous germ cell tumours. Such cells characteristically occur in the basal portion of seminiferous tubules that show decreased or absent spermatogenesis and have been termed 'carcinoma-*in-situ*' (**Figure 26**). However, since they can give rise to tumours that are not strictly 'carcinomas,' the term that a committee of pathologists recommended for this lesion is 'intratubular germ cell neoplasia, unclassified type' (IGCNU). Follow-up studies of patients who have had IGCNU identified on testicular biopsy verify a high rate of progression to an invasive germ cell tumour of either seminomatous or non-seminomatous type. The estimated rate of progression is 50% at 5 years, but some experts feel that all cases of IGCNU will eventually progress to an invasive tumour if sufficient time is allowed. Occasional cases of invasive testicular germ cell tumour have been reported in patients 15–20 years after a biopsy showing IGCNU.

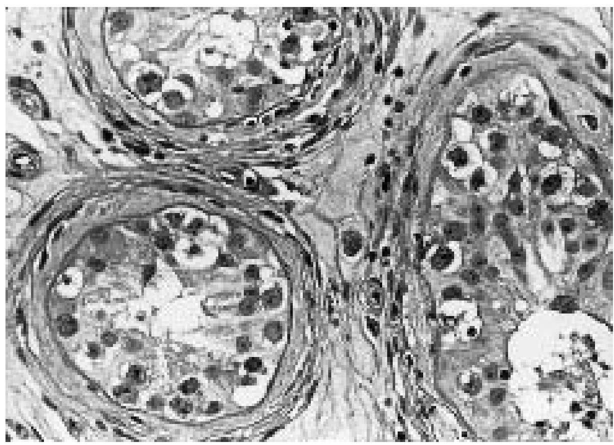


Figure 26 Intratubular germ cell neoplasia, unclassified type (also termed 'carcinoma-*in-situ*'). Note the basal proliferation of seminoma-like cells with clear cytoplasm and enlarged nuclei with prominent nucleoli in seminiferous tubules that lack spermatogenesis.

The evidence that IGCNU gives rise to testicular germ cell tumours is strong. Apart from the follow-up studies already mentioned, IGCNU is identified in the great majority of adult testes harbouring germ cell tumours (but, interestingly, not in prepubertal boys with testicular germ cell tumours). It also has the same immunohistochemical, ultrastructural and cytogenetic findings as are seen in seminomas, and it occurs with high frequency in patients who are known to be at increased risk for testicular germ cell tumours (see below).

Seminoma is the most common form of testicular germ cell tumour and occurs at an average age of 40 years. Most patients present with a testicular mass, but occasional patients present with metastatic tumour, most commonly manifest as vague abdominal or back pain because of retroperitoneal spread. A small group of patients actually have smaller than normal testes, reflecting pre-existing atrophy.

On gross examination, seminoma tends to have a fairly uniform, solid, white/grey to tan cut surface and a lobular outline (**Figure 27**). Some tumours, however, may diffusely replace the testis or show foci of haemorrhage or necrosis that produce a more variegated appearance.

On microscopic examination, the tumour cells are typically arranged in diffuse sheets in an overall lobulated configuration. The sheets are usually subdivided by fibrous septa that contain a lymphocytic infiltrate with occasional plasma cells (**Figure 28**). Some tumours may have a cord-like pattern of growth, especially at the tumour periphery.

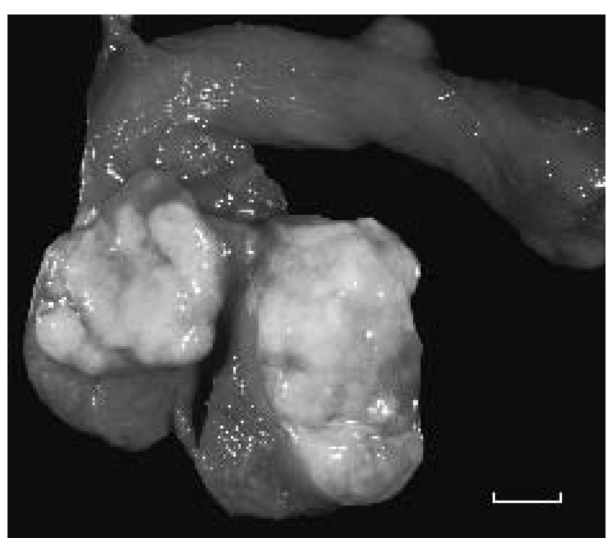


Figure 27 Cut surface of a testis with seminoma showing a solid, lobulated tumour with the typical, homogeneous, cream-coloured to light tan appearance. Bar, 1 cm.

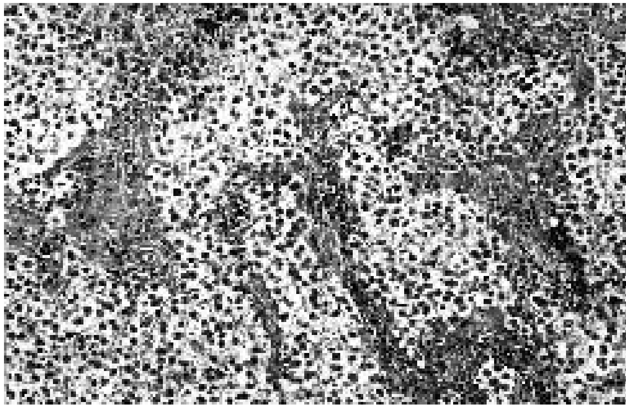


Figure 28 Typical seminoma, showing a sheet-like arrangement of cells with clear cytoplasm that is interrupted by fibrovascular septa with a lymphocytic infiltrate.

Occasional seminomas have a prominent interstitial growth pattern.

The tumour cells have round, often vesicular nuclei, with occasional flattened edges and one or more prominent nucleoli. They have polygonal shapes and generally abundant clear cytoplasm with well-defined cell membranes, although sometimes the cytoplasm is more dense and the nuclei more pleomorphic. The mitotic rate is usually brisk.

Small collections of epithelioid histiocytes occur in about 50% of seminomas, and rarely a similar granulomatous reaction may efface most of the tumour, making careful search for residual tumour cells necessary. Syncytiotrophoblast cells, usually multinucleated with eosinophilic cytoplasm, can be identified in about 10% of seminomas by light microscopy, but are more common if an immunohistochemical stain for human chorionic gonadotropin (hCG) is employed. They are often associated with small foci of haemorrhage, and their presence correlates with elevated levels of serum hCG.

Seminomas, like many other forms of testicular germ cell tumour, are cellular and friable, making artifactual tumour implants common on tissue surfaces and in vessels. Since these may lead to a false-positive diagnosis of extratesticular extension or lymphatic involvement, it is important to distinguish this phenomenon from legitimate examples of these processes. Artifactual implants on tissue surfaces occur at the periphery of tissue sections and in isolated tissue spaces that are discontinuous from clearly invasive tumour. The cells tend to have a poorly cohesive ‘floating’ arrangement and incite no tissue reaction. This is also true for artifactual vascular invasion, whereas true vascular invasion shows more cohesive cellular nests that conform to the shape of the vessel and may be attached to its wall and have an associated thrombus.

Spermatocytic seminoma is an entirely different entity from classic seminoma. It usually occurs in older men, with a mean age in the sixth decade, who present with testicular masses. Although it rarely ‘dedifferentiates’ into a sarcoma that behaves aggressively, the usual spermatocytic seminoma virtually never metastasizes; hence radical orchiectomy alone is adequate treatment. On gross examination, the tumours are often large, grey to haemorrhagic, multinodular, solid and cystic, and have a myxoid quality. Microscopically, the tumour cells are most typically arranged in sheets that are interrupted by oedema and have a paucity of lymphocytes and granulomas. The tumour cells vary in size from small, round, lymphocyte-like cells to intermediate-sized cells to giant cells that may be multinucleated. Despite the almost invariably benign biology of these cases, numerous mitotic figures, including abnormal forms, may be seen.

There are several types of non-seminomatous germ cell tumour; a detailed description of these tumours, however, is beyond the scope of this chapter and can be obtained from other sources (Ulbright *et al*., 1999). In general, the gross appearances of the non-seminomatous tumours are more heterogeneous than those of seminomas (**Figure 29**), reflecting the more common occurrence of haemorrhage, necrosis and cystic degeneration in the former.

Embryonal carcinoma, like the other forms of non-seminomatous germ cell tumour, tends to occur in patients who average 30 years of age. Although, like patients with seminoma, most present with testicular masses, a higher proportion have symptoms secondary to metastatic spread at presentation, reflecting the greater degree of aggressiveness of embryonal carcinoma. The tumour cells may be arranged in solid, glandular, and/or papillary patterns (**Figure 30**). They have large, pleomorphic, often irregularly shaped nuclei that appear crowded, with adjacent nuclei seeming to touch or even overlap. The

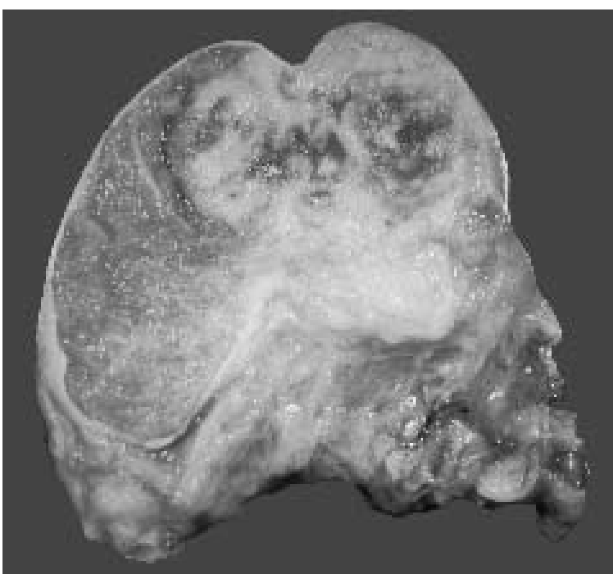


Figure 29 An embryonal carcinoma with a variegated, haemorrhagic appearance.

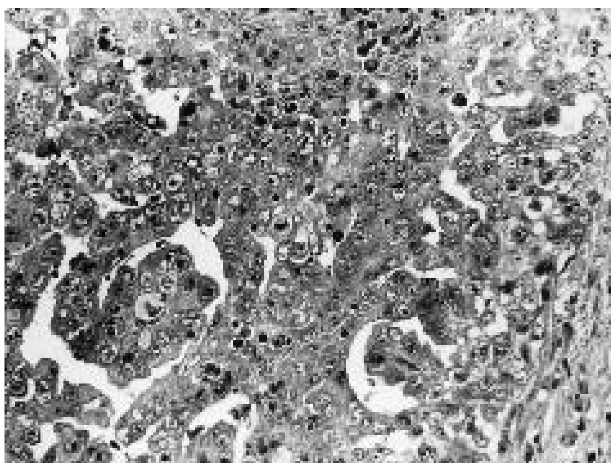


Figure 30 Papillary pattern of embryonal carcinoma. Note the crowded nuclei and indistinct cell borders.

cytoplasm is usually denser that that of seminoma cells and the cytoplasmic borders are ill-defined. The mitotic rate is brisk, and necrosis is frequent and often extensive. Intratubular embryonal carcinoma is often seen in the adjacent parenchyma and may be confused with intravascular tumour, although the more uniform size and shape of the rounded intratubular tumour and frequent comedo-type necrosis contrast with the features of intravascular tumour.

Yolk sac tumour is an uncommon form of pure non-seminomatous germ cell tumour in adult patients, but it represents about 70% of testicular germ cell tumours in children in whom it presents at an average age of 1.5 years. It is a common component of mixed germ cell tumours in postpubertal patients. Yolk sac tumour is the form of germ cell tumour that most consistently produces α-foetoprotein (AFP), which therefore serves as an extremely useful tumour marker. There are numerous yolk sac tumour patterns, and most tumours show several of them. The most common one is a microcystic arrangement wherein vacuolated tumour cells interconnect (**Figure 31**). Solid, myxoid, glandular, papillary, endodermal sinus-like (resembling structures normally identified in the placenta of rodents), macrocystic, hepatoid, sarcomatoid, polyvesicular vitelline, and parietal (basement membrane-rich) patterns also occur. In general, the neoplastic cells are less pleomorphic and atypical than those seen in embryonal carcinoma. The presence of hyaline, eosinophilic globules (**Figure 31**) and bands of basement membrane are helpful ancillary features for the recognition of yolk sac tumour.

Choriocarcinoma, as a pure tumour, is extremely rare and is even unusual as a component of mixed germ cell tumours. Unlike other forms of germ cell tumour, patients with choriocarcinoma usually present with metastatic tumour rather than a palpable mass. Even with known metastatic choriocarcinoma, careful clinical examination of the testes may fail to identify any testicular abnormality. This is because choriocarcinoma is the most common tumour to undergo regression in the testis, so that only foci of scarring with haemosiderin deposits may be present as evidence of pre-existing tumour. Choriocarcinoma is consistently associated with very high levels of hCG, and this can lead to gynaecomastia and thyrotoxicosis, which may therefore be presenting features. On gross examination, a haemorrhagic nodule is typical. Microscopically, syncytiotrophoblast cells and cytotrophoblast cells are intermixed, almost always with associated haemorrhage. Angioinvasion is frequent.

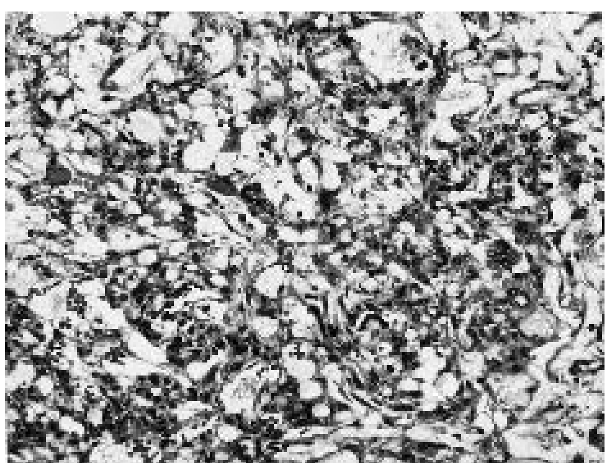


Figure 31 Typical microcystic pattern of yolk sac tumour. There are numerous hyaline globules.

Teratoma is uncommon as a pure tumour in adult patients, but it is the second type of germ cell tumour (in addition to yolk sac tumour) that occurs in children, representing about 30% of the total in paediatric patients. In children it is benign, but in adults it may be associated with metastases of either teratomatous or non-teratomatous germ cell tumours. The reason for this unexpected occurrence is that in postpubertal patients teratoma develops through a process of differentiation from an invasive malignant germ cell tumour, probably embryonal carcinoma most commonly. In postpubertal patients, therefore, teratoma has an aneuploid DNA content and cytogenetic abnormalities similar to those of other testicular germ cell tumours, whereas in prepubertal patients, it is diploid with a normal karyotype.

On gross examination, the tumours may be either solid or cystic. Microscopically, a large variety of tissues may be seen. Cartilage is common, as are enteric-type glands, squamous nests, and smooth muscle stroma. Immaturity is usually manifest as a cellular, mitotically active stroma or islands of neuroepithelium, but it is prognostically irrelevant, given the derivation of teratoma in postpubertal patients from an invasive malignant germ cell tumour and the benign outcome in prepubertal patients. Occasional overgrowth of primitive neuroepithelium or embryonic-appearing skeletal muscle is justification for a diagnosis of teratoma with a secondary malignant component (primitive neuroectodermal tumour or embryonal rhabdomyosarcoma, respectively).

Mixed germ cell tumours are fairly common and, by definition, represent those neoplasms having more than one of the types of tumours already discussed. Those cases with both seminoma and a non-seminomatous component are considered to be non-seminomatous in type because the latter component is most important in determining its behaviour.

Sex cord–stromal tumours of the testis are uncommon, representing about 5% of testicular tumours. They include Leydig cell tumour, Sertoli cell tumour, fibroma, granulosa cell tumour, mixed sex cord–stromal tumour, and unclassified variants. Most patients present with testicular masses, but some develop hormonal symptoms, which may be the presenting feature. This is particularly true for children who may develop pseudoprecocity or gynaecomastia. These tumours occur over a wide age range.

Grossly, sex cord–stromal tumours are usually solid, although cystic degeneration may be present. The colour

varies from grey/white to tan to yellow. Microscopically, a variety of patterns may occur. Leydig cell tumours most commonly have a diffuse arrangement of cells with abundant eosinophilic cytoplasm and sometimes have intracytoplasmic crystals (crystals of Reinke). The hallmark of Sertoli cell tumours is tubule formation. The granulosa cell tumours may have the range of patterns as seen in the more common ovarian counterpart, including microfollicular, trabecular, diffuse and gyriform. Fibromatous tumours again resemble the much more common ovarian lesions composed of gonadal stromal cells arranged in short fascicles and storiform patterns. Those showing substantial amounts of two or more well-recognized lines of differentiation are placed in the mixed category, whereas those with poorly formed elements such that much of the tumour can no longer be recognized in a reliable fashion as having a particular line of differentiation are placed in the unclassified category.

Malignant behaviour can occur with tumours in the sex cord–stromal category and generally correlates with a number of features, including large tumour size (>4–5 m), elevated mitotic rates, tumour necrosis, lymphovascular space invasion, significant cytological atypia and extratesticular growth. In general, about 10% of sex cord–stromal tumours are malignant.

A number of other malignant tumours may occur in the testis, but a detailed description of them is beyond the scope of this work, and most are similar to those occurring more commonly at other sites. These include malignant lymphoma, plasmacytoma, granulocytic sarcoma, leukaemic infiltrates, metastatic carcinoma and melanoma, malignant mesothelioma derived from the mesothelium of the tunica vaginalis, ovarian-type epithelial tumours of either 'borderline' or frankly malignant nature, the recently described desmoplastic round cell tumour and a variety of soft tissue sarcomas of paratesticular origin, including embryonal rhabdomyosarcoma in children.

Immunohistochemical stains are helpful in the evaluation of testicular tumours. Stains for placental alkaline phosphatase (PLAP) highlight the cytoplasmic membranes of most seminomas and are frequently positive in the other germ cell tumour types, with the exception of spermatocytic seminoma, although typically more focally. CD30 (BerH2) is a useful marker for embryonal carcinoma, but is virtually absent in other types. Epithelial membrane antigen is typically absent in germ cell tumours, with the exception of choriocarcinoma and teratoma. AFP is positive in yolk sac tumour and occasionally seen in embryonal carcinoma and teratoma, but it is absent in other types of germ cell tumours. hCG is present in syncytiotrophoblast cells, whether in choriocarcinoma or other types of germ cell tumour. Inhibin is a very useful marker for tumours in the sex cord–stromal category, typically staining most Leydig cell tumours diffusely and somewhat more than half of Sertoli cell tumours focally.

Epidemiology and Aetiology

Testicular cancer is overwhelmingly of germ cell origin and occurs predominantly in young men with a mean age of about 30 years (**Figure 32**). It is a disease that is most common in highly industrialized countries and in the Caucasian population. The highest incidence is reported from Denmark, Switzerland and Germany. Testicular cancer has increased in incidence over much of the twentieth century in those countries where it is common (**Figure 32**).

There are several well-recognized risk factors for testicular cancer (**Table 2**). Cryptorchidism remains one of the strongest associations. Some of the more recent studies suggest that patients with cryptorchidism have about a four times increased risk of testicular cancer that is not alleviated by orchiopexy, and that an elevated risk also applies to the non-cryptorchid testis. This risk, however, does not manifest until late adolescence. Swerdlow *et al*. (1997) presented data showing a higher relative risk, although the small number of abdominal testes in that study did not

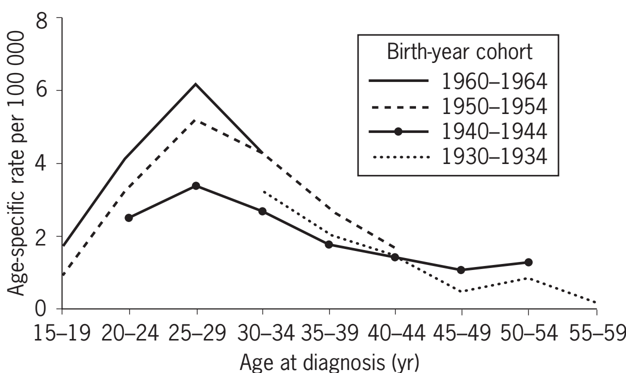


Figure 32 The incidence of testicular germ cell tumours with age in different birth cohorts in Ontario. Note the increasing incidence with later birth year. (Adapted from Weir *et al*., 1999.)

Table 2 Risk factors for testicular germ cell tumours

Intersex syndromes
Prior testicular germ cell tumour
Familial history of testicular cancer
Testicular atrophy
Cryptorchidism
Infertility
Exposure to high levels of ostrogen *in utero*
Caucasian race
Young adult age

Table 3 Estimates of increased risk of cancer in testes associated with various forms of maldescent. (From Swerdlow *et al.*, 1997, *British Medical Journal*, **314**, 1507–1511.)

Position of testes	No. of testes	Observed/expected No. of cancers	Relative risk (95% CI)
Descended opposite maldescended[a]	718	1/0.48	2.1 (0.1–9.2)
Unilaterally maldescended	697	4/0.47	8.5 (2.6–19.8)*
Maldescended opposite maldescended	708	7/0.49	14.4 (6.2–27.8)**
Abdominal maldescended[b]	199	0/0.13	0 (0–28.4)
Non-abdominal maldescended[b]	1206	11/0.84	13.0 (6.8–22.3)**
Maldescended[b]	1405	11/0.97	11.3 (5.9–19.4)**

*$P < 0.01$; **$P < 0.001$.
[a]Includes 21 descended testes for which the contralateral maldescended testis had been excised before the follow-up period.
[b]Regardless of position of other testis.

allow an accurate assessment of their potential for malignant change (**Table 3**).

Patients with certain intersex syndromes have a remarkably high incidence of gonadal germ cell tumours. These syndromes include two main types: gonadal dysgenesis in patients with a Y-chromosome and androgen insensitivity syndrome. Patients with the former often have ambiguous external genitalia and gonads with 'streak' morphology or containing mixtures of ovarian-type stroma and immature-appearing seminiferous tubules. They often develop germ cell tumours in childhood or adolescence; hence, gonadectomy is indicated shortly after the diagnosis is established. Patients with androgen insensitivity syndrome are phenotypic females but genetic males. They have maldescended testes that are at high risk of germ cell tumours after puberty. One study reported that 22% of patients with androgen insensitivity syndrome who were over 30 years of age had developed a germ cell tumour (Morris and Mahesh, 1963).

Patients with a history of testicular germ cell tumour have an increased risk of a second germ cell tumour in the remaining testis, especially if that testis is atrophic. Overall, about 5% of patients develop bilateral neoplasia (Dieckmann and Loy, 1998), but the frequency of contralateral neoplasia increases to about 20% if the remaining testis is atrophic. A similar frequency of contralateral neoplasia is also seen in patients who have a positive family history (see below). At minimum, therefore, continued follow-up of the contralateral testis is indicated for testicular cancer patients.

Males with a family history of testicular cancer are at increased risk (see chapter on *Inherited Predispositions to Cancer*). The risk is greater if the relative is a brother rather than a father. Overall, there is a 2.2% frequency of testicular cancer in the first-degree male relatives of patients with testicular cancer (Tollerud *et al.*, 1985). As mentioned above, a family history also increases the risk of bilateral involvement. A recent study of a large group of testicular cancer families has, through linkage analysis, implicated a testicular cancer gene with a dominant pattern of inheritance that is present on the X-chromosome (Rapley *et al.*, 2000).

Despite these strong associations with testicular cancer, most cases do not occur in patients with well-recognized risk factors. Numerous epidemiological studies have found only weak associations with certain occupations or industrial exposures (see chapter on *Identifying Cancer Causes through Epidemiology*). More recently, a number of studies have implicated *in utero* exposure to high levels of circulating oestrogens as a significant factor for testicular carcinogenesis. Hence, there is increased risk in males born from first pregnancies, in those whose mothers had gestational hyperemesis and in those with a history of neonatal jaundice – all conditions associated with high oestrogen levels during foetal life. In addition, there are numerous factors that likely increase environmental oestrogen exposure, including the increased therapeutic use of oestrogenic substances, the oestrogenic supplementation of livestock and oestrogenic activity of some pesticides.

Molecular Genetic Findings

Most testicular germ cell tumours have a hyperdiploid to hypotriploid DNA content and, on cytogenetic study, they frequently have one or more copies of an isochromosome derived from the short arms of chromosome 12 (i (12p)). This distinctive marker chromosome is considered specific for germ cell tumours and can be used to establish a germ cell tumour diagnosis in problematic cases. Even in those germ cell tumours lacking i (12p), there is an increased amount of DNA derived from 12p, so this region of the genome appears to play a key role in the pathogenesis of germ cell tumours.

In addition to gain of 12p DNA, there are other nonrandom chromosomal changes. These include loss from 11p, 12q, 13, 18 and Y and gain of 7, 8, 21 and X. It has been noted that the DNA content of seminomas is greater

than that of non-seminomatous tumours, and it has been hypothesized that the non-seminomatous tumours derive from seminomas as a consequence of gene loss. As mentioned in the discussion of familial testicular cancer, a putative testis cancer gene has recently been localized to the X-chromosome through linkage analysis (Rapley *et al*., 2000).

Prognostic Factors

The great majority of testicular germ cell tumours are now curable through a combination of surgery, chemotherapy and/or radiation. Prior to the development of effective chemotherapy, many patients with metastatic seminoma could be effectively treated with radiation, but those with metastatic non-seminomatous tumours had a poor prognosis. Now, however, more than 80% of patients with metastatic non-seminomatous tumours are cured.

Stage remains of paramount prognostic importance. Patients with early stage tumours are almost uniformly cured, whereas those with bulky retroperitoneal disease experience about a 20% overall mortality. There are prognostic models to help divide patients into different categories that factor in features such as the levels of serum hCG and AFP, sites of involvement and size of metastatic lesions.

Occasional patients with early stage disease have an unexpectedly aggressive course or fail initial therapy for unclear reasons. It would be of great interest to identify such cases prospectively, but this goal has not yet been achieved.

Several investigators have attempted to define groups of testicular cancer patients who are at high or low risk for metastatic disease after an orchiectomy for clinical stage I tumour (i.e. having no clinical evidence of metastatic tumour after thorough staging, including radiographic evaluation of the retroperitoneum and chest and serum marker studies). Based on either follow-up studies or the results of retroperitoneal lymph node dissections, it is known that about 30% of such patients have occult retroperitoneal metastases. A successful division would permit simple follow-up for those patients at low risk, whereas high-risk patients could be recommended for staging retroperitoneal lymph node dissection (non-seminomatous patients) or adjuvant radiation (seminoma patients). Thorough pathological evaluation of the primary tumour for features such as lymphovascular invasion, extratesticular extension, percentage or volume of embryonal carcinoma and tumour size have shown strong correlations with either pathologically confirmed metastatic tumour or relapse. Alternative methods of evaluation, including flow cytometric studies for tumour S-phase and ploidy, and immunohistochemical staining for determination of proliferative fraction have also been applied with varying results.

Overview of Present Clinical Management

Patients with a clinical diagnosis of a testicular neoplasm receive radical orchiectomy. For those with seminoma, most receive adjuvant radiation to the retroperitoneum, although there is some interest in surveillance management (i.e. close follow-up but no adjuvant therapy) for clinical stage I patients with favourable pathological features. Radiation is usually given for those seminoma patients with small volume, retroperitoneal metastases, but cisplatin-based chemotherapy is used for those with bulkier disease or supradiaphragmatic spread. Patients with non-seminomatous tumours who are clinical stage I may either be placed on surveillance management or receive limited, nerve-sparing retroperitoneal lymph node dissection, again depending in part on pathological features. For those with known metastatic tumour, cisplatin-based chemotherapy is the mainstay followed by surgical excision of residual masses, if necessary. A patient with a teratomatous component in the testicular primary appears to be at higher risk of requiring post-chemotherapy resection because of the increased likelihood of teratomatous metastases and the non-responsiveness of teratoma to chemotherapy.

Patients who have sex cord–stromal tumours and who are clinical stage I may receive either surveillance management or retroperitoneal lymphadenectomy. It is sensible to manage patients whose tumours lack the features associated with malignant behaviour by surveillance and to reserve nodal dissection for those clinical stage I patients who have tumours with one or more of the pathological features that are associated with metastases. For patients with sex cord–stromal tumours that have metastasized, surgical excision by retroperitoneal lymphadenectomy or other procedures provides the best current approach. Chemotherapy and radiation have not proved effective for this situation.

REFERENCES

Albertsen, P. C., *et al*. (1999). Statistical considerations when assessing outcomes following treatment for prostate cancer. *Journal of Urology*, **162**, 439–444.

Amin, M. B., *et al*. (1993). Putative precursor lesions of prostatic adenocarcinoma: fact or fiction? *Modern Pathology*, **6**, 476–483.

Bostwick, D. G. and Brawer, M. K. (1987). Prostatic intra-epithelial neoplasia and early invasion in prostate cancer. *Cancer*, **59**, 788–794.

Brawn, P. N. (1982). Adenosis of the prostate: a dysplastic lesion that can be confused with prostate adenocarcinoma. *Cancer*, **49**, 826–833.

Bussemakers, M. J., *et al*. (1999). DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Research*, **59**, 5975–5979.

Catalona, W. J., *et al*. (1994). Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630. *Journal of Urology*, **151**, 1283–1290.

Catalona, W. J., *et al*. (1997). Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination. Enhancement of specificity with free PSA measurements. *Journal of the American Medical Association*, **277**, 1452–1455.

Cheng, L., *et al*. (1998). Atypical adenomatous hyperplasia of the prostate: a premalignant lesion? *Cancer Research*, **58**, 389–391.

Cubilla, A. L., *et al*. (1998). Basaloid squamous cell carcinoma: a distinctive human papilloma virus-related penile neoplasm: a report of 20 cases. *American Journal of Surgical Pathology*, **22**, 755–761.

Cubilla, A. L., *et al*. (2000). Warty (condylomatous) squamous cell carcinoma of the penis: a report of 11 cases and proposed classification of 'verruciform' penile tumours. *American Journal of Surgical Pathology*, **24**, 505–512.

Dieckmann, K. P. and Loy, V. (1998). The value of the biopsy of the contralateral testis in patients with testicular germ cell cancer: the recent German experience. *APMIS*, **106**, 13–20.

Doll, J. A., *et al*. (1999). Genetic analysis of prostatic atypical adenomatous hyperplasia (adenosis). *American Journal of Pathology*, **155**, 967–971.

Epstein, J. I., *et al*. (1986). Prognosis of untreated stage A1 prostatic carcinoma: a study of 94 cases with extended followup. *Journal of Urology*, **136**, 837–839.

Gaudin, P. B., *et al*. (1998). Sarcomas and related proliferative lesions of specialized prostatic stroma: a clinicopathologic study of 22 cases. *American Journal of Surgical Pathology*, **22**, 148–162.

Gleason, D. F. (1966). Classification of prostatic carcinomas. *Cancer Chemother. Rep., Part 1*, **50**, 125–128.

Gleason, D. F. (1992). Histologic grading of prostate cancer: a perspective. *Human Pathology*, **23**, 273–279.

Gleason, D. F. and Mellinger, G. T. (1974). Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *Journal of Urology*, **111**, 58–64.

Habuchi, T., *et al*. (2000). Association of vitamin D receptor gene polymorphism with prostate cancer and benign prostatic hyperplasia in a Japanese population. *Cancer Research*, **60**, 305–308.

Horenblas, S. and van Tinteren, H. (1994). Squamous cell carcinoma of the penis. IV. Prognostic factors of survival: analysis of tumour, nodes and metastasis classification system. *Journal of Urology*, **151**, 1239–1243.

Johansson, J. E., *et al*. (1997). Fifteen-year survival in prostate cancer. A prospective, population-based study in Sweden. *Journal of the American Medical Association*, **277**, 467–471.

Leis, P. F., *et al*. (1998). A c-rasHa mutation in the metastasis of a human papillomavirus (HPV)-18 positive penile squamous cell carcinoma suggests a cooperative effect between HPV-18 and c-rasHa activation in malignant progression. *Cancer*, **83**, 122–129.

Levi, J. E., *et al*. (1998). Human papillomavirus DNA and p53 status in penile carcinomas. *International Journal of Cancer*, **76**, 779–783.

Maiche, A. G., *et al*. (1991). Histological grading of squamous cell carcinoma of the penis: a new scoring system. *British Journal of Urology*, **67**, 522–526.

Makridakis, N. M., *et al*. (1999). Association of mis-sense substitution in SRD5A2 gene with prostate cancer in African-American and Hispanic men in Los Angeles, USA. *Lancet*, **354**, 975–978.

McDougal, W. S. (1995). Carcinoma of the penis: improved survival by early regional lymphadenectomy based on the histological grade and depth of invasion of the primary lesion. *Journal of Urology*, **154**, 1364–1366.

McNeal, J. E. (1981). The zonal anatomy of the prostate. *Prostate*, **2**, 35–49.

McNeal, J. E. and Bostwick, D. G. (1986). Intraductal dysplasia: a premalignant lesion of the prostate. *Human Pathology*, **17**, 64–71.

McNeal, J. E. and Yemoto, C. E. (1996). Spread of adenocarcinoma within prostatic ducts and acini. Morphologic and clinical correlations. *American Journal of Surgical Pathology*, **20**, 802–814.

McNeal, J. E., *et al*. (1988). Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *American Journal of Surgical Pathology*, **12**, 897–906.

Morris, J. M. and Mahesh, V. B. (1963). Further observations on the syndrome, "testicular feminization". *American Journal of Obstetrics and Gynecology*, **87**, 731–748.

Nupponen, N. and Visakorpi, T. (1999). Molecular biology of progression of prostate cancer. *European Urology*, **35**, 351–354.

Ormsby, A. H., *et al*. (2000). Primary seminal vesicle carcinoma: an immunohistochemical analysis of four cases. *Modern Pathology*, **13**, 46–51.

Pan, C. C., *et al*. (2000). The prognostic significance of tertiary Gleason patterns of higher grade in radical prostatectomy specimens: a proposal to modify the Gleason grading system. *American Journal of Surgical Pathology*, **24**, 563–569.

Paris, P. L., *et al*. (1999). Association between a CYP3A4 genetic variant and clinical presentation in African-American prostate cancer patients. *Cancer Epidemiology Biomarkers Prevention*, **8**, 901–905.

Rapley, E. A., *et al*. (2000). Localization to Xq27 of a susceptibility gene for testicular germ-cell tumours. *Nature Genetics*, **24**, 197–200.

Renshaw, A. A. (1998). Correlation of gross morphologic features with histologic features in radical prostatectomy specimens. *American Journal of Clinical Pathology*, **110**, 38–42.

Schoen, E. J. (1991). The relationship between circumcision and cancer of the penis. *CA – A Cancer Journal for Clinicians*, **41**, 306–309.

Schoen, E. J. (1997). Benefits of newborn circumcision: is Europe ignoring medical evidence? *Archives of Disease in Childhood*, **77**, 258–260.

Srigley, J. R. (1988). Small-acinar patterns in the prostate gland with emphasis on atypical adenomatous hyperplasia and small-acinar carcinoma. *Seminars in Diagnostic Pathology*, **5**, 254–272.

Swerdlow, A. J., *et al*. (1997). Risk of testicular cancer in cohort of boys with cryptorchidism. *British Medical Journal*, **314**, 1507–1511.

Thomas, R., *et al*. (1982). Aid to accurate clinical staging-histopathologic grading in prostatic cancer. *Journal of Urology*, **128**, 726–728.

Tollerud, D. J., *et al*. (1985). Familial testicular cancer and urogenital developmental anomalies. *Cancer*, **55**, 1849–1854.

Ulbright, T. M., *et al*. (1999). *Tumours of the Testis, Adenexa, Spermatic Cord and Scrotum: Atlas of Tumour Pathology*, Third Series. (Armed Forces Institute of Pathology, Washington, DC).

Young, R. H., *et al*. (2000). *Tumours of the Prostate, Seminal Vesicle, Male Urethra and Penis: Atlas of Tumour Pathology*. (Armed Forces Institute of Pathology, Washington, DC).

FURTHER READING

Bostwick, D. G. (1997). Neoplasms of the prostate. In Bostwick, D. G. and Eble, J. N. (eds). *Urologic Surgical Pathology*. 349 (Mosby Year Book, St. Louis).

Bostwick, D. G. (1999). Prostatic intraepithelial neoplasia is a risk factor for cancer. *Seminars in Urology and Oncology*, **17**, 187–198.

Bostwick, D. G., *et al*. (2000). Prognostic factors in prostate cancer. College of American Pathologists Consensus Statement 1999. *Archives of Pathology and Laboratory Medicine*, **124**, 995–1000.

Catalona, W. J. and Smith, D. S. (1998). Cancer recurrence and survival rates after anatomic radical retropubic prostatectomy for prostate cancer: intermediate-term results. *Journal of Urology*, **160**, 2428–2434.

Cubilla, A. L., *et al*. (1998). Basaloid squamous cell carcinoma: a distinctive human papilloma virus-related penile neoplasm: a report of 20 cases. *American Journal of Surgical Pathology*, **22**, 755–761.

Cubilla, A. L., *et al*. (2000). Warty (condylomatous) squamous cell carcinoma of the penis: a report of 11 cases and proposed classification of 'verruciform' penile tumours. *American Journal of Surgical Pathology,* **24**, 505–512.

Kim, I., *et al*. (1985). Leydig cell tumours of the testis. A clinicopathological analysis of 40 cases and review of the literature. *American Journal of Surgical Pathology*, **9**, 177–192.

Langman, J. (1969). *Medical Embryology and Human Development – Normal and Abnormal*, 2nd edn. 164 (Williams and Wilkins, Baltimore).

Lee, F., *et al*. (1989). The role of transrectal ultrasound in the early detection of prostate cancer. *CA – A Cancer Journal for Clinicians,* **39**, 337–360.

McNeal, J. E. (1997). Prostate. In: Sternberg, S. S. (ed.). *Histology for Pathologists*, 2nd ed. 1001 (Lippincott-Raven, Philadelphia).

Randolph, T. L., *et al*. (1997). Histologic variants of adenocarcinoma and other carcinomas of prostate: pathologic criteria and clinical significance. *Modern Pathology*, **10**, 612–629.

Romanes, G. J. (1981). *Cunningham's Textbook of Anatomy*, 12th edn. 555 (Oxford University Press, Oxford).

Romanes, G. J. (1986). *Cunningham's Practical Anatomy, vol. 2: Thorax and Abdomen*, 15th edn. 110 (Oxford University Press, Oxford).

Ruijter, E., *et al*. (1999). Molecular genetics and epidemiology of prostate carcinoma. *Endocrine Reviews*, **20**, 22–45.

Ulbright, T. M. (1993). Germ cell neoplasms of the testis. *American Journal of Surgery and Pathology*, **17**, 1075–1091.

Ulbright, T. M. (1999). Testis risk and prognostic factors. The pathologist's perspective. *Urologic Clinics of North America*, **26**, 611–626.

Ulbright, T. M., *et al*. (1999). *Tumours of the Testis, Adenexa, Spermatic Cord and Scrotum: Atlas of Tumour Pathology*, Third series, Fascicle 25. (Armed Forces Institute of Pathology, Washington, DC).

Young, R. H., *et al*. (1998). Sertoli cell tumours of the testis, not otherwise specified: a clinicopathologic analysis of 60 cases. *American Journal of Surgical Pathology*, **22**, 709–721.

Young, R. H., *et al*. (2000). *Tumours of the Prostate, Seminal Vesicle, Male Urethra and Penis: Atlas of Tumour Pathology*, Third series, Fascicle 28. (Armed Forces Institute of Pathology, Washington, DC).

Weir, H. K., *et al*. (1999). Trends in the incidence of testicular germ cell cancer in Ontario by historic subgroup, 1964–1996. *Canadian Medical Association Journal*, **160**, 201–205

Web Sites with Information on Cancer of the Prostate, Testis and Penis

Prostate

http://www3.cancer.org/cancerinfo/load_cont.asp?ct=36&prevURL=load_cont.asp&language=ENGLISH from http://www.cancer.org/

http://www.cancerlinksusa.com/prostate/index.htm

Testis

http://rex.nci.nih.gov/WTNK_PUBS/testicular/index.htm

http://www.cancerlinksusa.com/testicular/

Penis

http://www3.cancer.org/cancerinfo/load_cont.asp?ct=35

http://www.oncolink.upenn.edu/disease/penile/

Chapter 46

Lymph Nodes

Hans Konrad Müller-Hermelink and German Ott
University of Würzburg, Würzburg, Germany

CONTENTS

- The Anatomical Structure of the Lymph Node
- A Functional Approach to the Recognition of Reactive and Neoplastic Lymph Node Patterns
- Reactive Patterns of the Lymph Node
- Ancillary Methods Necessary for the Definition of Malignant Lymphomas
- Genotypic Studies
- Lymphoma Classification
- Lymphomas of B Cell Lineage
- Peripheral B Cell Lymphomas
- Lymphomas of T Cell Lineage
- Predominantly Leukaemic (Disseminated) Peripheral T and T/NK Cell Neoplasms
- Predominantly Extranodal T/NK Cell Lymphomas
- Hodgkin Lymphoma

THE ANATOMICAL STRUCTURE OF THE LYMPH NODE

Lymph nodes are organs with highly organized compartments, the knowledge of which is a prerequisite for the recognition of reactive and neoplastic lymph node processes.

They are surrounded by a fibrous capsule, where afferent lymphatics penetrate and reach the subcapsular (marginal) sinus. Within the lymph node parenchyma, the cortex contains the B cell follicles, which are embedded in the interfollicular paracortex, organized into the T cell areas and the perisinusoidal and vascular zones. The vascular hilus, where arterial branches reach and efferent lymphatic and venous vessels leave the lymph node, is the localization of the central medulla containing collecting sinuses and the perisinusoidal pulp cords.

The **primary B cell follicle** is composed either of small lymphocytes of mantle zone type, in older patients (or in inactive conditions) or sometimes marginal zone lymphocytes (marginal zone cell nodules). When confronted with antigenic stimulation, germinal centres develop within the primary follicles at the interface of paracortical T cell areas, then called **secondary follicles**. They are surrounded by small lymphocytes of the follicular mantle zone. The germinal centres are usually sharply demarcated, and are composed of a network of follicular dendritic reticulum (FDC) cells, the professional antigen presenting cells of the B cell follicle. The B cells within the germinal centres have the appearance of either large transformed cells, the so-called centroblasts, and smaller to intermediate-sized cleaved cells, the centrocytes. Characteristically, centroblasts and centrocytes are not haphazardly distributed, but display a zonation phenomenon with the appearance of a ‘dark’ zone predominantly composed of centroblasts (the proliferative compartment of the germinal centre), and a ‘light’ zone containing centrocytes (which is normally oriented towards the marginal sinus). Within the dark zone, multiple mitotic figures are found, and there may be a prominent admixture of large histiocytes with pale cytoplasm, thus creating a ‘starry sky’ appearance. These macrophages may be densely packed with pyknotic nuclear debris (‘tingible body macrophages’). Furthermore, a specific type of T cell is contained in germinal centres (the CD4 +, CD57+ T cell).

Reactive germinal centres, as a rule, are surrounded by the **follicular mantle zone**, composed of small lymphocytes with small cytoplasm and round inconspicuous nuclei. In some localizations, predominantly in abdominal (mesenteric) lymph nodes, and rarely also in superficial nodes after antigenic stimulation, another structural component of the follicular zone, termed the so-called **marginal zone** is present. This outermost part of the perifollicular B cell area may occasionally become apparent, but usually cannot be seen in all lymph nodes. Marginal zones, however, can be readily detected and form a characteristic part of the follicular B cell areas in the spleen. By morphology, the marginal zone, as compared with the ‘dark’ lymphocytes of the mantle zone, which are derived from germinal centre precursors or naive B lymphocytes, is composed of slightly larger cells with a broad

pale cytoplasm, which represent different populations of memory B cells.

On immunohistochemistry, all these different components of the follicle are composed of B cells, and therefore stain strongly positive for B cell-associated antigens (see **Table 2**), but may be distinguished by their different phenotypes.

The **nodal paracortex** is composed of T cells areas (or nodules) and the perivascular (perisinusoidal) cortical pulp areas. The paracortical T zone, which may become a nodular structure after chronic (antigenic) stimulation, is composed of small T lymphocytes and intermingled, frequently pale-appearing, interdigitating reticulum cells (the professional antigen presenting cells for T cell activation). In the paracortical pulp, post-capillary high endothelial venules and some transformed (blastic) cells are found which are a mixture of T and B cells or plasma cell precursors. In the paracortex, CD4+ T cells normally are more numerous than CD8+ T cells.

The central portion of the lymph nodes, the **medulla**, contains the collecting sinuses, which are often filled with histiocytes and a special type of endothelial cells, the sinus lining cells. The perisinusoidal pulp cords contain the plasma cells, some T cells and macrophages.

A FUNCTIONAL APPROACH TO THE RECOGNITION OF REACTIVE AND NEOPLASTIC LYMPH NODE PATTERNS

Lymph nodes are filter organs of interstitial fluids and lymph, where circulating molecules, particles, cells and microorganisms are subjected to cognate cells and interaction with the phagocytic and immune systems, and major effector functions of these systems are initiated. In particular, immune stimulation and acceleration, where complex interactions of highly specialized cell types are needed to generate and regulate potentially dangerous effector functions, usually take place in the organized structures of the lymphoid organs (tonsils, Peyer's patches, lymph nodes, spleen). The lymph reaches the lymph nodes via afferent lymph vessels by the marginal sinus which acts like a sieve – there are only few shunts to the collecting ducts, so that everything has to pass through the meshwork of the interstitial reticular tissue and re-enter the collecting sinus. Macrophages take up particles or macromolecular constituents by pinocytosis. Dendritic cells and possibly also sinus lining cells may transport and present antigens to the organized T and B cell areas for further immune reactions.

The blood arteries enter the lymph node at the opposite medullary hilus and branch via lobular vessels to the capillary bed of the cortical lymphoid follicle, where fenestrated capillaries allow plasma proteins including antibodies to leave the circulation and attach to the follicular dendritic cells. Migrating lymphocytes leave the blood stream after specific attachment to the high endothelial vessels penetrating the T and B cell areas and leave the lymph nodes via the collecting sinuses at the hilus through efferent lymph vessels, together with the veins.

Specific antigenic stimulation of naive reactive T and B cells takes place in the highly organized microenvironment of T and B cell areas by professional antigen presenting cells, the interdigitating cells of the T cell area and the follicular dendritic cells of the B cell follicle.

Within the germinal centre of the B cell follicle, antigenic stimulation of naive B cells leads to transformation of these resting B cells into highly proliferating centroblasts that mature into centrocytes. By antigen-dependent selection and somatic mutation of immunoglobulin receptor genes, cells with high receptor affinity become selected, whereas cells with low antigen reactivity die by apoptosis and become phagocytosed by the so-called starry sky macrophages. Selected B cells leave the germinal centre as memory B cells or plasma cell precursors. The T cell influence through different cytokines is crucial for these precursors. Somatic mutation is confined to the germinal centres, and stops in the emigrant cell population.

In the T cell areas, naive T cells become activated and either cooperate with B cells at the interface of T cell areas and B follicles to induce the germinal centre reaction by TH_2 cytokines, or T effector functions by TH_1 mediated macrophages and granulomatous reactions or MHC Class I restricted cytotoxic T cell responses. Effector T cells will leave the lymph node collecting medullary sinuses and form the recirculating lymphocyte pool.

REACTIVE PATTERNS OF THE LYMPH NODE

Table 1 gives an overview on major reactive patterns in lymph nodes. Some of them are described in greater detail, since they represent important differential diagnoses to malignant lymphomas.

Progressive transformation of germinal centres is a relatively rare, but distinct, condition in which in the background of reactive follicular hyperplasia few or many follicles are greatly enlarged. They are characterized by a prominent mantle zone and ill-defined germinal centres, mainly consisting of small lymphocytes. Remnants of germinal centre cells may be found, and are difficult to recognize. In some instances, epithelioid histiocytes are interspersed. This condition sometimes bears a close resemblance to nodular paragranuloma, but no lymphocytic and histiocytic (L&H) cells are present.

Castleman's disease of hyaline vascular type: in contrast to the lesions described above, Castleman disease (CD) represents a distinct clinicopathological entity.

Table 1 Reactive hyperplasias and reactions in different compartments in chronic lymphadenitis

B cell compartment
Follicular hyperplasia
Sinusoidal B cell reaction (immature sinus histiocytosis)
Marginal cell reaction/nodules
Plasmacytosis
Progressive transformation of germinal centres
T cell compartment
Paracortical nodular T zone hyperplasia
Diffuse cortical hyperplasia
Plasmacytoid monocyte reaction
Sinus reactions
Sinus histiocytosis
Foreign body reaction
Mastocytosis
Sinus lymphocytosis
Macrophage reactions
Foreign body reactions
Epithelioid cell reaction
Histiocytic-suppurative granulomas

Usually, this type of CD presents as a solitary lesion, most commonly found in the mediastinum, abdomen or cervical lymph nodes.

The morphological hallmarks of CD are regressively changed small follicles, increased vascularity in the interfollicular area and the presence of plasmacytoid monocytes. In addition, sinuses are lacking. Follicular centres frequently are 'onion skin' shaped with a predominance of follicular dendritic cells and few lymphocytes, and the mantle zone is prominent, sometimes exhibiting a curious 'Indian file' pattern with small lymphocytes being arranged in concentric circles. In the interfollicular areas, postcapillary venules of the high endothelial cell type are numerous, and sometimes one or more of these vessels are seen penetrating into the regressive germinal center. Usually, there is a mixture of lymphoid cells, histiocytes and few plasma cells present, and the so-called plasmacytoid monocytes are diffusely interspersed or may be recognized as small cell nests. They are of medium size and have round to oval nuclei with fine chromatin distribution and pale eccentric cytoplasm, reminiscent of plasma cells. On immunohistochemistry, they stain positive for CD68 and CD43.

In contrast, **multicentric Castleman's disease**, also called multicentric angiofollicular lymphoid hyperplasia, or plasma cell-rich variant of Castleman disease, consistently involves multiple peripheral lymph nodes and may also manifest in the bone marrow, liver, kidney and CNS. The germinal centres may be expanded, or regressively changed as in the hyaline-vascular variant. High endothelial venules are a prominent feature in the interfollicular tissues, as is an enormous proliferation of plasma cells. In most cases, these plasma cells are polyclonal; however, they may be monoclonal in HIV-infected individuals in which the disease is associated with human herpes virus type 8 (HHV-8) infection.

In some patients, the so-called 'POEMS' syndrome may be diagnosed, the acronym summarizing the principal disease features of polyneuropathy, organomegaly, endocrinopathy, M proteins and skin lesions (Rosati and Frizzera, 1997).

ANCILLARY METHODS NECESSARY FOR THE DEFINITION OF MALIGNANT LYMPHOMAS

Immunophenotypic Studies

Immunohistochemistry has gained particular importance both with respect to the differentiation of the various subtypes of malignant lymphomas and in the differentiation of reactive versus neoplastic lymphoproliferative disorders. Immunophenotypic studies have also become accepted as a valuable help in the recognition of lymphoma entities, and have also contributed substantially to our understanding of the histogenesis and pathogenesis of haematopoietic neoplasms in general.

They are not only useful in the primary distinction of a neoplastic versus a benign lymphoid infiltrate, but are also important in the revelation of a preserved or destroyed normal architectural pattern. With the use of antibodies to kappa and lambda immunoglobulin light chains, and the determination of the light chain ratio, plasma cells or other immunoglobulin-producing lymphoid cells can be judged as polyclonal or monoclonal, the latter at least documenting a monoclonal expansion of a B cell population, if not its malignant nature.

By the application of other antibodies, certain biological features (e.g. the Ki67 antigen reflecting the proliferative index of a neoplasm) or prognostic features such as the atypical expression of tumour-suppressor genes (e.g. a mutated *p53* gene) or the formation of a fusion gene (e.g. *NPM-ALK*) may be determined.

Table 2 gives an overview of antibodies useful in the daily diagnostic practice of lymphoid neoplasms. This table lists antibodies to CD antigens and other antibodies reliably working on paraffin sections provided that antigen retrieval is performed. The specificity indicated represents only basic reactivity. Some of the antibodies, therefore, may have a broader spectrum of reactions.

It should be strictly kept in mind, however, that immunophenotypic studies can only be an adjunct to morphological diagnosis. For example, aberrant monoclonal B cell populations may be identified in non-neoplastic disorders, e.g. in certain autoimmune diseases. Therefore, 'monoclonality' by no means equates with 'malignancy'.

Table 2 Overview of antibodies useful in diagnosis of lymphoid neoplasms

Antigen	**Specificity**
CD2	T cells
CD3	T cells
CD4	Helper T cells
CD5	T cells, B cell subpopulation
CD8	Cytotoxic T cells
CD10	Germinal centre B cells
CD15	Neutrophils, Hodgkin and Reed–Sternberg cells
CD20	B cells
CD21	B cell subpopulation, follicular dendritic cells
CD23	Follicular mantle cells, follicular dendritic cells
CD30	Activated B and T cells, Hodgkin and Reed–Sternberg cells
CD34	Precursor myeloid and lymphoid cells
CD43	T cells, B cell subpopulation, myelomonocytic cells
CD45	Leukocyte common antigen
CD45R0	T cells
CD56	T/NK cells, NK cells
CD57	NK cells
CD68	Monocytes, macrophages
CD79a	B cells, plasma cells
CD138	Plasma cells
Others	**Specificity**
TdT	Precursor B and T cells
TIA 1, granzyme B	Cytolytic/cytotoxic cells
DBA44	Hairy cells
Bcl-2	Proto-oncogene product
p53	Suppressor gene product
p27	Cdk inhibitor
p21	Cdk inhibitor
Cyclin D1	Cdk activator
VS38c	Plasma cells

GENOTYPIC STUDIES

The detection of a clonal population of B or T cells refers to the capability of these cells to rearrange physically either their immunoglobulin heavy (IgH) and light (IgL) chain genes or their T cell receptor genes. Some other gene rearrangements, such as in chromosomal translocations, also alter the spatial arrangement of the DNA in the nucleus, thus being accessible to DNA rearrangement studies. The Southern blot technique as a means of determining clonality in a lymphocytic proliferation may detect clonal populations as low as 1–5% of cells relative to the total number of cells in a (fresh) specimen, thus clearly exceeding the threshold of pure morphological analysis. In addition to the detection of a clonal cell population represented by a non-germline band (or bands), the technique may also be used for distinguishing T or B lineage neoplasms.

The most important shortcoming of this technique, the requirement for fresh material, can be overcome by the use of the polymerase chain reaction (PCR) technique. The PCR, in its basic principle, represents a technique in which small amounts of DNA can be amplified *in vitro*, provided that the DNA sequences flanking the regions looked for are known. Because of the requirement of only minimal amounts of DNA (or small degraded DNA particles), this technique can be used for the detection of clonal cell populations or DNA rearrangements also in paraffin-embedded formalin-fixed material (in which the DNA normally is largely degraded). Because of this inherent advantage, the PCR may also be used in the monitoring of minimal residual disease, especially if clonotypic primers are used.

By DNA sequence analysis of Ig receptor genes, non-mutated (naive) prefollicular and mutated (memory) postfollicular B cells may be distinguished. In follicular cell populations, and some of their descendants, various 'ongoing' somatic mutations show a micropolymorphism of the B cell receptor repertoire. The detailed analysis of Ig receptor genes, therefore, permits conclusions on the status of antigen-dependent selection and mutation as well as the V_H gene repertoire.

Karyotypic Studies

Malignant lymphomas, especially the non-Hodgkin lymphomas of B type, are fairly well characterized – at least with respect to their primary genetic alterations – on the cytogenetic level. Mitelman's catalog of 'Chromosome aberrations in cancer' lists over 2000 lymphoid neoplasms on which conventional cytogenetic (banding) studies have been performed. Apart from proving the neoplastic nature of a lymphoid cell proliferation, the description of characteristic cytogenetic aberrations in certain types of malignant lymphomas has greatly added to our understanding of the biology of malignant lymphomas and has also influenced taxonomy, as may best be exemplified by the close association of mantle cell lymphoma to the translocation t(11;14)(q13;q32), the recognition of which ultimately led to the worldwide acceptance of mantle cell lymphoma as an entity on its own.

Table 3 lists the most common and characteristic chromosome aberrations in malignant lymphoma.

At present, malignant lymphoid tumours can not be classified according to their primary genetic aberrations alone, because (1) in some entities, no characteristic aberrations have been detected so far, and (2) because a recurring chromosomal translocation may be encountered in different lymphoma entities, e.g. the t(14;18)(q32;q21) in follicular lymphoma and diffuse large B cell lymphoma. In addition, cytogenetically detectable alterations are recognized only in a fraction of a given lymphoma entity.

More recently, the use of fluorescent dye-conjugated DNA probes in *in situ* hybridization has overcome some

Table 3 Characteristic chromosome aberrations in malignant lymphoma

Diagnosis	Chromosome aberration	Genes involved
Precurser B cell lymphoblastic lymphoma/leukaemia	t(1;19)(q23;p13)	c-*ski* (1q23)
	t(4,11)(q21;q23) del(6q)	c-*ets1* (11q23-24)
	t(9;22)(q34;q11)	*bcr-abl*
B cell chronic lymphocytic leukaemia/small lymphocytic lymphoma	del(13)(q14)	
	Trisomy 12	
Mantle cell lymphoma	t(11;14)(q13;q32)	*Cyclin D1*
Follicular lymphoma	t(14;18)(q32;q21)	*BCL-2*
Extranodal marginal zone B cell lymphoma of MALT type	t(11;18)(q21;q21)	*API2/MLT1*
Splenic marginal zone B cell lymphoma	t/del(7)(q22-32) del(10)(q22-24)	
Lymphoplasmacytic lymphoma (immunocytoma)	t(9;14)(p13;q32)	*PAX-5*
Diffuse large B cell lymphoma	t(3;14)(q27;q32)	*BCL-6*
	t(14;18)(q32:q21)	*BCL-2*
Burkitt lymphoma	t(8;14)(q24;q32)	c-*myc*
	t(2;8)(p12;q24)	
	t(8;22)(q24;q11)	
Plasmacytoma	t(4;14)(p16;q32)	*FGFR3*
	t(6;14)(p25;q32)	*MUM-1/IRF 4*
Precursor T cell lymphoblastic lymphoma/leukaemia	14q11	*TCR* genes
	7p15 or 7q34-35	*TCR genes*
T cell prolymphocytic leukaemia	inv(14)(q11q32)	*TCL-1*
Angioimmunoblastic T cell lymphoma	+3, +5, +X	
Anaplastic large cell lymphoma	t(2;5)(p23;q35)	*NPM/ALK*
Hepatosplenic gd T cell lymphoma	i(7)(q10)	

limitations of conventional cytogenetics, i.e. the need for viable dividing cells. With the use of the fluorescence *in situ* hybridization (FISH) technique, cytogenetic investigations can today also be performed on interphase cells, thus allowing for the recognition of amplifications, deletions or even translocations without the need to cultivate cells. This technique is sometimes also referred to as 'molecular cytogenetics'. With the aid of the FISH technique, it has been possible to define independent prognostic parameters in certain lymphoma entities which, in some cases, may be equivalent to the items of the International Prognostic Index. *ATM* or *p53* deletions in B cell chronic lymphocytic leukaemia, for example, are strong negative prognostic indicators.

The comparative genomic hybridization (CGH) technique allows for an overview of genetic imbalances (overrepresentations and deletions) in malignant tumours and may be used to calculate the overall genomic instability, which appears to constitute a new important prognostic feature.

LYMPHOMA CLASSIFICATION

History

The history of the recognition and classification of tumours that today are called malignant lymphomas is long, controversial and complicated. In 1832, Sir Thomas Hodgkin first identified malignant tumours of lymph nodes. Among those were cases that we would now classify as Hodgkin lymphoma. The definition of 'leukaemia' followed in 1845 by Rudolf Virchow and in 1863 he laid down the concepts of lymphosarcoma and lymphoma. At the turn of the nineteenth century, Sternberg and Reed, in 1898 and 1902, defined morphologically the tumour cells of Hodgkin disease. About 25 years later, Brill and Symmers described follicular lymphoma and in 1948 Burkitt and in 1960 O'Connor and Davis described African lymphoma, now called Burkitt lymphoma. Within this short list, of course, only historical milestones have been cited, constituting only those steps that still remain valuable in the twenty-first century.

Since 1960, modern immunology has greatly influenced our knowledge of lymphoid tumours. The biology of lymphoma cells, reflecting the immunological activities of their normal counterparts, has been clarified to a great extent and has allowed the establishment of a biological basis for new and more comprehensive classification systems of malignant lymphomas. Two classification systems have been widely used until recently, the Kiel Classification of Non-Hodgkin Lymphomas (Stansfeld *et al.*, 1988; Lennert and Feller, 1992) and the Working Formulation for Clinical Usage (Non-Hodgkin's Lymphoma Pathologic Classification Project, 1982). The Kiel classification was based on the exact morphological description and immunological identification of the normal cellular counterparts of tumour cells and was updated several times, introducing

new findings and more comprehensive knowledge. The so-called Working Formulation, on the other hand, was based on historical clinical survival data and, therefore, was not updated, although its usage was also adapted to modern findings. Of importance, there was a geographic split of categories in the diagnosis of lymphoid tumours, the Kiel classification in Europe and the Working Formulation in the USA.

An international group of experienced haematopathologists, the International Lymphoma Study Group (ILSG) formulated a new proposal for a modern lymphoma classification and published it as the so-called Revised European American Lymphoma (REAL) Classification (Harris *et al*., 1994). In the following years, this proposal was tested for its applicability and reliability and then accepted worldwide. It is, therefore, largely identical (with some minor corrections and additions) with the now proposed WHO classification.

Basic Taxonomic Principles

The taxonomic unit of the new WHO classification is the **disease entity**. Distinct entities in malignant lymphomas can be recognized by pathologists and, ideally, are of clinical relevance. For each disease entity, a combination of morphological, immunophenotypic, genetic and clinical features is needed for its definition. The relative value of each of these features may vary among different disease entities. Variations in grade and aggressiveness, that may exist within a given disease entity, and may be related to patients' survival and treatment response, must be distinguished from 'different diseases' (Jaffe *et al*., 1998; Harris *et al*., 1999, 2000).

The basic rules for the definition of disease entities in this proposal are comparable to the general rules of tumour classification as used in many other organs and organ systems. A given lymphoma entity is defined in first line by recognition of the predominant differentiated cell type using morphological and immunological features, a principle that follows the rules of the Kiel classification. In addition, the importance of the primary site of involvement, which is not only a feature of staging, but also an easy, clinically relevant and important biological distinction, has been explicitly stated.

Second-line principles of classification are important for some entities. These are **aetiological features,** such as the association of certain infectious agents, like the Epstein–Barr virus, Helicobacter pylori or HTLV I, **primary cytogenetic abnormalities** or **specific clinical features**. Daily clinical and pathological experience shows minor or more evident exceptions to the proposed rules, leading to the well-recognized heterogeneity of each type of lymphoid neoplasia. Therefore, within many entities, specific morphological or clinical subtypes are mentioned which are of clinical importance. Morphological variants reflect the diagnostic spectrum of a disease, which is important to recognize in order to establish a correct differential diagnosis.

Table 4 lists the recent WHO classification of B and T cell lymphomas and of Hodgkin lymphoma.

The non-Hodgkin lymphomas are divided primarily into those of the B and the T cell system. In both lineages, there is a primary distinction of lymphomas that arise from precursor cells (the lymphoblastic lymphomas or acute lymphoid leukaemias) and from peripheral cells.

Clinical Relevance

The clinical relevance of the WHO proposal was intriguingly confirmed by the results of the Non-Hodgkin's Lymphoma Classification Project (1997), which involved the pathomorphological and immunophenotypic investigation of 1403 lymphomas in eight different sites around the world. These cases were taken from the years 1988–1990, and were reviewed by experienced haematopathologists attempting to compare different classification systems, the recognition of epidemiological variations in the occurrence of lymphomas and the correlation of treatment results with histological diagnoses. The results of this study clearly established that the criteria formulated in the REAL classification and applied in the WHO classification resulted in a high inter-observer accuracy and were of significant prognostic value for the recognition of diseases with different clinical courses and behaviour. It turned out that immunophenotyping was less important in some of the diseases (e.g. in follicular lymphoma), but absolutely essential in others (mantle cell lymphoma, T cell lymphomas). Some diseases were only reliably diagnosed if clinical data were available (e.g. mediastinal B cell lymphoma). In other cases, the differential diagnosis was at least greatly improved by the knowledge of clinical features and presentation (Armitage and Weisenburger, 1998).

For some disease entities and variants, however, diagnostic accuracy was less satisfactory. In particular, the distinction of Burkitt lymphoma and 'Burkitt-like' non-Hodgkin lymphoma involved a high inter-observer disagreement of more than 40%, clearly establishing that additional and better criteria will have to be proposed for the definition of certain disease entities. Similarly, the subclassification of peripheral T cell lymphomas showed high inter-observer variability.

Prognostic Factors

As was clearly shown by the international Non-Hodgkin's Lymphoma Classification Project (1997) and many other multicentre trials published in recent years, the most important prognostic factor today is the definition of the disease entity, that is, the exact type of non-Hodgkin lymphoma. Within these individual diseases, prognostic

Table 4 The WHO classification of lymphoid neoplasms

B cell neoplasms

Precursor B cell lymphoblastic leukaemia/lymphoma (Precursor B cell acute lymphoblastic leukaemia)

Peripheral B cell neoplasms
B cell lymphocytic leukaemia/small lymphocytic lymphoma
B cell prolymphocytic leukaemia
Lymphoplasmacytic lymphoma
Mantle cell lymphoma
Follicular lymphoma
Cutaneous follicle centre lymphoma
Marginal zone B cell lymphoma of mucosa-associated lymphoid tissue (MALT type)
Nodal marginal zone B cell lymphoma ($\pm$ monocytoid B cells)
Splenic marginal zone B cell lymphoma ($\pm$ villous lymphocytes)
Hairy cell leukaemia
Diffuse large B cell lymphoma
Variants: Centroblastic
Immunoblastic
T cell or histiocyte-rich
Anaplastic large B cell
Subtypes: Mediastinal (thymic) large B cell lymphoma
Intravascular large B cell lymphoma
Primary effusion lymphoma
Burkitt lymphoma
Plasmacytoma
Plasma cell myeloma

T cell neoplasms

Precursor T cell lymphoblastic leukaemia/lymphoma (Precursor T cell acute lymphoblastic leukaemia)

Peripheral T cell and NK cell neoplasms
T cell prolymphocytic leukaemia
T cell large granular lymphocytic leukaemia
Aggressive NK cell leukaemia
NK/T cell lymphoma, nasal and nasal type
Sezary syndrome
Mycosis fungoides
Angioimmunoblastic T cell lymphoma
Peripheral T cell lymphoma (unspecified)
Adult T cell leukaemia/lymphoma (HTLV 1+)
Anaplastic large cell lymphoma (T and null cell types)
Primary cutaneous CD30 positive T cell lymphoproliferative disorders
Variants: Lymphomatoid papulosis (Type A and B)
Primary cutaneous ALCL
Borderline lesions
Subcutaneous panniculitis-like T cell lymphoma
Enteropathy-type T cell lymphoma
Hepatosplenic γ/λ T cell lymphoma

Hodgkin lymphoma (Hodgkin disease)

Nodular lymphocyte predominance Hodgkin lymphoma

Classical Hodgkin lymphoma
Hodgkin lymphoma, nodular sclerosis
Classical Hodgkin lymphoma, lymphocyte-rich
Hodgkin lymphoma, mixed cellularity
Hodgkin lymphoma, lymphocyte depletion

factors may influence clinical outcome (Shipp *et al.*, 1993). Prognostic factors and variations in grades within diseases should be distinguished from different diseases. They may be histological, biological or clinical in nature, such as stage or the International Prognostic Index. Histological grading is one method to define types of prognostic factors. Usual approaches include the determination of cell size, nuclear features, mitotic rates and growth pattern. In recent years, biological markers, such as genetic features, have turned out to be important prognostic factors and may even be more powerful than clinical or morphological features. Some of them may be recognized today by interphase cytogenetics (such as *p53* and *ATM* deletions in B-CLL), or by immunohistochemistry (such as the presence of the t(2;5/ALK) rearrangement in anaplastic large-cell lymphomas of T and O cell types, or the determination of the proliferative index in mantle cell lymphomas using the Ki67 antibody). More recently, exciting new data have become available taking into account mutated or unmutated IgVH genes in B-CLL or the subdivision of diffuse large B cell lymphomas according to their mRNA expression profiles.

These examples show how dependent a given classification is on new facts and findings that may be used for the definition of diseases in the borderline and grey zone of different diseases. It does not need great fantasy to predict that the new approaches of molecular biology will further modify our concepts of the biology of malignant lymphoma, and that in the not too distant future our concepts of lymphoma classification will be stepwise modified and possibly reverted to a molecular definition of neoplastic processes in different lymphoma entities (Alizadeh *et al.*, 2000).

LYMPHOMAS OF B CELL LINEAGE

Precursor B Cell Lymphoblastic Leukaemia/Lymphoma (Precursor B Cell Acute Lymphoblastic Leukaemia)

Clinical Features

The peak incidence of lymphoblastic lymphomas, irrespective of B or T cell lineage, is in the second and third decades, and the male gender is over-represented in the range 2–10:1. The majority of patients are diagnosed with advanced disease stages, and central nervous system involvement is common. About 60% of patients with no obvious bone marrow infiltration at presentation are reported to develop marrow infiltration and leukaemic conversion. Only 20% of the so-called lymphoblastic lymphomas are of B cell lineage, in contrast to 80% of acute lymphoblastic leukaemias.

Morphology

Precursor B cell neoplasias, in most cases, show a diffuse infiltration and destruction of the normal tissue structure. The tumour cells are medium-sized with a scant cytoplasm and round to oval or convoluted nuclei, often with distinct nuclear membranes, a finely dispersed, sometimes 'dust'-like chromatin and inconspicuous nucleoli. In some cases, a 'starry sky' pattern may be seen due to the presence of tingible body macrophages. Usually, numerous mitotic figures are encountered with numbers of 10–20 per high-power field (HPF). Owing to their aggressive nature, they often infiltrate the lymph node capsule and invade the surrounding perinodal fatty or fibrous tissues.

Immunophenotype

The majority of precursor B cell lymphoblastic lymphomas express B cell antigens similar to or identical with B cell acute lymphatic leukaemia. They express HLA-DR antigens and B cell-associated antigens CD19 and/or CD22 and may be positive for IgM and Ig light chains. In paraffin sections, most of them are positive for CD20 and CD34, as well as for CD10 (cALLA). Of pivotal importance in the diagnosis of precursor B (and T) cell neoplasias is the detection of nuclear expression of the terminal deoxynucleotidyl transferase (TdT), thereby clearly differentiating these neoplasms both from chronic lymphocytic leukaemia/small lymphocytic lymphomas and diffuse large cell lymphomas.

Genetic Features

The majority of precursor B cell neoplasms express clonal IgH or IgL gene rearrangements. Characteristic cytogenetic abnormalities in precursor B cell neoplasias are numerical chromosome changes (hyperploidy), t(1;19)(q23;p13) involving c-*ski*, t(4;11)(q21;q23) involving c-*ets1*, del(6q), del(12p) and t(9;22)(q34;q11) involving *bcr-abl*.

PERIPHERAL B CELL LYMPHOMAS

Chronic Lymphocytic Leukaemia of B Type (B-CLL)/Small Lymphocytic Lymphoma (B-SLL)

Clinical Features

B-CLL comprises 90% of chronic lymphoid leukaemias in Europe and the USA. Most patients will present with generalized lymphadenopathy, bone marrow and blood involvement and enlarged spleen. Hepatomegaly is frequent. A minority of patients initially present with a leukaemic nodal involvement.

Morphology

Lymph nodes in patients with B-CLL show a characteristic infiltration pattern in roughly 90% of cases which may be recognized at low-power magnification. In most B-CLL, there is a so-called 'pseudo-follicular' growth pattern. In these pseudo-follicles (also termed proliferation centres), slightly larger lymphocytes (prolymphocytes and paraimmunoblasts) are found in contrast to the majority of the infiltrate which consists of small lymphoid cells with scant cytoplasm and slightly clumped chromatin. In the proliferation centres, which only very occasionally show some remnants of follicular dendritic cells, the larger cells possess a small cytoplasmic rim and a lighter nuclear chromatin, so that these pseudo-follicles stand out, imparting 'light zones' to the infiltrate (**Figure 1**). Only about 10% of B-CLL/SLL are characterized by a diffuse infiltration without pseudo-follicles, and this represents the predominant pattern in extranodal infiltrates of the disease.

Variants include cases with a larger number of prolymphocytes and/or paraimmunoblasts (B-CLL/PL), tumours in which a sometimes marked nuclear irregularity can be seen (and which may be difficult to differentiate from mantle cell lymphomas), and cases with a secretory differentiation (monotypic cytoplasmic light-chain expression) corresponding to the lymphoplasmacytoid immunocytoma of the Kiel classification system.

Immunophenotype

The tumour cells of B-CLL are characteristically positive for B cell-associated antigens (CD19, CD20 (weak), CD79a) and IgM ± IgD, and express CD5, CD23, and CD43. Recognition of CD23 expression is important in the differential diagnosis of mantle cell lymphoma, in which CD23 (on paraffin or frozen sections) is not expressed. In B-CLL, CD23 expression may be pronounced in the proliferation centres, and it is also more often preserved than CD5 in transformed large B cell lymphomas having evolved from B-CLL.

Genetic Features

Ig heavy- and light-chain genes are rearranged. Of importance, recent studies have shown that, by the analysis of somatic mutations in IgVH genes, B-CLL is not a homogeneous disease, but may contain at least two important subentities. In roughly 50% of cases, the tumour cells do not display somatic mutations and hence may be referred to as derived from naive or virgin B cells. In the other 50% of cases, somatic mutations have occurred, indicating a germinal centre passage of the tumour cells. As judging from clinical data available to date, the mutated form of B-CLL may have a better prognosis.

In contrast to many other non-Hodgkin lymphomas, B-CLL is not characterized by a reciprocal chromosome

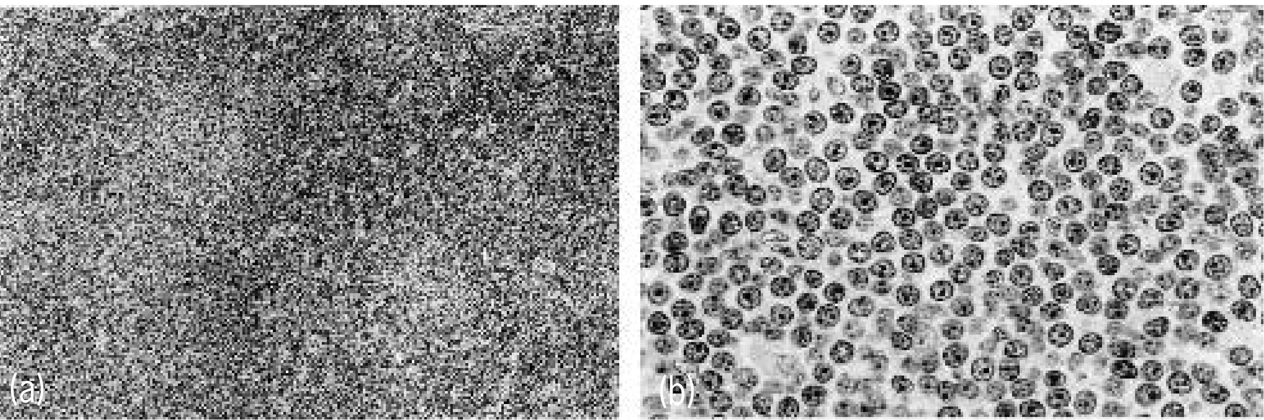


Figure 1 B cell chronic lymphocytic leukaemia. (a) The normal lymph node architecture is replaced by a pseudo-follicular infiltrate. (b) The infiltrate consists of small cells with round nuclei and few prolymphocytes.

translocation. The most common cytogenetic/molecular cytogenetic aberration is a deletion in chromosomal band 13q14 encountered in up to 40% of cases. Other recurring aberrations include trisomy 12, 6q deletions and deletions/mutations in 17p13 and 11q22/23. Deletions in 17p frequently target the *p53* gene and 11q deletions commonly result in the loss of one copy of the *ATM* (mutated in Ataxia telangiectasia) gene. Both aberrations have been shown to be prognostically relevant in defining patient subgroups with inferior prognosis.

In B-CLL, prolymphocytoid transformation or transformation to large B cell lymphomas (Richter syndrome) may occur. Transformation to Hodgkin disease was also observed and should be distinguished – by the appropriate background infiltrate and the characteristic immunophenotype of Hodgkin and Reed–Sternberg cells – from the occasional occurrence of Hodgkin- and Reed–Sternberg-like cells in B-CLL.

B Cell Prolymphocytic Leukaemia (B-PLL)

Clinical Features

B-PLL is a rare disease, occurring in older patients over 60 years of age with a marked male predominance presenting with massive splenomegaly, bone marrow infiltration and marked lymphocytosis in the peripheral blood (>55% prolymphocytes). Lymph nodes are rarely involved. Anaemia and thrombocytopenia are frequent.

Morphology

The infiltrate in the bone marrow consists of diffusely infiltrating medium-sized prolymphocytic or paraimmunoblastic cells with a small rim of slightly basophilic or pale cytoplasm, round to oval nuclei and a central prominent eosinophilic nucleolus. No pseudo-follicles are seen.

Immunophenotype

The tumour cells in B-PLL express pan B cell markers, IgM $\pm$ IgD and CD5 in a proportion of cases. CD23 usually is negative.

Genetic Features

Clonal rearrangements of IgH and IgL chain genes are found. Most of the tumours have been shown to carry somatic mutations. The occurrence of cases with t(11;14)(q13;q32) points to an overlap with mantle cell lymphoma ('mantle cell leukaemia').

Mantle Cell Lymphoma (MCL)

Clinical Features

MCL is a disease of middle-aged to older persons. Most patients present in stages III and IV. The tumour tends to involve lymph nodes and Waldeyer's ring. The bone marrow is often infiltrated, and a proportion of patients have prominent splenomegaly and peripheral blood involvement. The gastrointestinal tract is often infiltrated (lymphomatous polyposis). Sometimes, a high peripheral leukocyte count is found (mantle cell leukaemia). Blastic transformation may occur (blastoid MCL).

Morphology

Mantle cell lymphoma, in most cases, presents with diffuse lymph node infiltrates of monomorphic small to medium-sized cells with scant, barely recognizable cytoplasm and irregular or cleaved nuclei. The chromatin is slightly dispersed and inconspicuous nucleoli are found. In rare cases, the cells are nearly round, rendering a differential diagnosis to B-CLL. Two blastoid variants have been described, one with medium-sized nuclei with a more dispersed chromatin resembling lymphoblasts and a high proliferative index

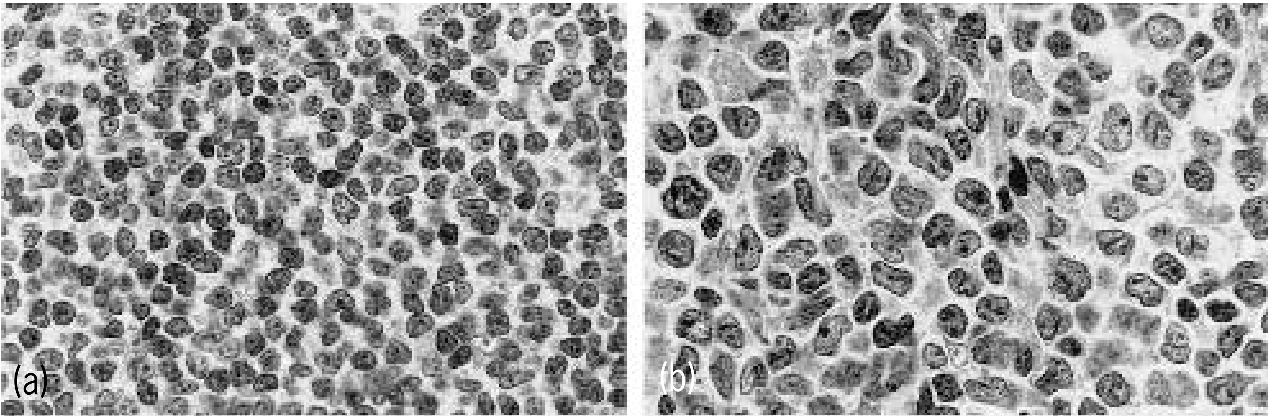


Figure 2 Mantle cell lymphoma. (a) In the classical variant, small- to medium-sized cells are seen, with small cytoplasm and irregular nuclei. (b) Pleomorphic variant of MCL. Cells are considerably larger and nuclei are deeply indented. The chromatin distribution is coarse.

(blastoid or lymphoblastoid type). Another variant variously termed as 'anaplastic,' 'large cell' or 'pleomorphic' is characterized by rather large cleaved cells with sometimes light or slightly basophilic cytoplasm and coarse chromatin distribution. This type obviously had been included in the Kiel classification as the 'centrocytoid' variant of centroblastic lymphoma (**Figure 2**).

In some cases, the infiltrate in mantle cell lymphoma may be found predominantly surrounding (partially) preserved germinal centres giving the impression of a 'mantle zone' pattern or may form sometimes vaguely circumscribed nodules.

Immunophenotype

The tumour cells of MCL are characterized by the expression, next to B cell antigens, of CD5 and CD43 in the absence of CD10 and CD23. Notably, the characteristic overexpression of Cyclin D1 can be recognized by suitable antibodies. Recently, it has been shown that the classical variant of MCL displays – in contrast to other non-Hodgkin lymphomas with comparably low proliferative indices – a marked down-regulation of nuclear p27 expression, a feature which is highly characteristic of and favours mantle cell lymphoma. Staining for follicular dendritic cells (using anti-CD21 or -CD23 antibodies) may show, especially in the cases with a nodular growth pattern, dispersed and disorganized meshworks of these cells.

Genetic Features

The cytogenetic hallmark of MCL is the chromosomal translocation t(11;14)(q13;q32), joining the Ig heavy-chain locus and the *BCL-1* locus and resulting in the overexpression of the cell-cycle protein Cyclin D1. It should be noted, however, that the t(11;14)/Cyclin D1 deregulation is not absolutely specific for mantle cell lymphomas, having also been described in plasmacytoma/multiple myeloma. Mantle cell lymphomas, in addition, show characteristic secondary aberrations, many of them targeting other cell cycle-related genes such as *p16*, *p53* and the *ATM* gene. Blastoid MCL of the pleomorphic type frequently are characterized by tetraploid chromosome clones.

Follicular Lymphoma (FL)

Clinical Features

Follicular lymphomas represent, together with diffuse large B cell lymphomas, the most frequent type of B cell lymphoma in Western countries. Both genders are equally affected. Systemic disease at presentation is common, including involvement of lymph nodes, spleen, bone marrow and, rarely, extranodal sites. PB involvement is rare and tends to occur only in final stages. The disease follows an indolent course.

Morphology

In the most common type with an almost exclusive follicular growth pattern, neoplastic follicular structures are found throughout the lymph node and also frequently penetrate the capsule and invade the perinodal fatty tissue. In contrast to reactive follicles, they are uniform in size and lack a well-defined mantle zone. Cytologically, they are composed of centrocytes and centroblasts. Centrocytes represent medium-sized cells with a small, barely recognizable cytoplasm and irregular contoured and indented nuclei. The chromatin is irregular, and small inconspicuous nucleoli are present. In contrast, centroblasts are large cells with a usually small, moderately basophilic cytoplasm, and round nuclei with vesicular chromatin and 1–3 nucleoli commonly found adjacent to the nuclear membrane. Centrocytes and centroblasts are arranged haphazardly, and a well-formed dark and light zone as seen in reactive germinal centres structures is not found (**Figure 3**). Likewise, in most cases, no tingible body

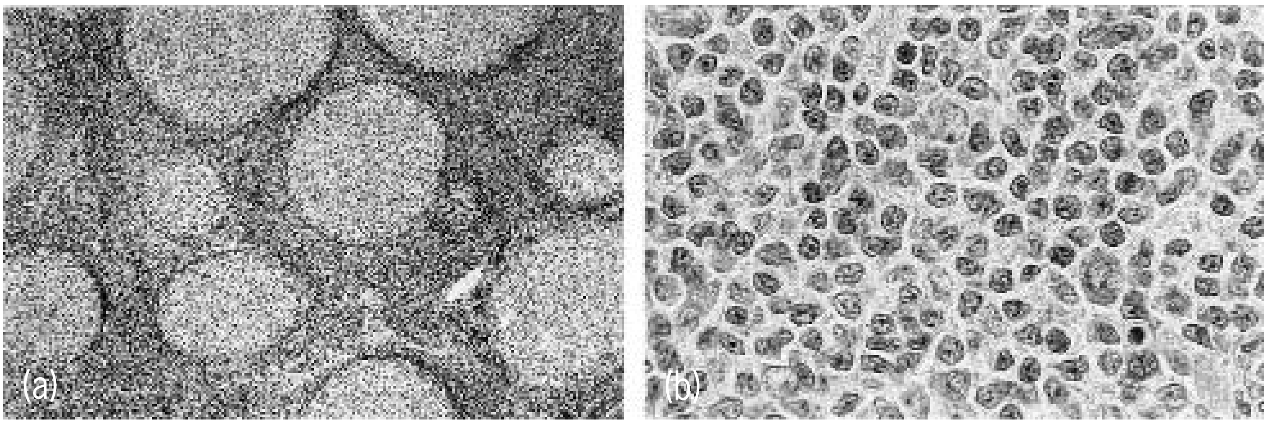


Figure 3 Follicular lymphoma. (a) The lymph node parenchyma shows an infiltration of atypical follicular structures. (b) On higher magnification, the neoplastic follicles are composed of small cleaved cells (centrocytes) with few interspersed blasts (centroblasts).

macrophages are present and the number of mitotic figures is low. Diffuse areas may be present in addition, and tumours may be classified as predominantly follicular (>75% follicular), follicular and diffuse (25–75% follicular) or predominantly diffuse (<25% follicular). Entirely diffuse follicular lymphomas, composed of centrocytes and centroblasts, are very rare neoplasms and caution has to be exerted in differentiating them from mantle cell lymphomas.

The WHO classification proposes a grading scheme for follicular lymphomas according to the content of centroblasts. Therefore, tumours with 0–50 centroblasts per 10 high-power fields (HPF) are grade 1, tumours with 50–150 centroblasts per 10 HPF are grade II and lymphomas with more than 150 centroblasts/10 HPF are grade III. Lymphomas with an exclusively follicular growth pattern and consisting entirely of centroblasts are very rare. They are more commonly found in association with a diffuse large B cell lymphoma and should be diagnosed as 'follicular lymphomas grade III with diffuse large B cell lymphoma'. There are some indications, however, that these tumours, on a genetic basis, are different from follicular lymphomas with varying centroblast content, but preserved maturation to centrocytes, being t(14;18) negative in 80% of the cases and more closely related to a 'follicular' variant of diffuse large B cell lymphomas. Follicular lymphomas grades I or II associated with a diffuse large cell component, in most cases, should be regarded as 'transformed' follicular neoplasias.

Some follicular lymphomas may show a more or less prominent 'marginal zone' differentiation, implying the occurrence of tumour cells slightly rounder than centrocytes and with a broad pale-staining cytoplasm reminiscent of monocytoid B cells. In most cases, these cells are accentuated in the outermost parts of the follicles, giving the impression of a marginal zone pattern. A differentiation to plasma cells (secretory differentiation) may be rarely seen in follicular lymphomas.

Immunophenotype

Most follicular lymphomas grades I and II express CD10 and show a nuclear reactivity for BCL-6, but lack CD5 and CD43 reactivity. In some tumours, CD23 may be expressed, and these lymphomas are frequently CD10-negative. One intriguing immunohistochemical finding in follicular lymphomas, and strongly suggesting a neoplastic process (in the differential diagnosis to reactive follicular hyperplasia), is the presence of CD10+ B cells in the interfollicular region. About 80–90% of follicular lymphomas express BCL2, a feature clearly differentiating them from reactive follicles which are always BCL2-negative. By staining of follicular dendritic cells using CD21 and/or CD23, FDC meshworks are usually dense and sharply demarcated. In BCL2-negative lymphomas, the diagnosis may require gene rearrangement studies.

Genetic Features

Apart from clonal rearrangements of (extensively hypermutated) IgH and IgL chain genes, follicular lymphomas are characterized by 'ongoing' somatic mutations, creating an intraclonal diversity. The cytogenetic hallmark of follicular lymphomas is the t(14;18)(q32;q21) chromosome translocation, that is found in 80–90% of grade I and II tumours. As already pointed out, follicular lymphomas grade III are t(14;18)-positive in only 50% of cases, and especially grade III lymphomas that are composed of centroblasts exclusively in the majority of cases are t(14;18) negative. On the molecular genetic level, the t(14;18) leads to the juxtaposition of the *BCL2* oncogene to the IgH chain gene promoter region, thus overexpressing *BCL2*, which is, as already mentioned, not expressed in centroblasts and centrocytes of reactive follicles. Secondary chromosome aberrations are common and well-defined in follicular lymphomas. Especially trisomies for chromosomes 1q, 7, 12, 18 and X are found, while deletions are encountered in chromosome regions 1p and 6q.

Transformed follicular lymphomas frequently are characterized by deletions in the short arm of chromosome 17, the site of the *p53* tumour suppressor gene. The retained allele has been found to be mutated in a fairly large number of transformed follicular lymphomas.

Marginal Zone B Cell Lymphomas (MZBL)

Marginal zone B cell lymphomas (MZBL) are B cell neoplasias with a presumed origin from the marginal zone cells of the B cell follicle. Three variants are discerned: extranodal MZBL of MALT type, splenic and nodal MZBL. Whereas the REAL classification regarded nodal and splenic MZBL as provisional entities, the WHO classification, taking into account the immunophenotypic and genetic differences in MZBL, regards them as true lymphoma entities.

Extranodal MZBL of MALT Type

Clinical Features

MALT-type MZBL frequently arise in extranodal sites commonly devoid of a regular lymphatic parenchyma. Therefore, MALT has to be introduced in organs such as the stomach, the lungs, the thyroid and salivary glands, the ocular adnexa, the skin or the mammary gland (the most frequent localizations) by chronic inflammatory processes such as chronic infections or autoimmune diseases prior to the development of these enigmatic tumours. The gastrointestinal tract is the most common site of involvement. Gastric MALT-type MZBL frequently (in 90% of cases) arise in the background of a chronic *Helicobacter pylori* (HP) gastritis. Upon eradication of the bacterium, about 80% of MALT-type MZBL regress and hence patients may be cured by this antibiotic therapy. However, clonal rearrangements for the IgH and IgL chain genes may be demonstrated also after complete regression of the lymphoma, the prognostic significance of which is not yet entirely understood. In rare cases, after eradication, a monoclonal plasma cell population may still be present in biopsies, representing a plasma cellular maturation of the tumour cells.

The majority of patients present with stage I or II disease, reflecting the particular feature of these tumours to be confined to their site of origin for long times.

Morphology

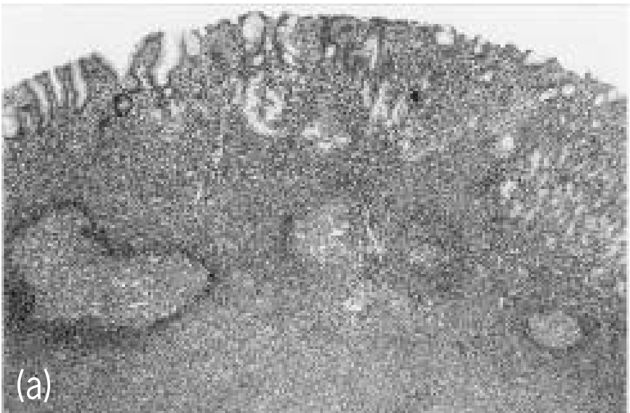
Extranodal MZBL, in early lesions, infiltrate the marginal zones of frequently preserved reactive follicular structures and, as a rule, present with an invasive and destructive growth potential forming so-called lymphoepithelial lesions (**Figure 4**). On the cytological level, marginal zone B cells are small- to medium-sized with a small to moderately broad, sometimes pale cytoplasm and slightly indented 'centrocytoid-like' nuclei. These tumours may be associated with a diffuse large B cell component and should then be designated as 'diffuse large B cell lymphomas with an extranodal MZBL of MALT-type component' (not as 'low- and high-grade MALT-type lymphoma' or 'transformed MALT-type lymphoma').

Immunophenotype

MALT-type MZBL are CD5−, CD10−, CD23−, CD43−/+ and IgD−. In contrast to mantle cell lymphomas (frequently presenting as primary gastrointestinal tumours, the so-called malignant lymphomatous polyposis), they are consistently cyclin D1 negative.

Genetic Features

The t(11;18)(q21;q21) chromosome translocation, juxtaposing the *API2* and *MLT1* genes, has been identified as the most common structural cytogenetic alteration in extranodal MZBL of MALT-type accounting for about 30% of cases (and even 50% of cases without any large-cell component). Interestingly, this translocation



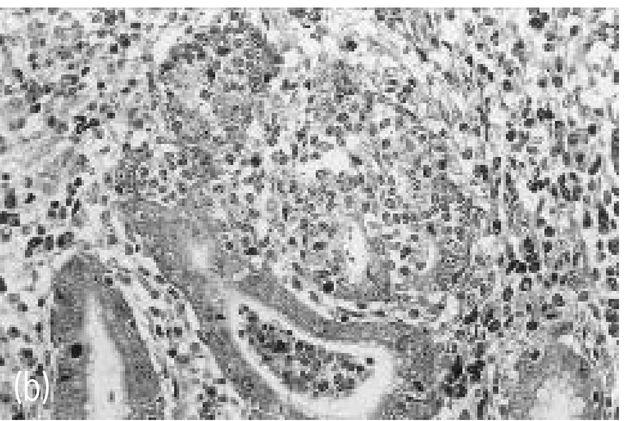


Figure 4 Extranodal marginal zone B cell lymphoma of MALT type. (a) Gastric lymphoma. The infiltrate colonizes the marginal zone areas of preserved follicles and spreads out into the mucosa and submocosa. (b) The tumour cells (centrocyte-like cells) invade and destroy epithelial structures, forming lymphoepithelial lesions.

has not been found in extranodal large B cell lymphomas with or without an extranodal MZBL MALT-type component.

Splenic MZBL ($\pm$Villous Lymphocytes)

Clinical Features

Splenic marginal zone lymphomas (SMZL), a distinctive type of primary splenic lymphoma, has to be differentiated from secondary splenic involvement occurring in B-CLL, B-PLL, mantle cell lymphoma, immunocytoma and follicular lymphoma. SMZL involves the spleen and frequently the bone marrow. Hilar splenic and abdominal lymph nodes may be infiltrated, and leukaemic dissemination often occurs. Peripheral lymphadenopathy is uncommon. SMZL is a rare disease and usually follows an indolent course.

Morphology

The tumour cells characteristically infiltrate the splenic marginal zone and, on the cytological level, resemble extranodal MALT-type MZBL cells. In most cases, the central parts of the splenic white pulp nodules are composed of small lymphoid cells with dark nuclei, and merge with surrounding, slightly larger, sometimes 'monocytoid-like' cells with a frequently pale and broader cytoplasm. Transformed blasts may be interspersed, and the splenic red pulp is involved to varying degrees. In splenic hilar lymph nodes, sinuses are generally preserved and distended, and the infiltrate surrounds germinal centres. About one third of cases may present with a leukaemic course and the appearance of 'villous' lymphocytes in the peripheral blood. Bone marrow infiltrates characteristically are found in an intertrabecular localization and sometimes sinusoidal infiltration may be encountered (**Figure 5**).

Immunophenotype

The tumor cells express pan B cell antigens and IgM, but are CD5−, CD10− and CD23−. In contrast to MALT-type MZBL, they are, in most cases, IgD-positive. Plasmacytic differentiation, as in other types of MZBL, may occur.

Genetic Features

On the cytogenetic level, splenic MZBL are consistently t(11;18)-negative, and some of the cases may show trisomy 3 and structural aberrations in 7q22–32 and 10q22–24. In fact, the apparent cytogenetic diversity of this neoplasm has elicited speculations on a different biological background of tumours designated as splenic MZBL.

Nodal MZBL

Clinical Features

Nodal marginal zone B cell lymphomas, by definition, are malignant lymphomas primarily manifesting in lymph nodes without extranodal or splenic involvement. It has been shown, however, that by careful clinical staging procedures, an occult extranodal manifestation may be present in a fraction of cases. Many patients present with advanced disease stages, but the disease is rather less aggressive.

Morphology

The small- to medium-sized tumour cells may resemble monocytoid B cells or centrocytoid B cells, and frequently show secretory differentiation with a plasmacytoid component. In most cases, large B cells are interspersed. In early stages, they colonize the perifollicular and interfollicular areas with remnants of germinal centres still being present (**Figure 6**). In later stages, however, these may be entirely destroyed and invaded by the tumour cells ('follicular colonization'). There may be cases in which the

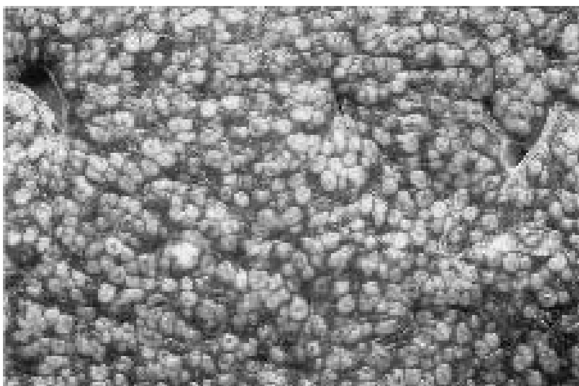


Figure 5 Splenic marginal zone B cell lymphoma. The cut surface of the spleen shows a prominence of tiny nodules representing infiltrated white pulp structures. Note that there is no merging of nodules in contrast to B-CLL splenic infiltrations.

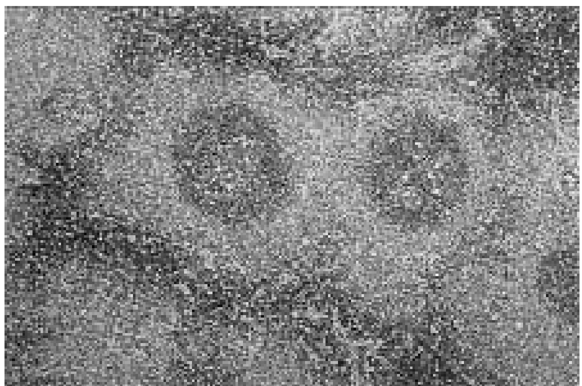


Figure 6 Nodal marginal zone B cell lymphoma. Preserved follicles are surrounded by neoplastic cells invading and broadening the perifollicular marginal zone.

tumour cells are entirely blastic in nature but with a preserved marginal zone growth pattern; we tend to classify these cases as ‘aggressive’ MZBL rather than diffuse large B cell lymphomas. Recently, two different types of nodal MZBL have been described, a ‘MALT’ type and a ‘splenic’ type. Apart from subtle morphological differences, splenic-type MZBL, as a rule, express IgD, whereas MALT-type MZBL are IgD-negative. It should be noted, however, that the designations of MALT and splenic types of nodal MZBL, at present, are only descriptive terms and it is not expected that these tumours represent secondary lymph node involvement by a primary extranodal or primary splenic lymphoma. While no splenic involvement has been found in the splenic-type nodal MZBL, around 50% of nodal MZBL designated as MALT type have been shown to be associated with an extranodal component, thus illustrating the present difficulties in discriminating true primary nodal MZBL from secondary lymph node involvement by a MALT-type lymphoma. On the other hand, characteristic clinical features have been described differentiating those tumours.

Immunophenotype

Most nodal MZBL are comparable to MALT-type lymphomas; however, some express IgD.

Genetic Features

Only single cases of presumably true nodal MZBL have been characterized on the cytogenetic level, and no unifying aberrations are known.

Lymphoplasmacytic Lymphoma (LPL)

Clinical Features

LPL is a disease of older adults, with a male predominance. Lymphoplasmacytic lymphomas are neoplasms only rarely manifesting as nodal tumours. They most commonly present with bone marrow and splenic infiltrations clinically corresponding to most cases of Waldenstrom macroglobulinaemia. Monoclonal serum paraproteins ($>3\ \mathrm{g\ dL^{-1}}$), autoimmune phenomena, cryoglobulinaemia, and hyperviscosity syndromes are frequent.

Morphology

This type of malignant lymphoma, by definition, lacks features of B-CLL, mantle cell, follicular, or marginal zone lymphomas. In the lymph node, there is a diffuse effacement of the normal architecture, sometimes however sparing preserved ‘naked’ germinal centres. The infiltrate consists of a mixture of small lymphocytes, plasmacytoid cells and mature plasma cells, and the demonstration of a monoclonal cytoplasmatic light chain is a *sine qua non* for the disease. Intranuclear immunoglobulin inclusions (Dutcher bodies) are frequently observed. In some cases, sinuses are preserved and even wide. In rare cases, a prominent epithelioid cell reaction may be seen.

Immunophenotype

These lymphomas are IgM (rarely IgG and IgA) cytoplasmic positive and IgD and CD5 negative. Some may express CD23. CD10 is always negative.

Genetic Features

A specific chromosome translocation, t(9;14)(p13;q32), has been demonstrated in 50% of cases, leading to the deregulation and over-expression of the *PAX 5* gene.

Diffuse Large B Cell Lymphomas (DLBL)

Clinical Features

Diffuse large B cell lymphomas constitute around 30–40% of adult non-Hodgkin lymphoma. Typically, patients present with a single, rapidly enlarging symptomatic mass at a single nodal or extranodal site. About one-third of DLBL are primary extranodal in origin, while the majority of cases arise in lymph nodes or other lymphatic organs, such as the tonsils or spleen.

Morphology

DLBL are principally composed of ‘large’ lymphoid cells, usually with nuclei at least twice the size of a small lymphocyte. The cytoplasm is moderately to deeply basophilic, and nuclei are large with a vesicular chromatin structure and prominent nucleoli. As can be seen from **Table 5**, diffuse large B cell lymphomas, in the WHO classification, are subdivided into distinct clinical subtypes and variants.

DLBL Variants

The **centroblastic** variant is characterized by a tumour cell population that may be either predominantly composed of cells with round to oval nuclei, a vesicular chromatin structure and membrane-bound nucleoli, or by a mixture of these centroblasts with immunoblasts. By convention, in the centroblastic variant, as much as 90% immunoblasts may be present (**Figure 7a**).

Table 5 Subtypes and variants in diffuse large B cell lymphomas

Variants:	Centroblastic
	Immunoblastic
	T cell- or histiocyte-rich
	Anaplastic large B cell
Subtypes:	Mediastinal (thymic) large B cell lymphoma
	Intravascular large B cell lymphoma
	Primary effusion lymphoma

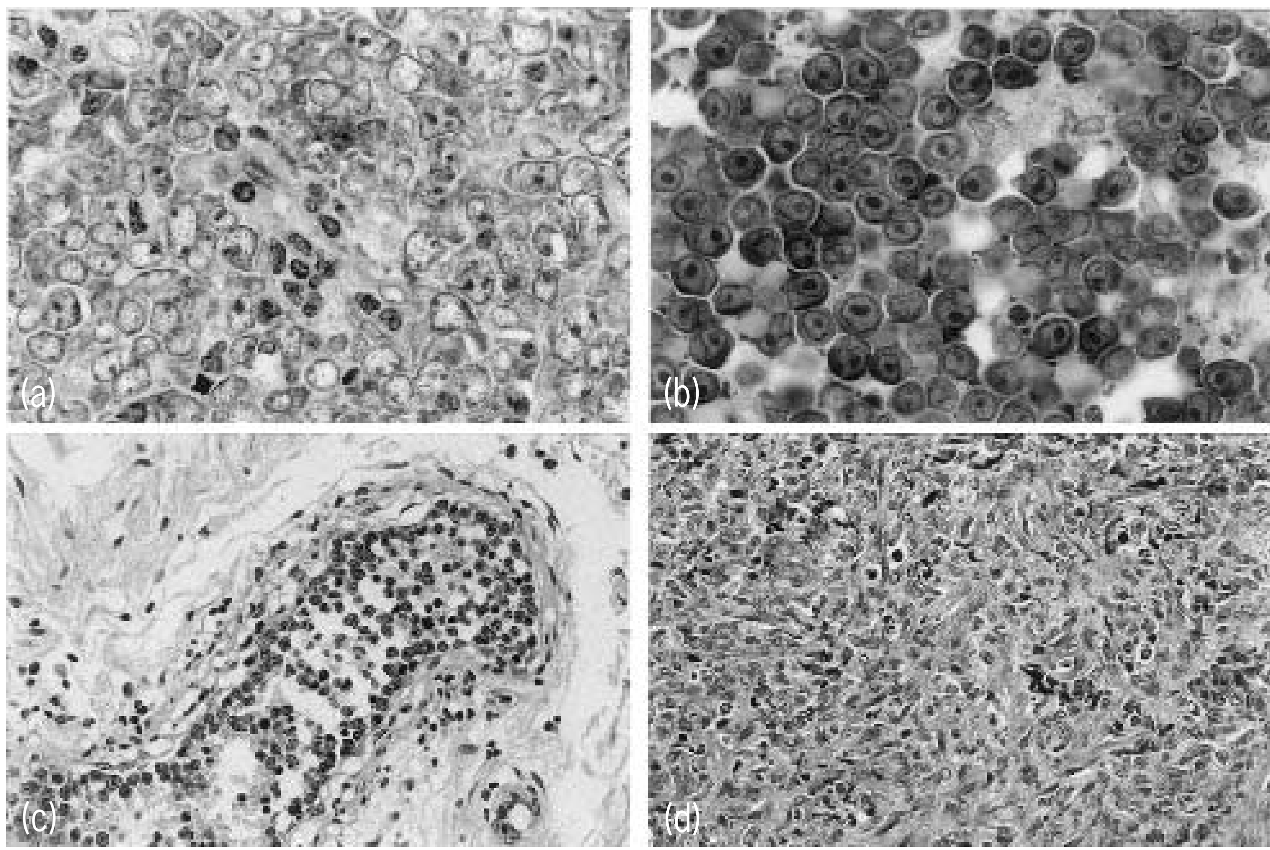


Figure 7 Diffuse large B cell lymphoma. (a) Centroblastic variant. Tumour cells are medium- to large-sized and possess a narrow rim of moderately basophilic cytoplasm. The nuclei are round to oval. Within the vesicular nuclear chromatin, 1–3 nucleoli are attached to the nuclear membrane. (b) Immunoblastic variant. Tumour cells are large with abundant cytoplasm, vesicular nuclei and single prominent central nucleoli. (c) DLBL, intravascular subtype. The blastic tumour cells are seen in the lumina of medium-sized vessels. (d) DLBL, mediastinal (thymic) large-cell subtype. Tumour cells of large size are embedded in a sclerosing stroma.

In the **immunoblastic** variant of DLBL, typical centroblasts comprise less than 10% of the total cellular infiltrate. The most cells are large with a central prominent nucleolus and a broad basophilic cytoplasm. Plasmablastic differentiation is common (**Figure 7b**).

The **anaplastic** variant comprises tumours with large, 'anaplastic,' oval or polygonal cells with pleomorphic, sometimes bizarre nuclei, sometimes reminiscent of Hodgkin or even Reed–Sternberg cells.

T cell- or **histiocyte-rich** large B cell lymphomas are characterized by a prominent reactive inflammatory infiltrate, comprising more than 90% of the total cell population.

Immunophenotype

Diffuse large B cell lymphomas, in all variants, consistently express B cell-associated antigens (CD20 and CD79a). Other B cell antigens are variably present. CD5 and CD23 immunoreactivity may be found in around 10% of cases, sometimes pointing to an underlying small lymphocytic lymphoma/B-CLL from which the diffuse large B cell lymphoma had evolved. CD10 expression is more commonly found in centroblastic variants (30%). CD30 is expressed in the anaplastic variant, but may also be demonstrated in centroblastic or immunoblastic types.

Genetic Features

Diffuse large B-cell lymphomas have their IgH and IgL chain genes clonally rearranged in a high percentage of cases and show a high load of somatic mutations. DLBL, therefore, in most cases, represent germinal or post-germinal neoplasias, having undergone germinal centre passage (and possibly antigen selection).

Chromosomal aberrations involving chromosomal band 3q27, the localization of the BCL-6 gene, seem to constitute the most frequent recurring genetic alterations in DLBL. In 20–30% of DLBL, the t(14;18) is found, and these tumours may have evolved from an (occult) follicular lymphoma. Little is known on the significance of the so-called secondary chromosomal aberrations in DLBL, although these are definitely non-random and may even be different in different variants. The most common secondary aberrations

involve deletions in the long arm of chromosome 6 and complete or partial trisomies for chromosomes 3, 5, 7, 11, 12, 18 and X. An exciting new finding in DLBL is the demonstration of different RNA expression profiles, delineating a 'germinal centre-like' and an 'activated B cell-like' subtype and, hence, possibly pointing to different transformation pathways. The recognition of these different pathways may even have profound prognostic importance.

DLBL subtypes

Primary mediastinal (thymic) B cell lymphoma is a tumour possibly derived from thymic B cells. The tumours commonly arise in the anterior mediastinum of younger, often female patients and frequently are characterized on the cytological level by large, clear cells. Prominent sclerosis is a characteristic structural feature.

Apart from the expression of B cell-associated antigens, these tumours are negative for immunoglobulins and may express CD30. The notion of a distinct biological entity has been further substantiated by the finding of characteristic chromosomal aberrations, namely a gain of material in the short arm of chromosome 9 and an amplification of the *REL* oncogene on chromosome 2p (**Figure 7d**).

Intravascular (angiotropic) B cell lymphoma is a rare disease characterized by the almost exclusive infiltration of small- to medium-sized blood vessels by large B blasts. Lymph nodes are rarely involved, and the diagnosis is commonly rendered in organ biopsies such as from the skin, the CNS or the renal parenchyma (**Figure 7c**).

Primary effusion lymphoma represents a rare body cavity-associated lymphoma and is constantly found in association with HIV and HHV8 (human herpes virus type 8) infection.

There are other characteristic clinico-pathological manifestations of diffuse large B cell lymphomas, at present regarded as provisional entities. Recognition as distinct entities is still pending and will require additional evidence for biological significance. Among those are ***AIDS-associated primary CNS lymphomas***, which are consistently associated with Epstein–Barr virus (EBV) infection, ***plasmablastic lymphomas***, frequently presenting in the oral cavity and, as a rule, evolving in the setting of HIV infection as well, and ***primary extranodal DLBL***, e.g. primary gastric or intestinal lymphomas, or primary cutaneous DLBL.

It should be pointed out, however, that neither reliable pathological or biological criteria for the subclassification of DLBL nor distinctive therapies recommended for clinical practice are available at present. For these reasons, the Clinical Advisory Committee of the WHO agreed that the subclassification of DLBL should be optional at present. However, it was agreed that the site of involvement should be clearly stated in the pathology report.

Burkitt Lymphoma (BL)

There is an endemic form (EBV-associated African Burkitt lymphoma), a so-called sporadic form occurring in Western countries, and tumours that occur in the setting of HIV infection. (See the chapters *Human DNA Tumour Viruses* and *RNA Viruses*.)

Clinical Features

Burkitt lymphoma is most common in children (one-third of paediatric lymphomas in Western countries). In African cases, the jaws and other facial bones are frequently involved. In non-endemic cases, BL tends to occur in extranodal sites, most commonly in the distal ileum, caecum and/or mesentery. In rare cases, the tumour presents as acute leukaemia. BL is a specific type of HIV-associated non-Hodgkin lymphoma.

Morphology

The tumour cells of Burkitt lymphoma are medium-sized and are characteristically arranged in cohesive sheets. They possess a deeply basophilic cytoplasm and round, monomorphic nuclei with several, usually prominent, nucleoli. Large macrophages with a broad cytoplasm and ingested apoptotic tumour cells ('tingible body macrophages') are regularly present, thus imparting a characteristic 'starry sky' appearance of the tumour. The proliferation index as measured by the Ki67 antibody is very high (>90–95% of cells) (**Figure 8**).

Immunophenotype

The tumour cells express B cell-associated antigens and are characteristically CD10+, BCL-6+ and BCL-2−.

Genetic Features

The cytogenetic hallmark of Burkitt lymphoma is the translocation t(8;14)(q24.1;q32) or one of the related translocations t(2;8) or t(8;22). On the molecular level, these translocations invariably involve the c-*myc* locus in chromosome band 8q24.1, but different breakpoints in the IgH or IgL chain genes. In endemic (African) cases, the breakpoint in chromosome 14 involves the IgH joining region in contrast to non-endemic (Western) cases in which the IgH switch region is targetted.

The WHO Steering Committee, after thorough discussion, decided to remove the tumours that in the REAL classification proposal were termed 'high-grade B cell lymphoma, Burkitt-like' (and regarded as a provisional entity) from the category of diffuse large B cell lymphomas and include those as variants under the general term of (atypical) Burkitt lymphoma. These variants commonly show, in contrast to classical Burkitt lymphoma, some unusual features either related to cell size (slightly larger), or to immunohistochemical features (cytoplasmic Ig

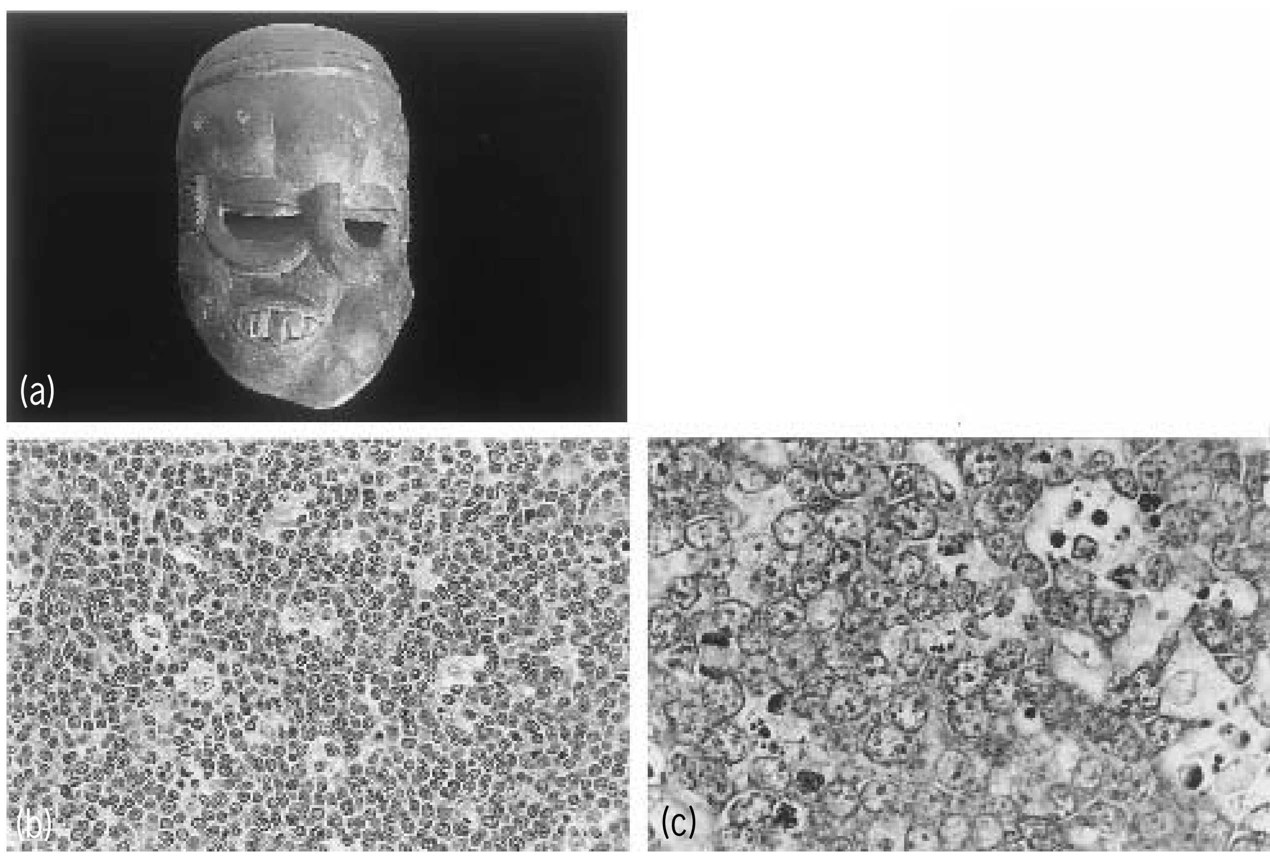


Figure 8 Burkitt lymphoma. (a) Gangoza mask from equatorial Africa (courtesy of Prof. Seeliger, Würzburg, Germany) illustrating the characteristic infiltration of the jaw bones and ensuing displacement of midfacial structures in endemic (African) cases. (b) Low-power view of a lymph node showing cohesive sheets of medium-sized blasts with interspersed macrophages imparting a starry sky pattern. (c) High-power view illustrating medium-sized blasts with round nuclei, coarsely reticulated chromatin and small nucleoli.

expression). Of note, extranodal 'Burkitt-like' lymphomas nevertheless frequently are t(8;14) positive and in the majority of cases are CD10+ and BCL2−. Therefore, these features, together with an exceedingly high proliferation index, should be considered when diagnosing 'Burkitt-like' lymphoma. The term for these tumours has now been changed to 'atypical Burkitt lymphoma'.

B Cell Lymphomas Rarely Infiltrating Lymph Nodes

Plasmacytoma rarely manifests in lymph nodes. In these cases, nodal infiltrates represent a secondary phenomenon in generalized multiple myeloma. Typically, the neoplastic plasma cells are diffusely infiltrating the lymphatic parenchyma without a lymphoid or lymphoplasmacytoid cell component. On immunohistochemistry, plasmacytomas are SIg−, CIg+ (G, A, rarely IgD or E or light chain only), CD20-negative and CD79a-positive in 50% of cases. VS38c and CD138 (syndecan) are expressed in the cytoplasm. Recently, the MUM-1 protein has been described as consistently expressed in the nucleus. Plasmacytomas represent postfollicular neoplasms with somatic hypermutations of IgH and IgL genes, but without ongoing mutations. Some specific cytogenetic aberrations have been reported, such as the t(4;14)(q16;q32) involving *FGFR3*, the t(6;14)(p25;q32) involving the *MUM-1/IRF 4* gene locus and t(11;14)(q13;q32) deregulating cyclin D1.

Hairy cell leukaemia is diagnosed only rarely in lymph nodes. Tumour cells predominantly infiltrate the spleen and bone marrow, resulting in splenomegaly and pancytopenia. In the peripheral blood, circulating tumour cells may be present with circumferential 'hairy' projections. Splenic involvement occurs in the red pulp. In lymph nodes, this neoplasm may primarily invade the subcapsular and intertrabecular sinuses. On the cytological level small- to medium-sized monomorphic lymphoid cells with a rather broad, pale cytoplasm and dense, bean-shaped nuclei are seen. The characteristic immunophenotype, CD5−, CD23−, CD11c+, CD25+, CD103+, DBA44+, along with its unique clinical features (prominent splenomegaly and infiltration of the bone marrow), helps in the diagnosis. No consistent cytogenetic alterations have been described so far.

LYMPHOMAS OF T CELL LINEAGE

Precursor T Cell Lymphoblastic Leukaemia/Lymphoma (Precursor T Cell Acute Lymphoblastic Leukaemia)

Clinical Features

Patients are adolescents or young adults, with male predominance. T-LBL accounts for roughly 40% of childhood lymphomas. T-LBL constitute 20% of precursor cell acute lymphatic leukaemias. Clinically, patients may present with large thymic tumours or peripheral lymphadenopathy. The disease is aggressive, but potentially curable.

Morphology

The morphology of the tumour cells is virtually identical with that of precursor B-LBL (**Figure 9**).

Immunophenotype

Most cases will express T cell-associated antigens CD3 (cytoplasmic) and CD7. CD4 and CD8 may be double-positive or double-negative. Sometimes, expression of NK cell antigens can be noted. Typically, TdT is positive.

Genetic Features

TCR genes are clonally rearranged. A number of recurring cytogenetic aberrations have been described, frequently involving 14q11 and 7p15 or 7q34-36 (the chromosomal bands in which the T cell receptor (TCR) genes are located); 25% of tumours have been reported to carry *SCL/TAL-1* rearrangements.

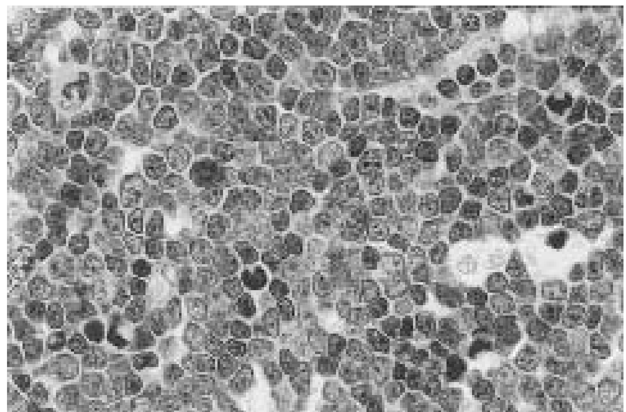


Figure 9 Precursor T cell lymphoblastic leukaemia/lymphoma. The tumour cells have round to irregular nuclei, a faint dust-like chromatin and scant cytoplasm. Note some interspersed large macrophages.

Peripheral T and T/NK Cell Neoplasias

Peripheral T and T/NK cell lymphomas are rare neoplasms in Western countries, but, they are more frequent in Oriental and Asian countries. Although the Clinical Advisory Committee of the WHO classification did not recommend a clinical grouping of the various entities of NHL, peripheral T and T/NK neoplasias may be subdivided according to their predominant clinical features into those that are primarily leukaemic or disseminated, primary nodal or primary extranodal in origin (Chan, 1999).

PREDOMINANTLY LEUKAEMIC (DISSEMINATED) PERIPHERAL T AND T/NK CELL NEOPLASMS

T Cell Prolymphocytic Leukaemia (T-PLL)

Clinical Features

T-PLL is a disease of older adults comprising only 1% of CLL. Patients present with marked leukocytosis ($100\times10^9/1$) and bone marrow, spleen, liver and lymph node infiltrates. Occasionally, cutaneous or mucosal involvement is seen. The disease is more aggressive than B-CLL.

Morphology

T-PLL diffusely infiltrates the paracortical lymph node areas, sometimes sparing pre-existing follicles. In contrast to B-CLL, no pseudo-follicular structures are present. Cytology is characterized by small- to medium-sized lymphocytes with scant cytoplasm and sometimes marked nuclear irregularity. One of the most powerful diagnostic hallmarks of T-PLL is the presence of numerous small high-endothelial venules. They often contain tumour cells, within both the lumens and vessel walls, so that the vessel walls appear to be transmigrated by the tumour (**Figure 10**). Splenomegaly is due to the infiltration of the splenic red pulp.

Immunophenotype

T-PLL cells display a mature T cell immunophenotype with positivity for CD2, CD3, CD5 and CD7. Most cases are CD4+.

Genetic Features

The *TCR* genes are clonally rearranged. In around 80% of cases, a characteristic chromosomal aberration, inv(14)(q11;q32), is present involving the *TCL-1* oncogene in 14q32. Frequently, there is also a partial trisomy for the long arm of chromosome 8. Deletion of the *ATM* gene has been observed in a considerable portion of cases.

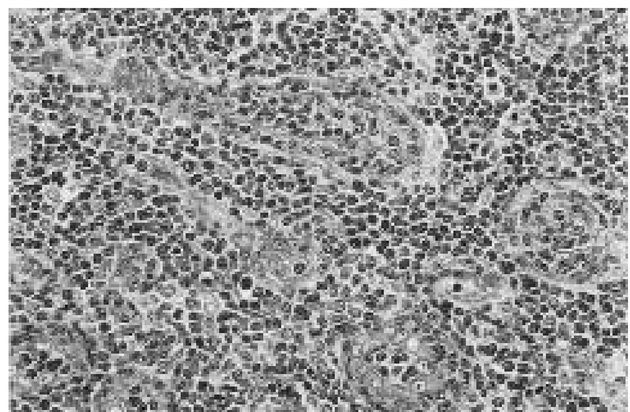


Figure 10 T cell prolymphocytic leukaemia. A characteristic hallmark of the disease is the transmigration of neoplastic cells through the walls of numerous high-endothelial postcapillary vessels.

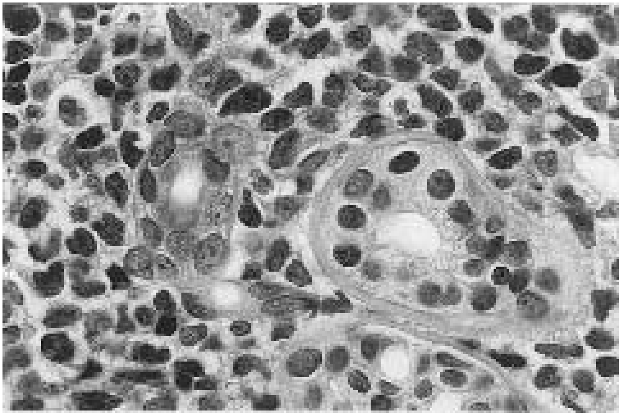


Figure 11 Aggressive NK cell leukaemia/lymphoma. Medium-sized blastic cells with irregular nuclear contours surround and invade adnexal structures in the skin. Note finely dispersed nuclear chromatin.

T Cell Large Granular Lymphocytic Leukaemia (T-LGL)

Clinical Features

T-LGL is a disease of adults frequently presenting with rheumatoid diseases and PB neutropenia and anaemia. Usually, there is mild leukocytosis ($20\times10^9/1$) and mild to moderate splenomegaly. Although the clinical course of T-LGL is usually indolent, more aggressive types may be found with an NK cell phenotype. T cell types are usually indolent, with morbidity related to cytopenias.

Morphology

T-LGL only rarely involves lymph nodes. Generally, there is infiltration of the bone marrow and leukaemic dissemination. Splenic red pulp and hepatic sinus infiltrations are common. In peripheral blood smears, T-LGL cells usually are small and characteristically show a wide rim of eccentric pale blue cytoplasm and the presence of azurophilic granules. The nuclei are round to oval, sometimes with nucleoli.

Immunophenotype

Most frequently, there is a T cell-associated phenotype with CD2+, CD3+, CD7+, CD16−/+, CD56−, CD57+ marker constellation. In rarer cases, an NK cell phenotype may be present with CD2+, CD3-, CD16+, CD56+, CD57+/−.

Genetic Features

In most of the cases with T cell phenotype, clonal rearrangements of the TCR genes may be found. NK cell cases usually are germline. There may be an association with Epstein–Barr virus in Asian cases.

Aggressive NK Cell Leukaemia/Lymphoma

Clinical Features

This is an exceedingly rare neoplasm in Western countries, while most of the cases have been described in Orientals. There apparently is some overlap with nasal NK/T cell lymphoma, however, there is infiltration of the bone marrow and peripheral blood lymphocytosis.

Morphology

The cytological picture has been described as composed of monotonous-appearing medium-sized lymphoblastoid cells with small cytoplasm, round to oval vesicular nuclei and inconspicuous nucleoli. Numerous mitotic figures are present (**Figure 11**).

Immunophenotype

Tumour cells are usually CD2+, CD3−, CD3ε+, CD56+, TIA-1+, granzyme B+.

Genetic Features

No clonal rearrangements of the TCR genes are present. As a rule, there is an association with Epstein–Barr virus infection.

Adult T Cell Leukaemia/Lymphoma (ATL/L)

Clinical Features

ATL/L is a very rare neoplasm in Western countries, and most cases have been reported in patients from south western Japan and the Caribbean. Patients are HTLV1

retrovirus positive. Several clinical variants have been described with differing clinical presentations such as an 'acute', 'lymphomatous', 'chronic' and 'smoldering' form. The acute form with high peripheral blood count, hypercalcaemia, lytic bone lesions and hepatosplenomegaly is most common.

Morphology

Histology may vary with respect to the cytological composition of the infiltrate. In most cases, a mixture of small and large cells, sometimes pleomorphic in shape, can be seen. Also, multinucleated giant cells may be present, reminiscent of Reed–Sternberg cells.

Immunophenotype

The tumour cells usually express T cell-associated antigens CD2, CD3 and CD5, but characteristically lack CD7, but are CD25+. A CD4+ phenotype is much more common than CD7+ types.

Genetic Features

The prerequisite for the diagnosis of ATL/L is the demonstration of integrated HTLV-1 genomes in virtually all cases. TCR genes are clonally rearranged. Some recurring structural and numerical chromosome aberrations have been reported, among them trisomies 3 and 5 and structural alterations involving chromosomal band 14q11.

Hepatosplenic γ/λ T Cell Lymphoma

Clinical Features

This primary splenic T cell lymphoma is a rare, albeit distinct, form of primary splenic lymphoma with cytotoxic features. The disease rarely affects peripheral lymph nodes, but may be suspected in leukaemic cases owing to its classical immunophenotye. The spleen shows moderate to distinct enlargement.

Morphology

In the spleen, there is a diffuse infiltration of the red pulp with sinus involvement by small- to medium-sized cells with sometimes more abundant pale cytoplasm and usually small round to ovoid nuclei. The white pulp is atrophic and sometime entirely absent. In the bone marrow, the infiltrate may manifest purely intrasinusoidal, a feature which in fact is a diagnostic hallmark of the disease in trephine specimens.

Immunophenotype

The tumour cells are CD3+ and also other T cell-associated antigens positive (CD2, CD5, CD7); however, they do not express either CD4 or CD8. No expression of TCRα/β receptor is present, but TCRγ/λ markers are expressed. Most characteristically, the tumour cells express both cytolytic/cytotoxic proteins, such as TIA-1, perforin and granzyme B, as well as NK cell-associated antigen CD56.

Genetic Features

The TCRα/β chain genes are germ-line, but the TCRγ and -λ chain genes are clonally rearranged. A characteristic chromosomal abnormality, i(7)(q10), has been recognized, sometimes associated with trisomy 8. Interestingly, rare TCRα/β variants of the disease have been described with identical chromosome aberrations.

Primary Nodal T/NK Cell Lymphomas: Angioimmunoblastic T Cell Lymphoma (AIL-T)

Clinical Features

AIL-T is a T cell lymphoma that can be suspected by clinicians because of its characteristic features. Typically, there is a generalized lymphadenopathy, fever, weight loss, localized or generalized erythema and polyclonal hypergammaglobulinaemia. Interestingly, occasional spontaneous remissions have been reported. Disease course is aggressive. Infectious complications are common, and progression to aggressive lymphomas of T cell or, rarely, B cell type may occur.

Morphology

In most cases, the nodal architecture is diffusely effaced, although rare cases have been described in which secondary follicles are still present, presumably representing early stages of the disease. The subcapsular sinuses may be preserved. The lymphoid infiltrate usually is composed of a mixture of small- and medium-sized lymphocytes, some immunoblasts and in some cases there may be an admixture of clear cells with pale or clear cytoplasm and pleomorphic nuclei. Epithelioid histiocytes, plasma cells and sometimes large numbers of eosinophils may be interspersed. Characteristically, a proliferation of small arborizing high endothelial venules with PAS-positive vessel walls is present throughout the lymph node, and also extending beyond the capsule (**Figure 12**). Sometimes, in routine H&E sections, expanded aggregates of follicular dendritic cells may be seen, in some cases forming characteristic onion-shaped 'burned-out'germinal centre nodules.

Immunophenotype

The tumour cells express T cell-associated antigens and usually are CD4+. Expanded and loosely structured

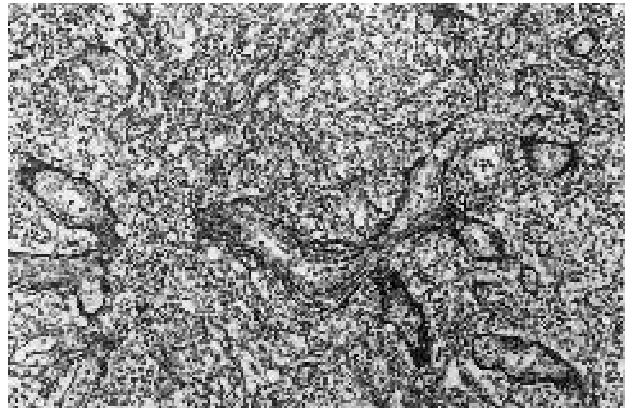


Figure 12 Angioimmunoblastic T cell lymphoma. The lymph node architecture is diffusely effaced. Note conspicuous proliferation of arborizing high-endothelial vessels.

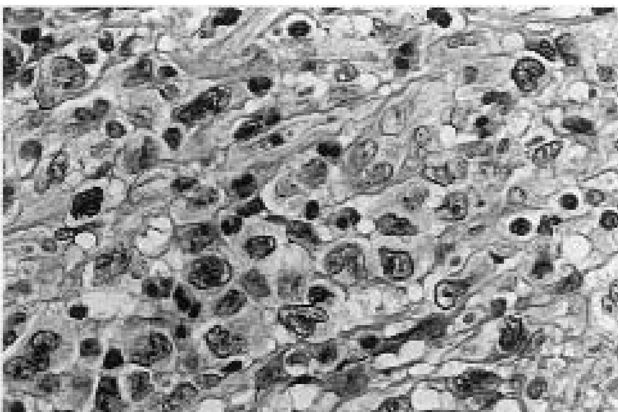


Figure 13 Peripheral T cell lymphoma (unspecified). There is a predominance of medium- to large-sized cells with moderate to broad cytoplasm and irregular nuclei varying in size and shape.

follicular dendritic cell clusters are recognized on staining with FDC markers such as CD21 or CD23. In rare cases, a prominent admixture of large atypical CD20+ B blasts may be present either haphazardly distributed in the infiltrate, or forming small to larger cell clusters. These cells, which often are reminiscent of Hodgkin and Reed–Sternberg cells, may either be CD30+, CD15+ or be CD20+, CD30+, and characteristically express LMP-1 due to Epstein–Barr virus infection.

Genetic Features

TCR genes are rearranged in most cases. Paradoxically, IgH rearrangements may be detected in 20–30% of cases. By *in situ* hybridization, EBV early repeat (EBER) transcripts have been noted, in varying numbers, in >95% of cases. Recurring chromosomal aberrations in AIL-T are trisomies 3 and 5 and, less frequently, an additional X chromosome.

Peripheral T Cell Lymphoma, Unspecified (PTCL-NOS)

This category comprises a large group of peripheral T cell neoplasms that cannot be subcategorized under one of the distinct, nodal or extranodal, entities. It comprises cases which in the Kiel classification have been described as T zone lymphoma, lymphoepithelioid (Lennert's) lymphoma, pleomorphic T cell lymphoma and T immunoblastic lymphoma. At present, it is not clear if this category contains several biological disease entities. The striking variability of the morphological picture, however, suggests that this could be in fact the case.

Clinical Features

PTCL-NOS is more frequent in advanced age groups. The disease is more common in Eastern countries, and in Western countries it comprise only 10–15% of all lymphomas. Aside from lymph nodes, the disease may frequently also involve the skin, liver, spleen and bone marrow. The disease is aggressive, but potentially curable.

Morphology

In most cases, the lymph node architecture is completely destroyed. In rare cases, however, an interfollicular infiltration pattern may be present with sparing of B cell follicles. Cytologically, peripheral T cell lymphomas typically are characterized by a mixture of small and large atypical cells. The appearance of the cytoplasm is variable, but in contrast to B cell lymphomas, the neoplastic cells, as a rule, have more irregular nuclei and vary considerably in size and shape. In some cases, large blasts reminiscent of Hodgkin and Reed–Sternberg cells are present (**Figure 13**). The background infiltrate may be equally variable in its cytological composition. There may be eosinophils or epithelioid histiocytes and plasma cells present. The number of high endothelial venules is not as striking as in AIL-T. In contrast to angioimmunoblastic T cell lymphoma, expanded meshworks of follicular dendritic cells usually are absent.

Immunophenotype

Peripheral T cell lymphomas are positive for T cell-associated antigens, but they may display selective loss of expression for all T cell markers, especially for CD7. Most cases are CD4+ and CD8−, but rare CD8+, CD4− peripheral T cell lymphomas may be encountered, often with an associated expression of cytotoxic molecules. Others may be CD4−, CD8−.

Genetic Features

The TCR genes are clonally rearranged in most cases. Ig genes are germ line. No unifying cytogenetic aberration

has been described in PCTL-NOS. Recurring chromosomal alterations involve chromosomes 7 and 14 (the bands in which the T cell receptor genes are localized). Recurring deletions have been described in the short arm of chromosome 1, the long arm of chromosome 6 and the short arm of chromosome 17. Gains are common in chromosomes 3, 7 and X. PTCL-NOS of the large cell type are frequently characterized by a tetraploid karyotype.

Anaplastic Large Cell Lymphoma (ALCL)

ALCLs were recognized in 1985 by their characteristic expression of the CD30 antigen. Although it soon became clear that other T cell or even B cell lymphomas may occasionally express CD30, the classical morphology of ALCL of T and null cell types justified its recognition as a distinct entity. CD30-positive B cell lymphomas with 'anaplastic' morphology have now been recognized as a variant of diffuse large B cell lymphoma.

Clinical Features

The majority of patients are children or adolescents, but a second peak in age distribution is seen in (older) adults. In the systemic form, lymph nodes and extranodal sites are involved, including the skin. The disease is moderately aggressive, but cure rates are high in the ALK+ variants.

Morphology

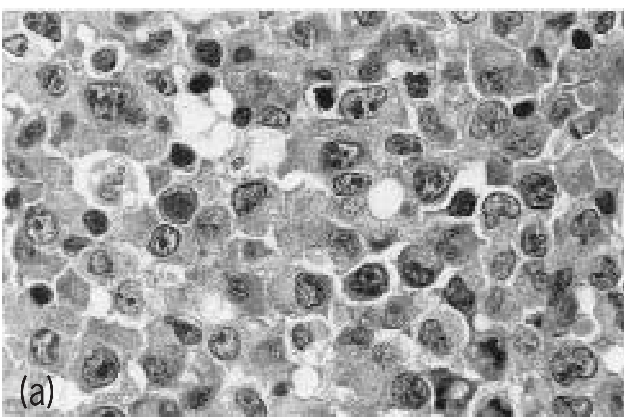
The majority of anaplastic large cell lymphomas is composed of large blastic cells with a broad eosinophilic cytoplasm on H&E sections (and grey–blue cytoplasm on Giemsa stain). The nuclei are large with a vesicular chromatin structure and a single prominent or multiple nucleoli. They may be round or, more often, horseshoe-shaped or pleomorphic. In contrast to peripheral T cell lymphomas, unspecified (which in some cases may also be CD30+), the cell size at least in the 'common' variant shows relatively little variation. The tumour cells are often arranged in cohesive sheets and infiltration of the lymph node sinuses, at first glance suggesting involvement of a solid tumour, is a key feature of the disease (**Figure 14**). In addition to this common type, a small-cell variant and a so-called lymphohistiocytic variant have been described.

In the small-cell variant, the infiltrate is mainly composed of small- to medium-sized cells with irregular nuclei and pale cytoplasm. However, typically, the characteristic large 'anaplastic' blasts are also present. In the lymphohistiocytic variant of the disease, the tumour cells comprise only a minority of the infiltrate. There are numerous histiocytes present, frequently with a broad and foamy cytoplasm, greatly obscuring the neoplastic cells. Other variants, e.g. sarcomatoid, granulocyte-rich and giant-cell types, may be encountered.

There may be an overlap with tumour cell-rich or syncytial variants of Hodgkin lymphoma (sometimes termed 'grey zone' or 'borderline' cases), in which a clear decision may be difficult. However, with the description of more and more antigens that are differentially expressed in Hodgkin lymphoma and ALCL, such as ALK-1, LMP-1, CD15 and others, in most cases a clear diagnosis can be rendered (**Table 6**).

Immunophenotype

The tumour cells, by definition, are CD30+, and in most cases also express the leukocyte common antigen. Irrespective of T or null cell origin, most of the tumours (roughly 75%) are positive for the epithelial membrane antigen (EMA) and CD25. Also, most cases of both T and null cell types express cytotoxic molecules (e.g. perforin, TIA-1 and granzyme B), pointing to the derivation of ALCL from cytotoxic cells. T cell-associated antigens may be variably present in T-ALCL. As a rule, CD20 and other B cell-associated antigens are negative.



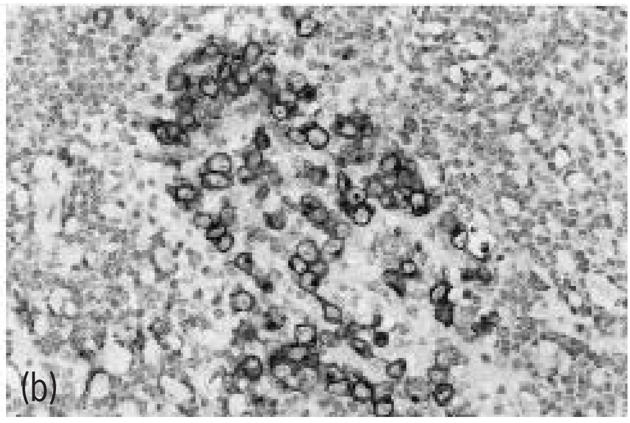


Figure 14 Anaplastic large cell lymphoma. (a) Diffuse proliferation of large cells with broad cytoplasm and round to oval nuclei and prominent nucleoli. (b) CD30 staining of the tumour cells highlighting the characteristic sinus infiltration pattern.

Table 6 Antigen expression patterns in classical Hodgkin lymphoma (cHL) and anaplastic large-cell lymphoma (ALCL)

	CD30	CD15	CD20	T cell markers	CD45	EMA	LMP	TIA1	ALK1
cHL	+	+/−	−/+	−	−	−/+	+/−	−/+	−
ALCL	+	−	−	+/−	+/−	+/−	−	+/−	+/−

Genetic Features

A reciprocal chromosome translocation, t(2;5)(p23;q35), has been detected in 40–60% of ALCL, generating a fusion protein that consists of the nucleophosmin (*NPM*) gene in 5q35 and the anaplastic lymphoma kinase (*ALK*) gene in 2p23. By virtue of the translocation, the ALK kinase is constitutively overexpressed and activated. After the generation of a monoclonal antibody directed against the kinase domain of the ALK protein, it became possible to detect the translocation even in paraffin sections. In cases with this classical translocation, the fusion protein is expressed both in the nucleus and the cytoplasm, while in rare variant translocations also involving ALK, but different fusion partners, the fusion protein is only seen in the cytoplasm. Although it was noted that ALK-positive ALCL comprise a wide morphological spectrum with common, lymphohistiocytic and small cell types being ALK-positive, it was shown that patients with ALK-positive ALCL are generally younger and predominantly of male gender. In comparison with ALK-negative ALCL, patients with tumours expressing the fusion protein show a distinctly favourable clinical course (with 10-year survival rates as high as 70–90%). Because of this important difference, ALK-positive anaplastic large-cell lymphomas are today regarded as a distinct disease entity.

PREDOMINANTLY EXTRANODAL T/NK CELL LYMPHOMAS

Primary Cutaneous Anaplastic Large Cell Lymphomas and Cutaneous CD30-positive Lymphoproliferative Disease (CD30 + LPD)

CD30-positive lymphoproliferative disorders in the skin represent a spectrum of diseases ranging from lymphomatoid papulosis to anaplastic large-cell lymphoma. Since a reliable distinction of these diseases on purely morphological grounds may be difficult, it was suggested to diagnose, in all of these cases, a primary cutaneous CD30-positive lymphoproliferative disease and to subcategorize lymphomatoid papulosis and primary cutaneous ALCL with respect to clinical features.

Morphology

In primary cutaneous CD30 + lymphoproliferative disease, the skin is infiltrated by sometimes cohesive sheets of large neoplastic cells with anaplastic morphology. In other cases, these cells are more loosely distributed, and there is a prominent reactive infiltrate of small lymphocytes and histiocytic cells rendering the differential diagnosis to lymphomatoid papulosis difficult. A single (large) tumour favours anaplastic large-cell lymphoma, while lymphomatoid papulosis tends to present with multiple, waxing and waning, small-sized lesions. By definition, primary cutaneous ALCL may only be diagnosed if the skin is the only site involved.

Immunophenotype

Primary cutaneous LPD express CD30, and frequently are positive for T cell-associated antigens. Cytotoxic markers (TIA1, perforin and/or granzyme B) are frequently found. In contrast to the systemic form, the epithelial membrane antigen is negative, and the cutaneous lymphocyte antigen HECA452 may be positive in most cases. Expression of the ALK kinase is an absolute rarity in this disorder. If ALK protein is expressed, one is more likely to be dealing with a cutaneous infiltration of primary systemic ALCL.

Genetic Features

TCR genes are clonally rearranged in most cases.

Nasal NK/T Cell Lymphoma and NK/T Cell Lymphoma of Nasal Type

Clinical Features

These tumours are rare in Western countries, but are more common in the East. They arise predominantly extranodal, in the nose, the nasopharynx and the paranasal sinuses. Because of their resemblance to nasal NK/T cell lymphoma, tumours with a similar morphology and immunophenotype in the skin, the soft tissue, the gastrointestinal tract or in other localizations, these tumours are diagnosed as ‘of nasal type’. The most frequent manifestation of the disease, however, is in the midfacial structures, hence the old term of ‘lethal midline granuloma’.

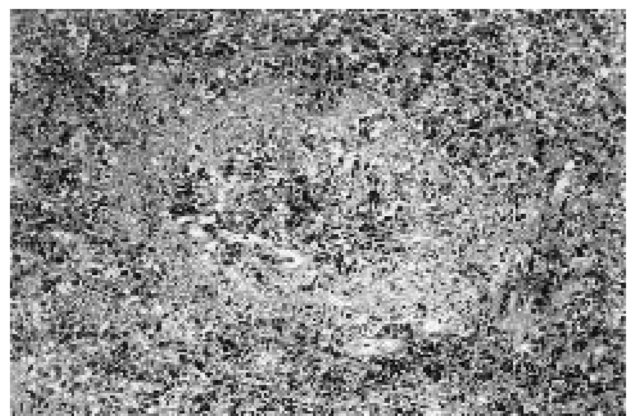


Figure 15 Nasal NK/T cell lymphoma. Tumour cells invade and destroy a larger blood vessel. Note necrosis of adjacent tissue compartments.

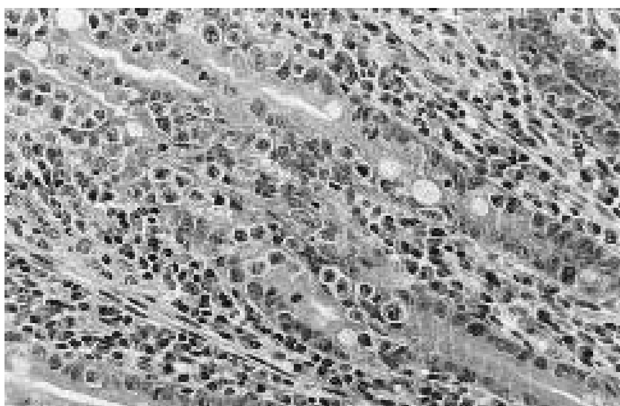


Figure 16 Enteropathy-type T cell lymphoma. The neoplastic cells invade preserved epithelial structures of the mucosa adjacent to a deeply ulcerated small intestinal tumour.

Morphology

Nasal and nasal type NK/T cell lymphomas are in many, but not all, cases characterized by a prominent angiocentric and angiodestructive growth pattern (**Figure 15**) leading to sometimes marked necrosis. The cytological spectrum may be variable with some tumours predominantly consisting of small cells, and some of medium- to large-sized cells. An admixed inflammatory infiltrate may be prominent. Frequently, the nuclei are irregular or pleomorphic, and the chromatin structure may be finely dispersed or granular. Because of their sometimes close similarity to T cell lymphomas, the immunophenotype plays a pivotal role in the diagnosis of the disorder.

Immunophenotype

Nasal NK/T cell lymphomas and NK/T cell lymphomas of nasal type invariably express the CD56 antigen and cytotoxic molecules. T cell-associated antigen CD2 is present in most cases, while CD3 is not found, but CD3ε(cytoplasmic) is expressed.

Genetic Features

Usually, no clonal rearrangements for TCR genes are found; however, in some cases they may be present pointing to a derivation of the tumour cells from T lymphocytes (hence the term NK/T cell lymphoma). In most cases, the tumour cells are infected with the Epstein–Barr virus, rendering EBER *in situ* hybridization an important diagnostic tool if the disease is suspected.

Enteropathy-Type T Cell Lymphoma

Clinical Features

Patients are adults, frequently with a history of gluten-sensitive enteropathy. In others, signs of enteropathy may only be observed in resection specimens, or be entirely absent. Disease course is aggressive. Most often, the disease manifests in the jejunal parts of the small intestine, and because of the difficult clinical diagnosis, may only be recognized if a small bowel segment is resected, frequently because of perforation. This disease was originally thought to represent a histiocytic disorder ('malignant histiocytosis of the intestine'), but has now been conclusively shown to be a T cell lymphoma.

Morphology

The tumours show a broad morphological spectrum ranging from small- and medium-sized to medium- and large-sized or anaplastic tumour cells. There may be diffuse infiltration throughout the bowel wall, and large ulcerations may be present. In the vicinity of these lesions (**Figure 16**), in preserved mucosal parts, there may be a characteristic infiltration of tumour cells into the crypt epithelium, sometimes reminiscent of lymphoepithelial lesions. The adjacent mucosa may or may not show villous atrophy, crypt hyperplasia and a high content of interepithelial T cells (which may be neoplastic as well).

Immunophenotype

T cell-associated antigens are usually expressed, as are cytotoxic molecules. A characteristic finding, which, however, is not present in all cases, is the expression of the mucosal homing receptor CD103 antigen. Some of the cases may be CD8+, and a rather monomorphic variant of small- to medium-sized tumour cells has been described with expression of CD56.

Genetic Features

The TCR genes are clonally rearranged in most cases.

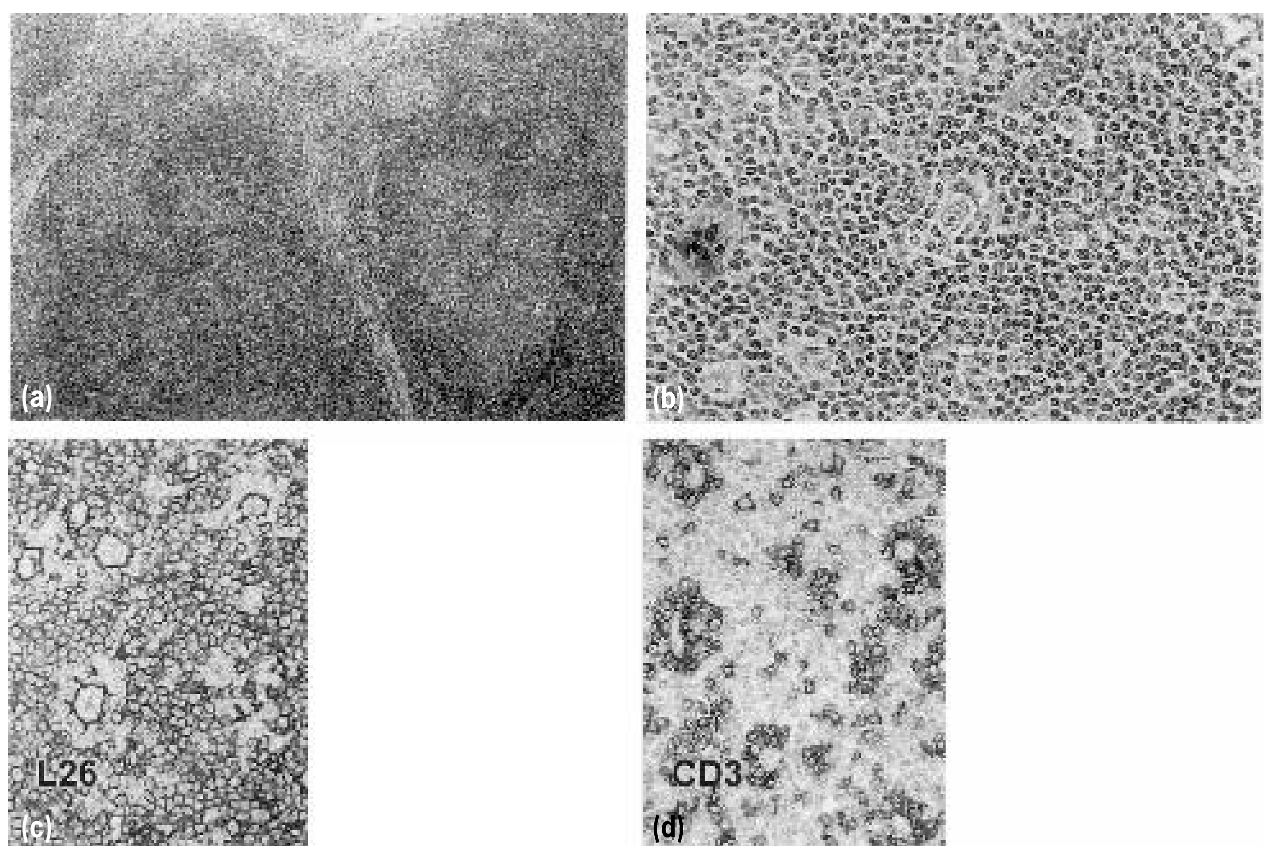


Figure 17 Nodular lymphocyte predominance Hodgkin lymphoma. (a) Low magnification showing large tumour nodules adjacent to areas of follicular hyperplasia. (b) High magnification illustrating lymphocytic and histiocytic (L&H) cells in a background of small lymphocytes. (c) L26 staining showing positivity of the L&H cells and of the background lymphoid cells for CD20. (d) CD3 staining illustrating 'rosetting' of T cells around the tumour cells.

HODGKIN LYMPHOMA

Hodgkin disease is histologically diverse and may be classified into two main entities, namely nodular lymphocyte predominance Hodgkin lymphoma (NLPHL or paragranuloma) and classical Hodgkin lymphoma.

Principally, the diagnosis of Hodgkin disease or Hodgkin lymphoma requires a malignant proliferation of Hodgkin and Reed–Sternberg cells (in the case of classical Hodgkin disease) or of the so-called lymphocytic and histiocytic cells (of NLPHD) and their variants in an 'appropriate background' of reactive, non-neoplastic bystander cells. The tumour cells as such comprise only a minority of the cellular infiltrate. The appropriate inflammatory background in the infiltrate, that is required for the diagnosis, varies widely in its structure and cytological composition and may contain variable numbers of small lymphocytes, eosinophils, histiocytes, plasma cells, neutrophils and epithelioid cells.

Nodular Lymphocyte Predominance Hodgkin's Disease (NLPHD, Paragranuloma)

Clinical Features

NLPHD occurs in all age groups. Usually, peripheral lymph nodes are involved. Often, the disease is localized (Stage I) at diagnosis. Prognosis, usually, is excellent, but transformation to DLBL may occur.

Morphology

NLPHD or paragranuloma mainly infiltrates the lymph nodes in a vaguely nodular pattern. Within this nodular proliferation, small lymphocytes, histiocytes and the characteristic tumour cells, the so-called lymphocytic and histiocytic or 'popcorn' cells, are seen (**Figure 17a and b**). The latter are large cells with a small- to medium-sized cytoplasm and large, usually folded or lobated, nuclei with

a vesicular chromatin structure and a single to several medium-sized nucleoli. In some cases, they may resemble classical Hodgkin or Reed–Sternberg cells. Sometimes, the tumour cells are surrounded by (in comparison with normal lymphocytes) slightly larger cells (expressing the CD57 antigen).

Diffuse areas may be present. When these comprise more than 30% of the lymph node, the case should be classified as nodular paragranuloma with diffuse areas. Diffuse paragranulomas comprise cases in which less than 30% of the infiltrate show a nodular pattern. Rare, purely diffuse paragranulomas are virtually indistinguishable from T cell-rich B cell lymphoma. Adjacent to the tumour infiltrate, a reactive follicular hyperplasia, sometimes with progressive transformation of the germinal centres, may be present (Mason *et al*., 1994).

Immunophenotype

The neoplastic cells in NLPHD constantly express B cell-associated antigens CD20 and CD79a and, in most cases, are positive for the J-chain. In roughly 60% of cases, the epithelial membrane antigen (EMA) is expressed; however, in contrast to classical Hodgkin disease, the tumour cells are negative for CD30, CD15 or the latent membrane protein (LMP-1) of the Epstein–Barr virus. In rare tumours, a monotypic immunoglobulin light-chain expression may be noted. In contrast to classical Hodgkin lymphoma, the small lymphocytes in the background infiltrate are also CD20-positive, and frequently coexpress IgD, like the cells of the normal perifollicular mantle zone. There is, however, a so-called 'rosetting' of CD3-positive T cells and CD57-reactive NK cells directly around the tumour cells of NLPHD (**Figure 17d**).

Genetic Features

By virtue of single-cell PCR analysis, a clonal rearrangement for IgH chain genes was found in all cases investigated. The finding of a high load of somatic mutations in the rearranged Ig genes and the presence of ongoing mutations suggests that NLPHD is derived from germinal centre B cells (Marafioti *et al*., 1997).

Classical Hodgkin Lymphoma

Clinical Features

Hodgkin lymphoma of nodular sclerosis type is most common in adolescents and young adults. The mediastinum is frequently involved. HD, mixed cellularity is a disease of adults, and more widespread disease is common involving lymph nodes, spleen, liver or bone marrow. Lymphocyte-depleted HD is the least common variant, predominantly occurring in older patients and in HIV-infected individuals. In spite of advanced disease stages, HD of all types are curable.

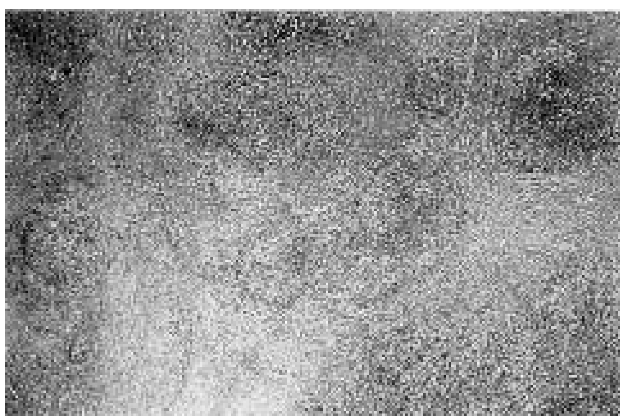


Figure 18 Classical Hodgkin lymphoma, nodular sclerosis type. Note cellular nodules surrounded by broad bands of densely packed collagen.

Morphology

Classical Hodgkin lymphoma is characterized by the presence of Hodgkin and Reed–Sternberg cells with a classical immunophenotype and an appropriate, albeit variable, background infiltrate. Several subtypes are now recognized, each presenting with a more or less unique infiltrate structure (Lukes and Butler, 1966).

Nodular sclerosing HL

Most Hodgkin lymphomas are of nodular sclerosis subtype, which is characterized by the formation of abundant collagen organized in broad bands surrounding cellular nodules (**Figure 18**). This collagen mantle shows birefringence on polarization and is PAS-positive. The thickened capsule normally is integrated into this fibrosing process, and even localized bands of fibrosis warrant the diagnosis of nodular sclerosis subtype, irrespective of the histological appearance of the remainder of the lymph node. The cellular nodules contain a sometimes wide morphological range of tumour cells, including Hodgkin, Reed–Sternberg and the so-called 'lacunar' cells. The lacuna-like appearance of these cells results from fixation artifacts. The reactive background is composed mainly of small lymphocytes and eosinophils. The so-called 'cellular phase' of nodular sclerosing Hodgkin lymphoma may be diagnosed if the capsule of the lymph node is thickened and tumour cells of lacunar type show a (frequently vague) nodular arrangement.

Classical Lymphocyte-rich Hodgkin Lymphoma, Nodular

Classical lymphocyte-rich Hodgkin lymphoma represents one end of a spectrum leading over mixed cellularity to lymphocytic depletion. In classical lymphocyte-rich Hodgkin lymphoma, the lymph-node structure is effaced and a varying degree of nodularity is present. The background infiltrate is mainly composed of lymphocytes,

while eosinophils and neutrophils are only rarely found. In the (more common) nodular form the lymphocytes are mainly B cells and, sometimes, in the centre of the nodules, regressively changed germinal centres may be found (**Figure 19**). In contrast, in the diffuse form the background infiltrate is mainly composed of T lymphocytes. The distinction from NLPHD is not always possible on histological grounds alone; however, immunohistochemistry will as a rule help in this distinction.

Mixed Cellularity

This subtype is histologically intermediate between classical lymphocyte-rich Hodgkin lymphoma and lymphocytic depletion (**Figures 20 and 21**). The diagnosis in fact is made whenever in a case of Hodgkin lymphoma the criteria for the other subtypes are not fulfilled. Partial or interfollicular involvement of the lymph node, by definition, is classified as mixed cellularity.

Figure 19 Classical Hodgkin lymphoma, lymphocyte-rich, nodular. The overview shows large tumour nodules with sometimes preserved germinal centres. Tumour cells are vaguely seen in the broadened follicular mantle zone.

Lymphocytic Depletion

Lymphocyte-depleted type Hodgkin lymphoma is rare (< 5% of all cases) and the diagnosis is often made in relapse or in patients with the acquired immunodeficiency syndrome. In lymphocytic depletion, there may be a diffuse fibrosis with the presence of disorderly orientated reticulin fibres that tend to surround individual cells. The overall cellularity is low, and diagnostic Reed–Sternberg cells may be rare. In the reticular variant, the key feature is the presence of numerous bizzare and anaplastic-appearing Hodgkin and Reed–Sternberg cells together with a depletion of small lymphocytes.

Immunohistochemistry

No significant differences in antigen expression are encountered between the subtypes of classical Hodgkin lymphoma. The tumour cells typically are positive for CD30, and in most of the cases also express CD15 and vimentin. In roughly 20% of the cases, CD20 may be positive, but usually is expressed only weakly and inconsistently in a part of the tumour cells. Other B cell-associated antigens (CD79a, J-chain) are absent. In 10–50% of cases (depending on the histological subtype), the latent membrane protein (LMP-1) of the Epstein–Barr virus is expressed. The positivity of the tumour cells for CD3 or cytotoxic proteins (TIA-1 or perforin) is a rarity. The cellular background is composed (with the exception of nodular lymphocyte-rich classical Hodgkin lymphoma) mainly of CD3-positive T cells and histiocytes. In contrast to paragranuloma, cytotoxic proteins are expressed in a significant number of the small background lymphocytes (**Table 7**).

Genetic Features

The cellular derivation of Hodgkin and Reed–Sternberg cells lymphoma remained enigmatic until, in 1994, by the use of single-cell PCR studies, both the clonal nature and

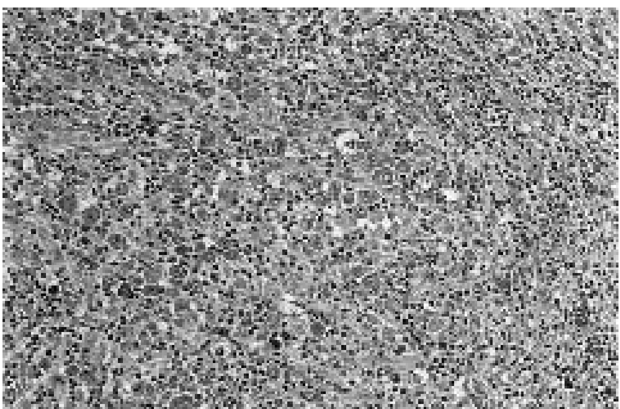


Figure 20 Classical Hodgkin lymphoma, mixed cellularity. Numerous Hodgkin and Reed–Sternberg cells are seen in a reactive background infiltrate.

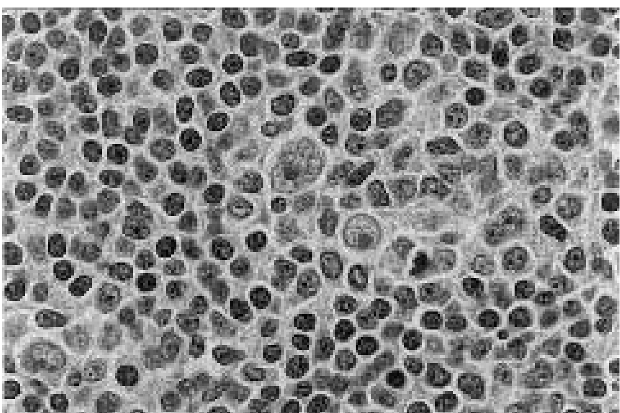


Figure 21 Classical Hodgkin lymphoma. One Hodgkin cell and one multinucleate Reed-Sternberg cell are seen embedded in a reactive background infiltrate.

Table 7 Antigen expression patterns in classical Hodgkin lymphoma (cHL), T cell-rich B cell lymphoma (TCRBCL) and nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL)

	Tumour cells						**Background**	
	CD30	**CD15**	**CD20**	**CD79a**	**J-chain**	**Vimentin**	**TIA/CD57 ratio**	**B cells**
cHL	+	+/–	Weak, uneven	–	–	+	↑	↓
TCRBCL	–	–	+	+/–	–/+	–	↑	↓↓
NLPHL	–/+	–	+	–/+	+/–	–	↓	↑

B lymphocyte origin of Reed–Sternberg cells was revealed in over 95% of cases (Küppers *et al.*, 1994). More recent results indicate that a very small proportion of cases may actually be derived from T cells. As in nodular lymphocyte predominant Hodgkin's disease, the immunoglobulin genes in classical Hodgkin lymphoma are highly mutated, giving evidence for a germinal centre origin of the tumour cells. Although Hodgkin and Reed–Sternberg cells are derived from B cells, they are not able to express immunoglobulins. It has been suggested that this perplexing phenomenon was due to so-called 'crippling' mutations in Ig genes. However, more recent results suggest that the Ig promotor region is not active, thereby preventing the transcription of the Ig genes. The expression of several antiapoptotic proteins may be the reason why the cells are able to escape from apoptosis.

The role of the Epstein–Barr virus in the pathogenesis of classical Hodgkin lymphoma still remains to be elucidated. The virus is found, with slightly varying frequencies in the subtypes recognized, in 40–60% of HL. It has been suggested that the latent membrane protein 1, that posseses transforming capacities *in vitro*, may be involved in tumorigenesis via its homologies to members of the tumour necrosis factor receptor molecules. By complexing in the cell membrane, LMP-1, therefore, is able to activate intracellular signal transduction factors ultimately leading to the activation of the nuclear factor (NF) κB. However, the presence of the EBV in the tumour cells of Hodgkin lymphoma does not necessarily imply that the virus is directly involved in the pathogensis of the neoplasm. A strong argument in favour of a pathogenetic role of EBV in HL is the monoclonal origin of the virus in the tumour cells. Large population studies showed that patients who developed Hodgkin lymphoma had abnormally high titres of some anti-EBV antibodies in prediagnostic sera, indicating that infection occurred prior to clonal expansion of the tumour and was not just due to a manifestation of immunosuppression secondary to the emergence of the neoplastic clone.

Lymphoproliferations in Immunologically Compromised Patients

Lymphoproliferative disorders in the immunosuppressed patient represent a spectrum of diseases different from sporadic lymphomas. The pathogenetic concepts involved result in problems of classification and diagnosis of these tumours and also of treatment options (Knowles, 1999).

Post-transplantation Lymphoproliferative Disorders (PTLDs)

In PTLDs, the allograft type and immunosuppressive regimen applied are most important. The incidence of PTLDs is approximately 1% and 2% for renal and hepatic transplants, respectively, whereas it varies from 2 to 10% for heart, combined heart–lung and bone-marrow transplants. These differences, however, appear to be more related to different degrees of immunosuppression than to the organ itself. With increasing organ mass and tissue type, entirely different factors become more important, such as the number of allogeneic passenger lymphocytes, frequency of allograft rejection and graft versus host disease. PTLD in conventional immunosuppression (antilymphocyte immunoglobulin and azathioprin) is about 1–5% in cardiac transplants. After the introduction of cyclosporin A, initially much higher frequencies of PTLD were observed (9–13%). More recently, as a result of dose reduction and serum level monitoring of the drug, lower frequencies in the range 1–2% are observed.

The majority of PTLDs occur shortly after transplantation, in most instances at a mean time of 6 months, within the first 2 years after transplantation. However, a small but steady number of PTLDs continue to occur throughout the following years. Whereas EBV-positive cases are seen early after transplantation, EBV-negative PTLDs and T cell lymphomas tend to occur in the later post-transplant period.

Clinical Features

There is considerable variability in clinical presentations of PTLDs. Very often the disease starts, especially in younger patients, with mononucleosis-like symptoms (tonsillar enlargement, cervical lymphadenopathy), rapidly transforming into general lymphadenopathy. About half of these patients in a survey of the University of Minnesota took a rapidly fatal course whereas in the other half the disease was self-limiting. Another type of presentation consists of a localized tumour mass occuring very often at extranodal sites, such as the gastrointestinal tract, the central nervous

system or unusual sites (liver, lung, oral cavity, skin or uterus). Involvement of the allograft itself is rare (10–15%) and in some cases was related to donor cell origin of the tumour. Donor cell origin is especially frequent in bone marrow transplantation and increased when T cell depletion of the graft had been performed in order to avoid graft-versus-host disease. Most other PTLDs are of host cell origin representing a reactivation or *de novo* infection by EBV.

PTLDs were considered in the early days as frankly malignant disease and treated (mostly unsuccessfully) by antitumoral combined cytostatic regimens. This view was challenged when spontaneous regression of tumours was reported upon reduction or discontinuation of immunosuppressive therapy. A favourable response is seen in 31% of PTLDs. Early onset lesions and polymorphous rather than monomorphous lymphoproliferations are more likely to respond. However, exact predictive diagnosis has to take into account morphological, molecular and karyotypic as well as virological data.

Pathological Classification of Post-transplant Lymphoproliferative Disorders

Early lesions of mononucleosis-like type or plasmacytic hyperplasias are terms used for PTLD occuring early after organ transplantation (mean time 3–4 months) and frequently involving adenoids, tonsils and superficial nodes, but also extranodal localizations, in children or young adults. The histological pattern is overlapping, showing plasmacytic hyperplasia without destruction of the underlying architecture of lymphoid or organ tissues. Plasma cells do not show light chain restriction. EBV can be demonstrated in most instances showing either a polyclonal, oligoclonal or monoclonal type of infection. These lesions tend to regress spontaneously or after reduction of immunosuppressive treatment, but rarely can progress to higher histological grades, accompanied by more aggressive clinical behaviour or even be fatal, as can infectious mononucleosis in other conditions.

Polymorphic PTLDs are invasive and destructive lesions leading to the effacement of the underlying normal architecture of lymphoid tissues or involved organs. In contrast to sporadic B cell lymphomas, these tumour masses consist of a mixture of cell types including the full range of B cell maturation from centroblasts to immunoblasts and plasma cells and a background inflammatory component of different T cell populations. There may be areas of necrosis, which may be prominent in some cases. The proliferative activity is high.

Immunophenotyping on paraffin sections reveals monotypic light-chain restriction and immunoglobulin secretion in some cases, but no immunoglobulin production in others. EBV may be detected in most cases, the antigen expression profile suggesting EBV latency type 2 or 3.

Polymorphic PTLDs may show partial or complete regression after discontinuation of immunosuppressive treatment or resolution with surgery, radiation therapy or chemotherapy less frequently. Other cases may progress in spite of therapy.

Most cases of **monomorphic PTLD** fulfil diagnostic criteria used for the designation of aggressive lymphomas on morphological grounds, and show a B cell phenotype. The designation of these lymphomas should follow the rules of the WHO classification, but include the term PTLD for clinical and prognostic reasons. Grossly, monomorphic PTLDs are aggressive and invasive tumours destroying the architecture of the lymph node or extranodal tissues. Most frequently these lymphomas are composed of large transformed immunoblasts or plasmablasts, with little or no differentiation toward plasma cells. The term also may be used in cases containing bizarre or multinucleated cells. Immunophenotyping reveals B cell-associated antigens in most cases. CD20 may be negative in plasmablastic lymphoma or anaplastic myeloma which, however, are rarely seen as PTLD. If immunoglobulin production is detected, light-chain restriction is present. The proliferative fraction is high. Most monomorphic PTLDs representing malignant non-Hodgkin lymphomas or multiple myelomas after organ transplantation tend to occur in a higher age group (after the age of 50 years) and do not regress after discontinuation of immunosuppressive treatment or progress despite cytostatic treatment.

EBV Infection

EBV infection is detected in 90% of PTLDs by EBER *in situ* hybridization techniques. Plasmacytic hyperplasias and early lesions of mononucleosis-like type usually show polyclonal EBV infection patterns involving increased numbers of EBV-positive small lymphocytes or activated blast cells. Polymorphic PTLD and monomorphic PTLD show monoclonal EBV integration and a positive result of EBER *in situ* hybridization in most if not all tumour cells. These cases usually also show LMP1 expression in at least some tumour cells, but in fewer cells than detected by EBER *in situ* hybridization. Half of the cases of polymorphic PTLD also express EBNA2, suggesting that these cases are latency type 3, while the others are of latency type 2. Latency type 1, lacking the expression of EBNA2 and LMP1, is seen in some of the malignant lymphomas or monomorphic PTLDs, respectively.

Clonality Studies

Using Southern blot hybridization techniques or PCR-based immunoglobulin heavy-chain CDRIII region amplification, early lesions of PTLDs usually show a polyclonal smear or, rarely, a weak clonal predominance. In polymorphic PTLD, clonality analyses usually demonstrates a monoclonal or oligoclonal result. The investigation of multiple tumour nodules within one organ (e.g.

Table 8 Molecular genetic features in PTLD

Category	IgH clonality	EBV/clonality	Oncogene/tumour-suppressor gene alterations	Clinical course
Plasma cell hyperplasia	Polyclonal	Absent Polyclonal Clonal	Absent	Non-aggressive
Polymorphic PTLD	Monoclonal	Clonal	Absent	?
Monomorphic PTLD	Monoclonal	Clonal	Yes	Aggressive

gastrointestinal tract) may show different clones in each lesion. Malignant (monomorphic) lymphoma occuring as PTLD harbour monoclonal tumour cell populations (Knowles *et al*., 1995).

Genetic Alterations

The status of proto-oncogenes and tumour suppressor genes has been intensively studied. Early lesions and polymorphic PTLD characteristically lack evidence of *BCL1*, *BCL2*, c-*MYC*, *RAS* and *p53* gene alterations, whereas monomorphic lymphomas after organ transplantation consistently contain structural alterations of one or more proto-oncogenes or tumour suppressor genes, most commonly involving *RAS*, c-*MYC* or *p53* genes. More recently, analysis of the *BCL-6* gene in these tumours revealed frequent mutations, the presence of which strongly predicted shorter survival and refractoriness to reduced immunosuppression and/or surgical excision (Cesarman *et al*., 1998). An overview on these findings is summarized in **Table 8**.

Rare types of post-transplantation lymphoproliferative disorders include EBV-negative B cell-related PTLDs, classical Hodgkin lymphoma and peripheral T cell lymphoma.

Iatrogenic Non-transplantation-induced Lymphoproliferative Disorders

Clinical Features

Iatrogenic non-transplantation-induced lymphoproliferative disorders have been recognized only very recently. The most characteristic clinical setting for the occurrence of these disorders is represented by methotrexate treatment in rheumatological diseases, especially rheumatoid arthritis. In contrast to solid organ transplant recipients, where the occurrence of PTLD is considered to be related to the kind and severity of medical immunosuppression, it is usually not known whether and to what extent a state of immunosuppression exists or is active in patients with rheumatoid arthritis receiving methotrexate therapy. Clearly, not all malignant lymphomas that occur in patients with rheumatological diseases are related to their immunomodulatory treatment.

Lymphoproliferative disorders related to immunosuppressive or immunomodulatory treatment appear to comprise only a minority of lymphomas in these patients. Most lymphomas that occur are likely coincidental with rheumatological disease and not the effect of immunosuppression. Most lymphomas in patients with rheumatoid arthritis, similar to other autoimmune diseases, represent malignant lymphomas that are EBV negative and in their histology are similar to lymphomas seen in patients without rheumatological disease. Most iatrogenic lymphoproliferative disorders are EBV positive.

Morphology

The morphological features of iatrogenic lymphoproliferative disorders reported are seen in three main categories:

Atypical polymorphous lymphoproliferative disorders occurring in lymph nodes or extranodal tissues tend to efface the organ architecture and contain a mixture of lymphoid cells at various stages of activation and maturation. These cases show clonal immunoglobulin gene rearrangement or light-chain restriction. Nevertheless, many of them regress after discontinuation of immunosuppressive treatment.

Diffuse aggressive non-Hodgkin lymphomas cytologically represent different types of diffuse large B-cell lymphomas either of large cell or of Burkitt-like type. Some of these proliferations reveal pleomorphic features and focally show Reed–Sternberg-like cells. EBV may be detected in almost all cases by EBER *in situ* hybridization in most, if not all, tumour cells.

Hodgkin lymphoma and lymphoproliferations resembling Hodgkin lymphoma show the most intriguing morphology in the group of iatrogenic lymphoproliferative disorders. The diagnosis of these tumors is difficult, since features of Hodgkin disease may be evident in one place and diffuse large B cell lymphoma in others. The immunophenotype of Hodgkin cells very often is positive for CD20 and CD30, which renders the distinction between classical Hodgkin lymphoma and lymphoproliferations resembling Hodgkin disease in this setting difficult. Hodgkin disease in iatrogenic lymphoproliferative disorders may show diagnostic features of mixed cellularity or nodular sclerosing variant of classical Hodgkin lymphoma including a typical phenotype of Hodgkin cells. The exact diagnosis requires morphological, immunophenotypic and detailed studies on EBV status, including the definition of the latency type.

'Sporadic' Atypical Lymphoproliferative Disorders

As has been pointed out, the pathological features of lymphoproliferative disorders after organ transplantation and/or immunomodulatory treatment differ from sporadic cases of malignant lymphoma. However, their morphology is by no means specific. Therefore, similar, if not identical, cases are also seen without a clinical report of the respective conditions. EBV is also found in 5–10% of aggressive non-Hodgkin lymphomas in the general population and polymorphic lymphoproliferations associated with EBV with or without Hodgkin disease-like features are definitely found outside manifest immunosuppression or immunomodulatory treatment.

Therefore, the question is still unresolved as to whether PTLD and iatrogenic lymphoproliferations represent the well-defined, quasi-experimental condition allowing one to recognize a specific morphology of immunodeficiency-related lymphomas and lymphoproliferative disorders which may also be seen in sporadic, less well-defined states of immunodeficiency or postchemotherapy immunosuppression in elderly patients predisposing to a so-far undefined type of 'sporadic' lymphoproliferative disorder.

REFERENCES

Alizadeh, A. A. *et al*. (2000). Distinct types of diffuse large cell lymphoma identified by gene expressing profiling. *Nature*, **403**, 503–511.

Armitage, J. O. and Weisenburger, D. D. (1998). New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *Journal of Clinical Oncology*, **16**, 2780–2795

Cesarman, E. *et al*. (1998). BCL-6 gene mutations in posttransplantation lymphoproliferative disorders predict response to therapy and clinical outcome. *Blood*, **92**, 2294–2302.

Chan, J. K. (1999). Peripheral T-cell and NK-cell neoplasms. An integrated approach to diagnosis. *Modern Pathology*, **12**, 177–199.

Harris, N. L. *et al*. (1999). World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting – Airlie House, Virginia, November 1997. *Journal of Clinical Oncology*, **17**, 3835–3849.

Harris, N. L. *et al*. (2000). Lymphoma classification – from controversy to consensus: the R. E.A. L. and WHO classification of lymphoid neoplasms. *Annals of Oncology*, **11**, Supplement 1, 3–10.

Harris, N. L. *et al*. (1994). A revised European–American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*, **84**, 1361–1392.

Jaffe, E. S. *et al*. (1998). World Health Organization classification of lymphomas: a work in progress. *Annals of Oncology*, **9**, 25–31.

Knowles, D. M. *et al*. (1995). Correlative morphologic and molecular genetic analysis demonstrates three distinct categories of posttransplantation lymphoproliferative disorders. *Blood*, **85**, 552–565.

Knowles, D. M. (1999). Immunodeficiency-associated lymphoproliferative disorders. *Modern Pathology*, **12**, 200–217.

Küppers, R. *et al*. (1994). Hodgkin disease: Hodgkin and Reed–Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. *Proceedings of the National Academy of Sciences of the USA*, **91**, 10962–10966.

Lennert, K. and Feller, A. C. (1992). *Histopathology of Non-Hodgkin's Lymphomas* (*Based on the Updated Kiel Classification*). (Springer, Berlin).

Lukes, R. J. and Butler, J. J. (1966). Natural history of Hodgkin's disease is related to its pathologic picture. *Cancer*, **19**, 317–344.

Marafioti, T. *et al*. (1997). Origin of nodular lymphocyte-predominant Hodgkin's disease from a clonal expansion of highly mutated germinal-center B cells. *New England Journal of Medicine*, **337**, 453–458.

Mason, D. Y. *et al*. (1994). Nodular lymphocyte predominance Hodgkin's disease. A distinct clinicopathologic entity. *American Journal of Surgery and Pathology*, **18**, 526–530.

Non-Hodgkin's Lymphoma Pathologic Classification Project (1982). National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a working formulation for clinical usage. *Cancer*, **49**, 2112–2135.

Non-Hodgkin's Classification Project (1997). A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. Blood, **89**, 3909–3918.

Rosati, S. and Frizzera, G. (1997). Pseudoneoplastic lesions of the hematolymphoid system. In: Wick, M. R., *et al*. (eds), *Pathology of Pseudoneoplastic Lesions*. 449–544 (Lippincott-Raven, Philadelphia).

Shipp, M. A. *et al*. (1993). A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *New England Journal of Medicine*, **329**, 987–994.

Stansfeld, A. G. *et al*. (1988). Updated Kiel classification for lymphomas. *Lancet*, **i**, 292–293 and 603.

FURTHER READING

Döhner, H. *et al*. (1999). Chromosome aberrations in chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. *Journal of Molecular Medicine*, **77**, 266–281.

Goossens, T. *et al*. (1998). Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proceedings of the National Academy of Sciences of the USA*, **95**, 2463–2468.

Heim, S. and Mitelman, F. (1995). *Cancer Cytogenetics*, 2nd edn. (Wales, New York).

Hockenbery, D. *et al*. (1990). Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature*, **348**, 334–336.

Mason, D. Y. *et al*. (1998). Nuclear localization of the nucleophosmin–anaplastic lymphoma kinase is not required for malignant transformation. *Cancer Research*, **58**, 1057–1062.

Müller-Hermelink, H. K. and Greiner, A. (1998). Molecular analysis of human immunoglobulin heavy chain variable genes (IgVH) in normal and malignant B cells [comment]. *American Journal of Pathology*, **153**, 1341–1346.

Pulford, K. *et al*. (1997). Detection of anaplastic lymphoma kinase (ALK) and nucleolar protein nucleophosmin (NPM)–ALK proteins in normal and neoplastic cells with the monoclonal antibody ALK1. *Blood*, **89**, 1394–1404.

Rüdiger, T. *et al*. (1998). Differential diagnosis between classic Hodgkin's lymphoma, T-cell-rich B-cell lymphoma, and paragranuloma by paraffin immunohistochemistry. *American Journal of Surgical Pathology*, **22**, 1184–1191.

Johansson, B. *et al*. (1995). Cytogenetic evolution patterns in non-Hodgkin's lymphoma. *Blood*, **86**, 3905–3914.

Chapter 47

Bones

Lawrence M. Weiss
City of Hope National Medical Center, Duarte, CA, USA

CONTENTS

- Introduction
- Normal Development and Structure
- Epidemiology and Aetiology
- Pathology of Benign Tumours
- Giant Cell Tumour of Bones
- Osteosarcoma
- Chondrosarcoma
- Ewing Sarcoma/Peripheral Neuroectodermal Tumour of Bone (PNET)
- Chordoma
- Malignant Fibrous Histiocytoma and Fibrosarcoma
- Adamantinoma
- Angiosarcoma
- Multiple Myeloma
- Malignant Lymphoma Presenting in Bone
- Langerhans Cell Histiocytosis

INTRODUCTION

Bone tumours are relatively uncommon, but comprise an interesting group of neoplasms (**Table 1**). As a whole, these tumours occur in a younger age group than most other tumours, particularly frequent at the time of maximal bone growth in the teens. Recent advances in modern chemotherapeutic regimens have significantly improved the prognosis in selected types. The chapter will summarize the pertinent pathological features of the most common types of bone tumours. Excellent specialized tests are recommended for rare entities not covered in this chapter (Dahlin and Unni, 1986; Fechner and Mills, 1993; Unni, 1996; Dorfman and Czerniak, 1998).

Table 1 Classification of common neoplasms of bone

Benign
Osteoid osteoma
Osteoblastoma
Osteochondroma
Enchondroma
Chondroblastoma
Chondromyxoid fibroma
Fibrous dysplasia
Osteofibrous dysplasia
Nonossifying fibroma
Benign fibrous histiocytoma
Haemangioma/lymphangioma/angiomatosis
Locally aggressive
Giant cell tumour
Malignant
Osteosarcoma
Chondrosarcoma
Ewing sarcoma
Chordoma
Malignant fibrous histiocytoma
Fibrosarcoma
Adamantinoma
Angiosarcoma
Multiple myeloma and plasmacytoma of bone
Malignant lymphoma
Langerhans cell histiocytosis

NORMAL DEVELOPMENT AND STRUCTURE

The bones are derived from the mesenchyme of the mesoderm. They can be classified into three main types: the flat bones, including the skull, scapula, clavicle, pelvis and sternum; the tubular bones, including most of the bones of the extremities and the ribs; and epiphysioid bones, including the carpal (wrist) and tarsal (ankle) bones and the patella. The tubular bones grow until adulthood by virtue of a cartilaginous growth plate located near the ends

of the bones. The region between the growth plate and the end of the bone is called the epiphysis, the region around the growth plate is called the metaphysis and the region between the metaphyses is known as the diaphysis. Different bone tumours often have distinct predilections for specific regions of the bone.

Mature bone is a combination of two main types of bone, including cortical bone, dense bone present at the exterior, and trabecular bone, a looser meshwork of bone enclosing the medullary cavity containing the bone marrow. The matrix of bone is composed of collagen, proteoglycans and calcium-containing hydroxyapatite. The cellular composition of bone includes osteoblasts, the specialized cells responsible for the synthesis of the bony matrix; osteocytes, osteoblasts that have been incorporated into the bony matrix and responsible for the maintenance and metabolic activity of bone; and osteoclasts, multinucleated cells of monocyte/macrophage lineage responsible for the resorption and remodelling of bone. Cartilage is present at the joint surfaces and the immature growth plate. It consists of S-100 protein-expressing, specialized cells called chondrocytes embedded in an avascular matrix rich in proteoglycans. The external lining of the bone is a fibrous membrane called the periosteum.

The tubular bones are usually formed by a process called endochondral ossification, in which a cartilaginous intermediary forms a scaffold for the deposition of bone. Growth of the bone occurs at the cartilaginous growth plate by calcification of the cartilaginous matrix accompanied by apoptosis of chondrocytes and osteoblastic differentiation, with conversion of the cartilage to immature bone. Flat bones are usually formed by a process called intramembranous ossification, in which fetal mesenchymal cells directly differentiate into osteoblasts without a pre-existing cartilage matrix. Immature bone that lacks a calcified matrix is known as osteoid. The first bone formed lacks normal organisation and is called woven bone. In time, woven bone is organized into concentric layers surrounding blood vessels (haversion canals) forming the osteons of mature lamellar bone.

EPIDEMIOLOGY AND AETIOLOGY

Bone tumours are relatively rare tumours, accounting for about 0.2% of all malignancies, or about 1 in 100 000 individuals (Dorfman and Czerniak, 1995). There has been no discernible trend toward a change in the frequency in recent years. In general, there is a bimodal age distribution, with one peak occurring in adolescence (at the time of greatest bone growth) and a second peak in patients older than 60 years. There is a slightly higher incidence of bone tumours in whites than blacks, and in males than females. However, each specific bone tumour has its own characteristic age, race and sex predilections. Since bone tumours are so rare, there are no effective programmes for screening and prevention.

Most bone tumours have no known aetiology. A small percentage of bone tumours are due to genetic predisposition. There are several well-known syndromes in which specific bone tumours are markedly increased in frequency. Multiple hereditary exostoses is a rare autosomal dominant disorder in which patients develop multiple osteochondromas associated with bone deformities of the affected sites. Some families with this syndrome have a genetic defect at a gene called *EXT-1* present on chromosome 8q24.1 (Ahn *et al*., 1995). Other families have a different genetic defect at a gene called *EXT-2* on chromosome 11p11–13, while still other families may have an abnormality on chromosome 19 (Strickens *et al*., 1996). About 20% of patients develop secondary malignancies, usually a chondrosarcoma arising in a previous osteochondroma. Ollier disease is a rare, nonhereditary congenital disorder characterized by multiple enchondromas. Malignant transformation, again usually with chondrosarcoma, occurs in about 20–25% of patients with Ollier disease (Goodman *et al*., 1984). Mafucci syndrome is similar to Ollier disease in the development of multiple enchondromas, but patients also have co-existing soft tissue angiomas (Bean, 1958). The incidence of complicating chondrosarcoma is even higher in these patients, at approximately 50%; in addition, a variety of other extraskeletal neoplasms have also been reported. Chondrosarcoma may also arise as a rare complication of solitary enchondromas and osteochondromas.

Patients with Li–Fraumeni syndrome carry a germline mutation in the *TP53* gene, leading to an increase in the frequency of numerous neoplasms, including osteosarcoma. In addition, families carrying mutations in the retinoblastoma (*RB*) gene also have a greatly increased risk of osteosarcoma. In addition, there may be a familial osteosarcoma syndrome, independent of the above two disorders. Finally, Rothmun–Thomson syndrome is an extremely rare syndrome characterized by skin, endocrine and neural abnormalities, in which there is an increased incidence of squamous cell carcinoma of the skin and osteosarcoma, which may be multicentric (Vennos *et al*., 1992).

Nongenetic aetiological factors include pre-existing Paget disease of bone, which predisposes to a variety of high-grade sarcomas; osteomyelitis with sinus tract formation, which predisposes to squamous cell carcinoma; bone infarcts, which predispose to bone sarcoma, usually malignant fibrous histiocytoma; and radiation injury, which predisposes to a variety of bone sarcomas. Finally, some investigators have suggested that Kaposi sarcoma-associated human herpesvirus (HHV)-8 may be present in bone marrow antigen presenting cells in cases of multiple myeloma (Berenson and Vescio, 1999). HHV-8 produces viral interleukin-8, which may act as a stimulator of plasma cells, the precursor cells of multiple myeloma.

However, the issue is highly controversial at present, as the findings are inconsistent within the numerous laboratories that have attempted to confirm the original findings.

PATHOLOGY OF BENIGN TUMOURS

Osteoid osteoma and osteoblastoma are the two main benign bone-forming tumours (Healy and Ghelman, 1986). Osteoid osteoma usually occurs in patients between the ages of 5 and 25 years, with a male predilection. It is usually less than 1 cm in size and typically occurs in the tubular bones in the extremities. Histologically, bone formation by benign osteoblasts is seen (**Figure 1; see colour plate section**). Osteoblastoma usually occurs in patients between the ages of 10 and 40 years, with a male predilection. It is usually larger than 1.5 cm in size and typically occurs in the vertebral column. The histological features are very similar to those seen in osteoid osteoma. Osteochondroma is more a hamartomatous (developmental) anomaly than a true neoplasm (Milgram, 1983). It is a common lesion that has a peak incidence between 10 and 30 years of age, with a male predilection. It typically occurs on the surface of the metaphyseal region of the large tubular bones of the extremities. Pathologically, there is an external benign cartilage cap, usually less than 1 cm in thickness, with an underlying support of benign trabecular bone (**Figure 2; see colour plate section**). The histological features resemble the normal sequence of endochondral ossification.

Enchondroma, chondroblastoma and chondromyxoid fibroma are the three most common benign tumours of cartilage. Enchondroma is common neoplasm that occurs in patients of a wide age range, with no clear sex predilection (Mirra *et al*., 1985). It usually occurs in the medullary cavity of tubular bones, particularly in the hands and feet. Clinically, it is usually unassociated with pain (in contrast to some other tumours of cartilage including chondrosarcoma), and histologically, it is composed of mature cartilage with a lobular architecture and containing chondrocytes with small nuclei that lack atypia (**Figure 3; see colour plate section**). Chondroblastoma is an uncommon neoplasm that typically occurs in patients between the age of 10 and 20 years, with a male predilection (Dahlin and Ivins, 1972). It usually occurs in the epiphysis of the large tubular bones of the extremities. It is often associated with pain clinically, and histologically consists of varying numbers of chondroblasts, chondrocytes and multinucleated giant cells in a matrix that varies from immature to mature cartilage, often with a distinctive 'chicken wire' pattern of calcification (**Figure 4; see colour plate section**). Chondromyxoid fibroma is a relatively rare neoplasm that typically occurs in patients between the age of 10 and 30 years, with a male predilection (Rahimi *et al*., 1972). It usually occurs in the metaphysis of the large tubular bones of the extremities. It is usually associated with pain, and histologically shows areas of immature myxoid mesenchymal tissue with variable areas showing primitive cartilage formation.

Fibrous dysplasia is a dysplastic disorder of bone characterized by the inability to form mature lamellar bone (Reed, 1963). It occurs in two major forms. The monostostic form occurs as a single focus in one bone, while the polyostostic form occurs as multiple foci in several bones, either unilateral in one area (monomelic) or widespread (polymelic). The Albright–McCune syndrome is the combination of polyostotic fibrous dysplaia, with endocrine abnormalies (typically precocious puberty in females). Mutations of signal-transducing G proteins may underly this syndrome (Shenker *et al*., 1994). Histologically, fibrous dysplasia displays abnormal immature woven bone with an intervening stroma of spindle cells. Osteofibrous dysplasia is a rare lesion that occurs in infants and children in the bones of the lower leg that histologically has a close resemblance to fibrous dysplasia.

Nonossifying fibroma (fibrous cortical defect) is a very common lesion that typically occurs in patients aged between 5 and 15 years, without a distinct sex predilection. It is often an incidental finding discovered during a radiological study, and usually occurs in the metaphyses of the long tubular bones. Histologically, a bland but cellular spindle cell proliferation is seen (**Figure 5; see colour plate section**). Left alone, most lesions regress spontaneously, evidence that nonossifying fibroma probably represents a developmental disorder due to incomplete ossification. Benign fibrous histiocytoma is a rare neoplasm that occurs in patients of all ages, with no sex predilection (Roessner *et al*., 1981). It most commonly involves the pelvis and the ribs. The histological features are identical with those seen in nonossifying fibroma, with the distinction made on its clinical and radiological features; however, benign fibrous histiocytoma probably represents a true neoplasm.

Haemangioma is a benign tumour of blood vessels (Wold *et al*., 1982). It most often occurs in adults, and is most commonly seen in the skull or the spine. Grossly, they are red–brown, well-demarcated, medullary lesions. Histologically, they are composed of capillary-sized (capillary haemangioma) or slightly larger (cavernous haemangioma) vessels lined by bland endothelial cells. Mixtures of the two patterns are most frequently seen. Lymphangioma is a rare benign tumour of lymphatics. Histologically, it consists of a dilated, thin-walled lymphatic filled with a proteinaceous fluid. Both haemangiomas and lymphangiomas may be multiple. Regional angiomatosis involves one or several bones in a single anatomical region, whereas disseminated (cystic) angiomatosis affects multiple sites within bones of the trunk. Multiple lymphangiomas are often asscoiated with soft tissue lymphangioma or systemic lymphangiomatosis. A rare sporadic syndrome known as Gorham disease is an

aggressive form of angiomatosis in which massive osteolysis occurs which eventually effaces the normal bone architecture. The disease may stabilize or, in some cases, lead to death.

GIANT CELL TUMOUR OF BONES

In contrast to the preceding tumours, giant cell tumour of bone is a locally aggressive neoplasm with a high propensity for local invasion, frequent recurrence and rare pulmonary metastases (Dahlin, 1985). It usually occurs in patients between 20 and 40 years of age, without a sex predilection, and it may have an increased incidence in Chinese patients. It typically affects the epiphyses of the large tubular bones (**Figure 6; see colour plate section**). Histologically, one sees sheets of plump, spindle cells in which numerous multinucleated giant cells are evenly interspersed (**Figure 7; see colour plate section**). The giant cells are distinctive and may have up to 100 nuclei, which are said to resemble the nuclei of the mononuclear spindle cells. The lineage of the mononuclear and giant cells is not yet entirely clear. The giant cells show some features of the monocyte/macrophage lineage, similar to benign osteoclasts, while the mononuclear cells have suggestive but not definite features of monocyte/macrophage lineage. Furthermore, it is not clear which cell population (or whether both) represents the neoplastic element. A high proportion of giant cell tumours are aneuploid, and the majority of cases show chromosome aberrations, particularly telomeric fusion. Patients with giant cell tumour are usually treated by thorough curettage and bone grafting. Rare cases of giant cell tumour may be complicated by secondary malignancy; many of these patients have received prior radiotherapy. Rarer cases represent *de novo* malignant giant cell tumour, characterized by overtly atypical histological features.

OSTEOSARCOMA

Osteosarcoma represents the most common sarcoma of bone, accounting for about one-quarter of all primary malignancies of bone and about one-third of all bone sarcomas (Dorfman and Czerniak, 1998). It is defined as a malignant mesenchymal neoplasm that shows differentiation towards bone formation. Osteosarcoma may be divided into the common intramedullary variant and the much rarer surface oestosarcomas. In conventional intramedullary osteosarcoma, there is a bimodal age distribution with a large peak in patients between the age of 10 and 20 years (corresponding to the peak of bone growth) and a second smaller peak in older patients; there is a male predilection. The locations of the neoplasms in the peak in adolescence are those areas with the greatest growth rate; therefore, the metaphyses of the large tubular bones are the most commonly affected sites. On radiological studies, osteosarcomas may be lytic or sclerotic, depending on the amount of bone formation in the lesion. The borders of the lesion are usually ill-defined, and a soft tissue component is often seen. Grossly, the cut section of an osteosarcoma is variegated, with areas of bony, chondroid and soft tissue, usually with foci of haemorrhage and necrosis (**Figure 8; see colour plate section**). Histologically, a wide variety of patterns can be seen, united by the presence of malignant cells forming bone, even if very focal (**Figure 9; see colour plate section**) (Dahlin and Unni, 1977). Thus, although an osteoblastic pattern with obvious bone formation is most common, chondroblastic, fibroblastic, malignant fibrous histiocytoma-like, osteoblastoma-like, giant cell-rich, small-cell, epithelioid and telangiectatic variants occur, which have little to no impact on prognosis, but may cause great difficulty in differential diagnosis. Mixtures of several patterns are often seen in an individual case. One histological variant that may have an impact on prognosis is well-differentiated intramedullary osteosarcoma. Microscopically, this variant has deceptive bland cytological features combined with a relatively mature pattern of bone. Although most of the histological variants of osteosarcoma are considered to be of high grade, well-differentiated intramedullary osteosarcoma is the one exception that is of low grade when it is the only histological component present.

DNA ploidy studies have usually demonstrated highly aneuploid populations in osteosarcoma (Mandahl *et al.*, 1993). Classical cytogenetic studies confirm this, often showing grossly aneuploid karyotypes, with numerous extra chromosomes, marker chromosomes or loss of chromosomes, particularly loss of chromosomes 3, 10 and 12 (Mertens *et al.*, 1993). Even cases with a normal complement of chromosomes usually have evidence of structural rearrangements, particularly in chromosome 1q. On a molecular level, a high proportion of cases of osteosarcomas has mutations in the *RB* gene, similar to those seen in patients with familial retinoblastoma (and possibly explaining the high frequency of osteosarcoma in these patients) (Miller *et al.*, 1996). In addition, many cases also show mutations in the *TP53* gene (possibly explaining the increased frequency of osteosarcoma in patients with the Li–Fraumeni syndrome). The double minute 2 (*MDM2*) gene is amplified in a high proportion of metastatic lesions, providing another mechanism of *TP53* gene inactivation in osteosarcoma (Ladanyi *et al.*, 1993).

Untreated, conventional osteosarcoma is always fatal, and if treated with surgery alone, patients with osteosarcoma have survival rates no higher than 10–20%. However, patients with resectable lesions and without evidence of metastases are now often treated with

aggressive preoperative chemotherapy followed by complete surgical excision (often limb-sparing) followed by postoperative chemotherapy (Link *et al.*, 1986). These patients now have 5-year survivals greater than 50%. The degree of tumour necrosis observed in the resection specimen is of critical importance as patients with 90% or more necrosis in their tumours have survival rates greater than 80%, whereas patients with less than 90% necrosis have survival rates less than 20%. Other prognostic factors include the site of the neoplasm (patients with tumours in the long bones do better than those with tumours in the trunk), tumour size, the degree of aneuploidy, the status of the *TP53* gene and the status of the multidrug resistance gene (*MDRI*) (upregulation is associated with chemoresistance and decreased survival) (Baldini *et al.*, 1995). Metastases are most commonly seen in the lung and liver. Patients with solitary or even several metastases may benefit from surgical resection and adjuvant chemotherapy.

There are three main types of surface osteosarcoma: parosteal osteosarcoma, periosteal osteosarcoma and high-grade surface osteosarcoma. High-grade surface osteosarcoma is a rare type of high-grade osteosarcoma that develops on the surface of a long bone without the usual medullary involvement. Other than the lack of medullary involvement, it has an epidemiology, pathology and natural history similar to those of conventional osteosarcoma. However, parosteal osteosarcoma and periosteal osteosarcoma are distinctive low-grade forms of osteosarcoma. Parosteal osteosarcoma is a rare variant characterized by location on the surface of long tubular bones on top of the periosteum and a distinctive microscopic appearance (Unni *et al.*, 1976a). The peak incidence is between 20 and 30 years, with a female predilection. These neoplasms are characteristically located on the posterior aspect of the distal portion of the femur, just above the knee joint, or just opposite the knee joint in the posterior aspect of the proximal portion of the tibia. Grossly and radiologically, there is an exophytic (mushroom-like) growth on the periosteum, without elevation of the periosteum. Microscopically, there is a bland spindle-cell proliferation admixed with well-formed bony trabeculae. These tumours when pure have an excellent prognosis following radical surgical excision alone; however, the presence of focal areas of high-grade osteosarcoma seen microscopically is associated with a much higher incidence of metastases, and patients with such neoplasms are often treated with postoperative chemotherapy. Periosteal osteosarcoma is a rare low- to intermediate-grade variant of osteosarcoma that develops on the surface of long tubular bones, but in contrast to parosteal osteosarcoma, it occurs beneath the periosteum (Unni *et al.*, 1976b). It is also rare, with a peak occurrence between 10 and 30 years, with a female predilection. These tumours characteristically occur in proximal tibia or the distal femur. Grossly and radiologically, the lesion is present on the surface of the cortex, with elevation of the periosteum above the tumour. This often induces periosteal reactive new bone formation as perpendicular striae within the tumour and parallel striae at the edges where it attaches to the cortex, the latter often forming a characteristic ‘Codman triangle’ that may be observed on radiographs. Histologically, one sees prominent cartilaginous differentiation, with only focal areas of bone formation, identifying the neoplasm as an osteosarcoma. The tumour nuclei are usually of intermediate to high grade. Patients with periosteal osteosarcoma are usually treated by radical surgical excision, with a survival of about 70%.

CHONDROSARCOMA

Chondrosarcoma of bone is the second most common sarcoma of bone, accounting for about one-fifth of all primary bone malignancies and about one-quarter of all bone sarcomas. It is defined as a malignant tumour forming cartilaginous matrix, without any evidence of bone formation directly synthesized by the neoplastic cells (Sanerkin, 1980).

Chondrosarcoma may be divided into the common conventional chondrosarcoma and the rarer dedifferentiated, clear-cell and mesenchymal variants. Conventional chondrosarcoma is a neoplasm of older adults, with an increasing incidence with increasing age; there is no sex predilection. It occurs most frequently in the pelvis, ribs and proximal extremities, and is very rare in the spine and craniofacial bones. Patients usually present with a dull aching pain (in contrast to most benign lesions of cartilage), often with local swelling due to a mass effect. Radiographically, it usually appears as a clearly demarcated radiolucent lesion with discrete calcified opacities originating in the medullary cavity. Rarely, the neoplasm may arise in a subperiosteal site on the bone surface, a variant termed juxtacortical chondrosarcoma. Grossly, chondrosarcoma usually consists of lobulated hyaline nodules, with areas of calcification, particularly at the periphery of the lobules (**Figure 10; see colour plate section**). An extraosseous component is often present, and high-grade lesions may show areas of haemorrhage and necrosis. Histologically, the hallmark of a chondrosarcoma is the presence of a hyaline cartilage matrix, usually in a lobulated architecture. There is often a variable degree of calcification in the cartilage matrix, and variable degrees of myxoid change (a loose, gelatinous matrix) may be present. Bone may be present in the lesion, but this bone arises from the cartilage via endochondral ossification, and not directly from the neoplastic cells; the presence of the latter would mandate a diagnosis of osteosarcoma.

The neoplastic cells in chondrosarcoma are chondrocytes of varying degree of atypia, graded I–III (Rosenthal *et al.*, 1984). Grade I chondrosarcomas have

cells whose nuclei differ only subtly or not at all from the chondrocytes in benign enchondromas, with the distinction between the two neoplasms best made based on the radiological appearance rather than the pathological findings (**Figure 11; see colour plate section**). Grade II chondrosarcoma features chondrocytes with a greater degree of cytological atypia. Grade III chondrosarcomas are rare, but feature chondrocytes with overtly malignant nuclear features. Both grade I and grade II chondrosarcomas may progress to higher grade chondrosarcomas when recurrence (or metastasis) occurs. Similar to normal chondrocytes, the neoplastic cells in chondrosarcoma express S-100 protein.

Essentially all grade I, and many grade II, chondrosarcomas are diploid, while many grade II and essentially all grade III chondrosarcomas have an aneuploid DNA population. Classical cytogenetic studies reveal complex karyotypes, particularly in high-grade neoplasms, with nonrandom abnormalities in chromosome 1p. High-grade chondrosarcomas have been reported to have mutations or abnormal overexpression of the *TP53* gene in many cases (Nawa *et al*., 1996).

The primary treatment for chondrosarcoma is complete surgical excision, with adjuvant chemotherapy and radiotherapy only effective in high-grade chondrosarcomas. Prognostic factors include histological grade, size, anatomical location, location (extremities better than axial skeleton), presence of aneuploidy, S-phase measurements and presence of TP53 mutation. In general, grade I chondrosarcoma has the potential to recur, but generally does not metastasize, while grade II chondrosarcoma has a higher potential to recur and metastasizes in 10–20% of cases. Overall, grade I and II chondrosarcomas have a 5-year survival greater than 80%, although the disease may recur years after treatment. Grade III chondrosarcoma is a high-grade tumour with a high propensity for recurrence and metastasis, and a 5-year survival of about 20%. Metastases most often occur in lungs and liver.

In about 10% of cases of chondrosarcoma, a phenomenon known as dedifferentiation may occur (McCarthy and Dorman, 1982). In a dedifferentiated chondrosarcoma, a grade I or grade II chondrosarcoma is associated with a distinct area of a high-grade sarcoma, with a clear demarcation between the two areas. Patients with dedifferented chondrosarcoma have similar epidemiological characteristics as those with conventional chondrosarcoma, but often give a clear history of an increase in pain, which may in about half of patients may represent a pathological fracture. The radiological and gross appearance usually shows an area characteristic of conventional chondrosarcoma with an adjacent lytic region corresponding to a fleshy mass. Microscopically, the high-grade sarcoma often has features of a malignant fibrous histiocytoma or osteosarcoma. DNA studies have shown that the low-grade component is usually diploid whereas the dedifferentiated component is aneuploid. Cytogenetic studies have shown abnormalities common to both components, with additional abnormalities in the high-grade component, consistent with a common origin. The dedifferentiated areas consistently show overexpression of the p53 protein, consistent with mutations in the *TP53* gene (Simms *et al*., 1995). In addition, loss of the *RB* gene and/or loss of Bcl-2 protein present in the low-grade component is seen in most cases. The prognosis of dedifferentiated chondrosarcoma is very poor, with few long-term survivors.

Clear-cell chondrosarcoma is a rare variant of chondrosarcoma with characteristic epidemiological and pathological features (Bjornsson *et al*., 1984). It has a peak of incidence between 20 and 30 years, with a strong male predilection. It occurs most often at the proximal ends of the long tubular bones, usually extending to the articular cartilage. Radiologically, a lytic defect is usually seen. Grossly, it is usually a well-circumscribed soft, grey mass, often with focal calcifications. Histologically, one observes distinctive variants of chondrocytes with abundant clear cytoplasm in a loose cartilaginous matrix, often with varying foci of calcification. Ultrastructural and histochemical studies demonstrate that the clear cytoplasm is due to the presence of abundant glycogen. Preliminary cytogenetic studies suggest that clear-cell chondrosarcoma may possess a hypodiploid karyotype distinct from other forms of chondrosarcoma. Clear-cell chondrosarcoma is a low-grade malignancy, but metastasis may occur.

Mesenchymal chondrosarcoma is a very rare variant of chondrosarcoma that has a peak of incidence between 20 and 30 years with no sex predilection (Nakashima *et al*., 1986). It usually affects the bones of the jaws, the vertebrae or the ribs. Radiologically, it most often is a lucent lesion with varying degrees of calcification, and grossly it is usually a well-circumscribed soft grey mass, with foci of calcification. Mesenchymal chondrosarcoma is defined by its characteristic histological appearance, with cellular areas of round to spindle cells and other areas showing cartilaginous differentiation of varying maturation, often with calcification (**Figure 12; see colour plate section**). The proportion of the two elements may vary widely from case to case. Only the cells in the cartiliginous component express S-100 protein. Mesenchymal chondrosarcoma is a high-grade neoplasm, with a propensity for recurrence and metastasis; the 5-year survival is under 50%.

EWING SARCOMA/PERIPHERAL NEUROECTODERMAL TUMOUR OF BONE (PNET)

Ewing sarcoma/PNET comprises about 15% of primary bone sarcomas. A similar neoplasm also occurs in soft

tissue sites. In bone, it essentially occurs in patients under the age of 30 years with a peak between 10 and 20 years (Kissane *et al*., 1983). There is a slight male predilection, and the neoplasm only rarely occurs in blacks. It occurs in all bones, but has a slight preference to involve the long tubular bones, pelvis and ribs; when it occurs in the ribs, it has been known as the Askin tumour. Radiologically, one sees an ill-defined, lytic lesion that involves the intramedullary spaces. When long bones are affected, the diaphysis is usually the site of origin. There is often an extensive soft tissue component. Grossly, the neoplasm is a grey–white tumour that fills the medullary cavity, permeates the cortex, grows subperiosteally and usually forms an extensive soft tissue mass (**Figure 13; see colour plate section**). Microscopically, Ewing sarcoma/PNET consists of a highly homogeneous population of small cells with a fine chromatin pattern, indiscernible nucleoli and a thin rim of cytoplasm; thus, it is the prototype of the 'small, round, blue cell tumour' (**Figure 14; see colour plate section**). The mitotic rate is generally high, and there is usually a high number of apoptotic cells. In a subset of cases, the cells focally form rosettes around stroma, a pattern termed the Homer–Wright rosette, which is suggestive evidence of neural differentiation. The ultrastructural correlate of the Homer–Wright rosette is the presence of cytoplasmic projections, with their organisation toward a central core. Other ultrastructural features that suggest neural differentiation in a subset of tumours is the presence of neurosecretory granules, neurofilaments or neurotubules. Ultrastructural and histochemical studies demonstrate that the cytoplasm contains abundant glycogen in most cases, whereas immunohistochemical studies demonstrate that the cell membrane has consistent (although not highly specific) strong staining for CD99, the *MIC2* gene product (Llombart-Bosch *et al*., 1996). The neoplasm shows a spectrum of expression of neuroendocrine markers, including the chromogranin family of proteins, CD57, protein gene product 9.5 (PGP9.5), neurofilaments, S-100 protein, neuron-specific enolase and synaptophysin. Some pathologists consider neoplasms that possess two or more markers of neuroendocrine differentiation, Homer–Wright rosettes, or ultrastructural evidence of neural differentiation to represent PNET, while neoplasms that possess none or one neuroendocrine marker and lack ultrastructural markers of neural differentiation to represent Ewing sarcoma, whereas other pathologists prefer to consider all cases to represent one neoplasm regardless of the degree of neural differentiation, using the term Ewing's sarcoma/PNET for all tumours.

A characteristic molecular feature of Ewing sarcoma/PNET is the presence a translocation involving the *EWS* gene on chromosome 11 in over 90% of cases (Lopez-Terrada, 1996). A t(11;22)(q24;q12) occurs in about 90% of cases, and involves the *FLI-1* gene on chromosome 11, while the t(7;22)(p22;q12) and t(21;22)(q22;q12) are both rare, and involve the *ETV-1* gene on chromosome 7 and the *ERG-1* gene on chromosome 21, respectively. FLI-1, ETV-1 and ERG-1 are all transcription factors with DNA binding domains. In all three translocations, there is creation of a new fusion protein that uses the promoter of EWS to upregulate the transcription factor. Interestingly, a translocation has not been found in about 10% of cases of Ewing sarcoma. Additionally, there is a distinct soft tissue tumour, the small-cell desmoplastic tumour, that utilizes *EWS* in a distinct t(11;12)(p13;q12) involving another gene on chromosome 11, *WT-1*, that also codes for a transcription factor.

Ewing sarcoma/PNET is a highly aggressive tumour and, if untreated, will rapidly lead to death. Current treatment protocols include a combination of surgery, pre- and postoperative multidrug chemotherapy and radiotherapy, leading to a dramatic increase in survival, with 5-year survival rates of about 70% for those with resectable disease. Surgical removal of one to a limited number of lung metastases may improve overall survival. Prognostic factors include stage (particularly, the presence of metastases at diagnosis), tumour size, site (trunk lesions worse than extremity lesions) and the degree of necrosis in the resection specimen following preoperative chemotherapy (less than 90% necrosis worse than 90–100% necrosis). The presence or absence of neural differentiation is a controversial prognostic factor; some studies show the presence of neural differentiation to be associated with more aggressive tumours, when strict criteria are used.

CHORDOMA

Chordoma is a relatively uncommon neoplasm, representing about 5% of primary bone tumours, and about 8% of bone sarcomas. It is thought to originate from the remnants of the primitive notochord that is present in embryonal development and which forms the forerunner of the vertebral column. It occurs in all age groups, with a peak from 50 to 60 years (Dahlin and MacCarty, 1952). There is perhaps a slight male predominance and it is very rare in blacks. It almost always involves the axial skeleton, particularly the base of the skull and the sacrococcygeal region. Patients with chordoma usually present with pain or symptoms related to compression of adjacent structures. Radiologically, chordoma usually appears as a lytic lesion with scattered calcifications. Grossly, the tumours are grey–tan gelatinous multilobulated masses. Microscopically, chordomas have a resemblance to the normal notochord, with cords and nests of vacuolated cells in a myxoid matrix (**Figure 15; see colour plate section**). Immunohistochemical studies reveal differentiation similar to the cells of the primitive notochord with expression of S-100 protein and the epithelial markers keratin and epithelial membrane antigen. Cytogenetic studies have revealed abnormalities, most frequently involving chromosome 21.

Complete surgical excision is the treatment of choice. However, as these tumours frequently arise in the base of the skull, complete excision is not always feasible; in these cases, radiotherapy is often used after tumour debulking. Chordoma is a moderately aggressive tumour, marked by both recurrences and metastases, most commonly in the lung.

Dedifferentiated chordoma is a tumour consisting of conventional chordoma and a distinct adjacent component of a high-grade sarcoma, similar to that seen in dedifferentiated chondrosarcoma (Meis *et al*., 1987). It may occur in primary or recurrent lesions of chordoma. Radiologically, the neoplasm appears as a destructive lytic lesion. Grossly, the dedifferentiated component is a fleshy mass adjacent to a gelatinous region typical of chordoma. Histologically and immunohistochemically, the sarcomatous component is undifferentiated. The prognosis of dedifferentiated chordoma is very poor, with widespread metastasis and few survivors.

Chondroid chordoma is a controversial entity that contains an admixture of chordoid and cartilaginous areas (Wojno *et al*., 1992). It occurs in bones commonly affected by pure chordomas. Immunohistochemical studies usually show chordoid differentiation (keratin, epithelial membrane, and S-100 protein positive) in both areas of the tumour. Patients with these tumours appear to have a longer survival time than those patients with conventional chordoma.

MALIGNANT FIBROUS HISTIOCYTOMA AND FIBROSARCOMA

Malignant fibrous histiocytoma is a rare sarcoma of bone, comprising approximately 3% of all primary bone tumours and 5% of bone sarcomas. It is similar to the more common entity seen in the soft tissues now thought to be of myofibroblastic differentiation rather than derived from the monocyte–macrophage lineage. It occurs in all age groups, with an increasing incidence with age; there is no sex predilection (Capanna *et al*., 1984). It may involve any bone, but there is a predilecton for the metaphyses of the long tubular bones. Radiologically, a lytic lesion is seen, usually with an associated soft tissue mass. The gross appearance is that of a sarcoma, with a fleshy grey mass, usually with haemorrhage and necrosis. Histologically, a spindle cell proliferation is seen, forming whorls and fascicles. Storiform–pleomorphic, giant cell and myxoid variants are recognized, but there is often intermixture of these components within a single case. Bone formation by the neoplastic cells is not present, as this would mandate the diagnosis of osteosarcoma. The individual cells vary from highly spindled to oval. Nuclear atypia is usually marked, but can show a spectrum from case to case and within a given case (**Figure 16; see colour plate section**). Immunohistochemical studies are commonly performed to rule out other entities, but there are no specific findings in malignant fibrous histiocytoma. Although expression of smooth muscle actin is often seen and used to support a myofibroblastic differentiation, many other sarcomas may also express this antigen. There have been no specific molecular abnormalities found in malignant fibrous histiocytoma; this is not surprising as the entity probably represents a 'waste-basket' of cases of bone sarcoma that do not show specific features of other sarcomas. It is usually treated by radical surgical excision followed by systemic chemotherapy and/or radiotherapy, although newer approaches are examining treatment protocols similar to those used for osteosarcoma. The prognosis is poor, particularly in those patients whose tumours have arisen as a transformation event in a lower grade sarcoma. Metastases most often occur in the lungs.

Some pathologists separate fibrosarcoma from malignant fibrous histiocytoma and regard it as a separate bone sarcoma, whereas other pathologists classify these cases within the category of malignant fibrous histiocytoma. The epidemiological, radiological and gross characteristics of fibrosarcoma are similar to those seen in malignant fibrous histiocytoma (Dahlin and Ivins, 1969). Histologically, those pathologists who distinguish fibrosarcoma from malignant fibrous histiocytoma look for a well-developed fascicular architecture, often with a 'herring-bone' pattern between the different fascicles. The fascicles are formed by spindled cells, without the oval cells typical of malignant fibrous histiocytoma. Collagen formation is variable. The immunohistochemical findings are similar to those seen in malignant fibrous histiocytoma, and the molecular findings are similarly nonspecific. The treatment is usually the same as that for malignant fibrous histiocytoma. The survival figures for patients with fibrosarcoma may be superior to those of patients with malignant fibrous histiocytoma, but this may be because a higher proportion of fibrosarcomas are of lower histological grade.

ADAMANTINOMA

Adamantinoma is a rare primary epithelial neoplasm of bone (Weiss and Dorfman, 1977). It is currently subdivided into classical and differentiated types. The classical type usually occurs in adults, without a sex predilection. It exclusively involves the tibia or, less often, the fibula, the long tubular bones of the lower leg. Patients present with pain in the involved site, often accompanied by a mass. Radiological studies show a mixed lytic and sclerotic mass that may either be intracortical or may show complete cortical disruption with involvement of the medullary cavity, soft tissue or both. Grossly, it appears as a fleshy mass. Histologically, adamantinoma features

epithelial cells arranged in various patterns, including basaloid, spindle, tubular, squamoid and osteofibrous dysplasia-like (**Figure 17; see colour plate section**). The tumour cells in all patterns are keratin-positive on immunohistochemical studies, and ultrastructurally show prominent desmosomes and evidence of keratinization, consistent with epithelial cells. Cytogenetic studies have revealed complex chromosomal abnormalities. Classical adamantinomas are indolent tumours with a high recurrence rate; metastasis, mostly to lungs, occurs in about one-quarter of patients.

Differentiated adamantinoma is extremely rare and, in contrast to classical adamantinoma, usually in patients under the age of 20 (Baker and Coley, 1953). It also exclusively involves the tibia and the fibula, and synchronous lesions may occur. The radiological appearance is identical with that of the classical subtype. Grossly, the tumour appears fibrous rather than fleshy. Histologically, the tumour looks very similar to osteofibrous dysplasia (see above), with the exception that single epithelial cells and small nests of epithelial cells are present within the fibrous stroma. These cells are best identified by immunohistochemical studies for keratin. Cases of differentiated adamantinoma have not yet been reported to metastasize.

ANGIOSARCOMA

Angiosarcoma is a malignancy of endothelial cells. It is a rare primary sarcoma of bone, representing less than 1% of cases. It is separated into classical and well-differentiated epithelioid subtypes. Both classical and well-differentiated occur in all age groups, with a peak of incidence in young adulthood; there is a male predilection (Volpe and Mazabraud, 1982; Wold *et al*., 1982). The long tubular bones of the lower extremity are most often involved, and multifocal lesions are frequent. Radiological studies show usually show a lytic lesion or multiple lesions. They are bright red on cut section. Histologically, classical angiosarcoma typically consists of irregular anastomosing channels lined by individual or heaped up masses of highly atypical endothelial cells. Solid sheets of these cells may also be present focally, extensively or comprising the entire lesion. In contrast, well-differentiated epithelioid angiosarcoma features plump (epithelioid) cells that form sheets, cords, and occasionally lines spaces. They are frequently vacuolated – the vacuoles representing abortive attempts at vessel formation. The nuclei usually show mild to moderate degrees of atypia. The proliferating cells of both subtypes of angiosarcoma express markers of vascular cells, including CD31, CD34 and factor VIII-related antigen. Expression of keratin and, occasionally, epithelial membrane antigen is seen in a subset of cases. Ultrastructural studies demonstrate Weibel–Palade bodies, cytoplasmic structures found in normal endothelial cells.

Classical angiosarcoma is a high-grade sarcoma, usually requiring radical surgery. The prognosis is poor, with a high propensity for metastasis, particularly to the lungs. In contrast, epithelioid angiosarcomas are indolent neoplasms, with local invasion and only occasional metastasis, (also to the lungs). They are usually treated by more conservative surgery with postoperative radiotherapy.

MULTIPLE MYELOMA

Multiple myeloma (plasma cell myeloma) represents a malignant neoplasm of plasma cells, the terminally differentiated cells of the B-lymphocyte lineage that secrete immunoglobulins (Bataille and Harousseau, 1997). Multiple myeloma is a relatively common tumour, occurring more frequently than all primary bone sarcomas combined. It occurs in adults, with an increasing frequency with age. There is a slight male predilection, and there is a higher incidence in blacks than whites. The spine, skull, pelvis and the ribs are most frequently involved. Patients typically present with bone pain and often have recurrent infections and symptoms of renal disease, either at presentation or during the course of the disease. Laboratory abnormalities include anaemia, hypercalcaemia, a monoclonal gammopathy in serum and/or urine and reduced levels of normal polyclonal immunoglobulins. The monoclonal gammopathy derives from the neoplastic plasma cells, and usually consists of IgG (about half of cases) or IgA (one-quarter of cases) in the serum and/or kappa or lambda light chains (called Bence Jones protein) in the urine. Radiological studies usually demonstrate multiple sharply delimited lytic lesions. Grossly, the lesions are usually soft and grey. Microscopically, a uniform proliferation of plasma cells is seen in biopsies from lesions, of varying levels of cytologic atypia (**Figure 18; see colour plate section**) (Bartl *et al*., 1987). Plasma cells are distinctive cells with a round, eccentrically placed nucleus with a clumped chromatin pattern and abundant cytoplasm with a paranuclear pale zone. In addition, an atypical plasmacytosis may also be seen in random biopsies or aspirates from random sites of bone such as the iliac crest that do not appear to be radiologically involved. Immunohistochemical studies usually demonstrate lack of the leucocyte common antigen CD45 and the B-lineage markers CD20 and CD19, because these markers are usually absent on terminal differentiated B-lineage cells. However, there is positivity for the B cell marker CD79a, which is expressed throughout the complete gamut of B cell maturation. In addition, there is usually expression of the plasma cell marker CD38 and the adhesion markers CD138 (syndecan-1), CD56 and CD58. Most importantly, there is almost always monotypic expression of immunoglobulin light and heavy chains, usually of the IgG or

IgA class. Some myeloma cells may express the immature B cell antigen CD10 (common acute lymphoblastic leukaemia antigen (CALLA)), myelomonocytic antigens or T cell antigens. In general, multiple myeloma has a relatively low proliferative index, indicating that the primary abnormality in this neoplasm may be lack of cell death due to loss of cell death mechanisms rather than active proliferation due to loss of growth control mechanisms.

Molecular studies demonstrate monoclonal rearrangements of the heavy and light chain immunoglobulin genes. Classical cytogenetic studies reveal consistent clonal structural abnormalities in about 40% of cases. The most common abnormalities involve chromosomes 11, 13 and 14, with a t(11;14) involving the *BCL1/CCND1* gene on chromosome 11 and the immunoglobulin heavy chain gene on chromosome 14 have been seen in about 25% of cases. High levels of c-Myc have been reported in a subset of patients. Mutations of the *TP53* gene have also been reported in some patients. Alterations of the *PAX-5* gene have also been reported, and may account for the loss of B-lineage markers in multiple myeloma.

Malignant myeloma is usually treated by chemotherapy, although radiotherapy and even surgery may play a role in the treatment of individual lesions that are particularly symptomatic. Prognostic factors include tumour burden (tumour mass), cytological features, expression of CD10, high proliferative index, specific cytogenetic abnormalities and the presence of complications such as renal dysfunction or amyloidosis.

There are several uncommon variants of multiple myeloma. Plasma cell leukaemia is a rare variant with a dominant leukaemic component (Dahlin and Unni, 1986). Solitary myeloma (plasmacytoma of bone) is defined as a neoplastic plasma cell proliferation present in one site of bone (Woodruff *et al*., 1979). Random bone marrow examinations lack plasmacytosis and a serum or urine gammopathy is usually absent. In the majority of cases, progression to multiple myelomas occurs within a few years. Smouldering myeloma lacks overt bone lesions but contains a moderate marrow plasmacytosis, while indolent myeloma has a lesser degree of marrow plasmacytosis but may have up to three bone lesions (Kyle and Griepp, 1980). Finally, monoclonal gammopathy of unknown significance is defined as the presence of a monoclonal gammopathy in the absence of bone lesions or significant marrow plasmacytosis (Kyle, 1978).

MALIGNANT LYMPHOMA PRESENTING IN BONE

Malignant lymphoma involves bones most commonly secondary to spread from other sites, but rarely may present as a primary lesion in bone (Baar *et al*., 1994). Primary bone lymphoma usually represents non-Hodgkin lymphoma. Most cases occur in adults with a slight male predilection. The large tubular bones, pelvis and the spine are most frequently involved, and multifocality may be seen (**Figure 19; see colour plate section**). About two-thirds of cases represent diffuse large B cell lymphoma, with peripheral T cell lymphoma, including cases of anaplastic large cell lymphoma being relatively rare. Hodgkin disease rarely presents in bone without extraskeletal involvement; it most commonly involves the lower spine or the pelvis (Gold and Mirra, 1979). The histological features of non-Hodgkin lymphomas and Hodgkin disease are identical with those seen in the more commonly involved sites.

LANGERHANS CELL HISTIOCYTOSIS

Langerhans cell histiocytosis (histiocytosis X) is a rare neoplastic proliferation of Langerhans cells (Lieberman *et al*., 1996). Most cases occur in childhood, with a predilection for males. The disease is rare in blacks. There are three overlapping syndromes, all of which may involve bone: unifocal disease (solitary eosinophilic granuloma), multifocal, unisystem disease (Hand–Schuller–Christian disease) and multifocal, multisystem disease (Letterer–Siwe disease). In unifocal disease, a single bone is affected, most commonly the skull, femur, pelvic bones or ribs). In multifocal, unisystem disease, there is involvement of several bones, usually the craniofacial bones. In multifocal, multisystem disease, multiple organ systems are involved, but the bones are commonly included, usually in a diffuse fashion. In general, the younger the patient at diagnosis, the more extensive is the pattern of involvement. Histologically, one sees a proliferation of Langerhans cells, in a milieu which typically includes eosinophils, histiocytes, neutrophils and lymphocytes. Langerhans cells have a characteristic grooved nucleus and abundant histiocyte-like cytoplasm (**Figure 20; see colour plate section**). Langerhans cells express CD1 and S-100 protein, and ultrastructurally show characteristic organelles called Birbeck granules. There is a wide spectrum of clinical behaviour, with the number of organs affected at presentation being the critical factor. Survival is greater than 95% in patients with unifocal disease, 80% in patients with multifocal unisystem disease and poor in patients with multisystem disease. The last type of patients are typically treated with multidrug chemotherapy.

REFERENCES

Ahn, J., *et al*. (1995). Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1). *Nature Genetics*, **11**, 137–143.

Baar, J., *et al*. (1994). Primary non-Hodgkin's lymphoma of bone: a clinicopathologic study. *Cancer*, **73**, 1194–1199.

Baker, H. and Coley, B. (1953). Chordoma of lumbar vertebra. *Journal of Bone and Joint Surgery*, **35A**, 403–408.

Baldini, N. *et al*. (1995). Expression of P-glycoprotein in high-grade osteosarcoma in relation to clinical outcome. *New England Journal of Medicine*, **333**, 1380–1385.

Bartl, R., *et al*. (1987). Histologic classification and staging of multiple myeloma: a retrospective and prospective study of 674 cases. *American Journal of Clinical Pathology*, **87**, 342–355.

Bataille, R. and Harousseau, J.-L. (1997). Multiple myeloma. *New England Journal of Medicine*, **336**, 1657–1664.

Bean, W. (1958). Dyschondroplasia and hemangiomata (Maffucci's syndrome). *Archives of Internal Medicine*, **102**, 544–550.

Berenson, J. R. and Vescio, R. A. (1999). HHV-8 is present in multiple myeloma patients. *Blood*, **15**, 3157–3159.

Bjornsson, J., *et al*. (1984). Clear cell chondrosarcoma of bone: observations in 47 cases. *American Journal of Surgical Pathology*, **8**, 223–230.

Capanna, R., *et al*. (1984). Malignant fibrous histiocytoma of bone: the experience at the Rizzoli Institute: report of 90 cases. *Cancer*, **54**, 177–187.

Dahlin, D. (1985). Caldwell Lecture: giant cell tumor of bone – highlights of 407 cases. *American Journal of Roentgenology*, **144**, 955–960.

Dahlin, D. and MacCarty, C. (1952). Chordoma: a study of fifty-nine cases. *Cancer*, **5**, 1170–1178.

Dahlin, D. and Ivins, J. (1969). Fibrosarcoma of bone: a study of 114 cases. *Cancer*, **23**, 35–41.

Dahlin, D. and Ivins, J. (1972). Benign chondroblastoma: a study of 125 cases. *Cancer*, **30**, 401–413.

Dahlin, D. and Unni, K. (1977). Osteosarcoma of bone and its important recognizable varieties. *American Journal of Surgical Pathology*, **1**, 61–72.

Dahlin, D. and Unni, K. (1986). *Bone Tumors: General Aspects and Data on 8,542 Cases*, 4th edn (Charles C. Thomas, Springfield, II).

Dorfman, H. D. and Czerniak, B. (1995). Bone cancers. *Cancer*, **75**, 203–210.

Dorfman, H. D. and Czerniak, B. (1998). *Bone Tumors* (Mosby, St. Louis).

Fechner, R. and Mills, S. (1993). *Atlas of Tumor Pathology: Tumors of the Bones and Joints*, 3rd Series (Armed Forces Institute of Pathology, Washington, DC).

Gold, R. and Mirra, J. (1979). Case report 101: primary Hodgkin disease of humerus. *Skeletal Radiology*, **4**, 233–235.

Goodman, S., *et al*. (1984). Ollier's disease with multiple sarcomatous transformations. *Human Pathology*, **15**, 91–93.

Healy, J. and Ghelman, B. (1986). Osteoid osteoma and osteoblastoma: current concepts and recent advances. *Clinical Orthopaedics*, **204**, 76–85.

Kissane, J., *et al*. (1983). Ewing's sarcoma of bone: clinicopathologic aspects of 303 cases from the Intergroup Ewing's Sarcoma Study. *Human Pathology*, **14**, 773–779.

Kyle, R. (1978). Monoclonal gammopathy of undetermined significance: natural history in 241 cases. *American Journal of Medicine*, **64**, 814–826.

Kyle, R. and Griepp, P. (1980). Smoldering multiple myeloma. *New England Journal of Medicine*, **302**, 1347–1349.

Ladanyi, M., *et al*. (1993). MDM2 gene amplification in metastatic osteosarcoma. *Cancer Research*, **53**, 16–18.

Lieberman, P., *et al*. (1996). Langerhans cell (eosinophilic) granulomatosis: a clinicopathologic study encompassing 50 years. *American Journal of Surgical Pathology*, **20**, 519–552.

Link, M., *et al*. (1986). The effect of adjuvant chemotherapy on relapse-free survival in patients with osteosarcoma of the extremity. *New England Journal of Medicine*, **314**, 1600–1606.

Llombart-Bosch, A., *et al*. (1996). Histology, immunohistochemistry, and electron microscopy of small round cell tumors of bone. *Seminars in Diagnostic Pathology*, **13**, 153–170.

Lopez-Terrada, D. (1996). Molecular genetics of small round cell tumors. *Seminars in Diagnostic Pathology*, **13**, 242–249.

Mandahl, N., *et al*. (1993). Comparative cytogenetic and DNA flow cytometric analysis of 150 bone and soft-tissue tumors. *International Journal of Cancer*, **53**, 358–364.

McCarthy, E. and Dorman, H. (1982). Chondrosarcoma of bone with dedifferentiation: a study of eighteen cases. *Human Pathology*, **13**, 36–40.

Meis, J., *et al*. (1987). "De-differentiated" chordoma: a clinicopathologic and immunohistochemical study of three cases. *American Journal of Surgical Pathology*, **11**, 516–525.

Mertens, F., *et al*. (1993). Cytogenetic findings in 33 osteosarcomas. *International Journal of Cancer*, **55**, 44–50.

Milgram, J. (1983). The origins of osteochondromas and enchondromas. *Clinical Orthopaedics*, **174**, 264–284.

Miller, C., *et al*. (1996). Alterations of the p53, Rb and MDM2 genes in osteosarcoma. *Journal of Cancer Research and Clinical Oncology*, **122**, 559–565.

Mirra, J., *et al*. (1985). A new histologic approach to the differentiation of enchondroma and chondrosarcoma of the bones; a clinicopathologic analysis of 51 cases. *Clinical Orthopaedics*, **201**, 214–237.

Nakashima, Y., *et al*. (1986). Mesenchymal chondrosarcoma of bone and soft tissue: a review of 111 cases. *Cancer*, **57**, 2444–2453.

Nawa, G., *et al*. (1996). Prognostic significance of Ki67 (MIB1) proliferation index and p53 over-expression in chondrosarcoma. *International Journal of Cancer*, **69**, 86–91.

Rahimi, A., *et al*. (1972). Chondromyxoid fibroma: a clinicopathologic study of 76 cases. *Cancer*, **30**, 726–736.

Reed, R. (1963). Fibrous dysplasia of bone: a review of 25 cases. *Archives of Pathology*, **75**, 480–495.

Roessner, A., *et al*. (1981). Benign fibrous histiocytoma of bone: light- and electron-microscopic observations. *Journal of Cancer Research and Clinical Oncology*, **101**, 191–202.

Rosenthal, D., *et al*. (1984). Chondrosarcoma: correlation of radiological and histological grade. *Radiology*, **150**, 21–26.

Sanerkin, N. (1980). The diagnosis and grading of chondrosarcoma of bone: a combined cytologic and histologic approach. *Cancer*, **45**, 582–594.

Shenker, A., *et al*. (1994). An activating Gs alpha mutation is present in fibrous dysplasia of bone in Albright–McCune–Sternberg syndrome. *Journal of Clinical Endocrinology and Metabolism*, **79**, 750–755.

Simms, W., *et al*. (1995). p53 expression in dedifferentiated chondrosarcoma. *Cancer*, **76**, 223–227.

Strickens, D., *et al*. (1996). The EXT2 multiple exostoses gene defines a family of putative tumour suppressor genes. *Nature Genetics*, **14**, 25–32.

Unni, K. (1996). *Dahlin's Bone Tumors: General Aspects and Data on 11,087 Cases*, 5th edn (Lippincott-Raven, Philadelphia).

Unni, K., *et al*. (1976a). Parosteal osteogenic sarcoma. *Cancer*, **37**, 2644–2675.

Unni, K., *et al*. (1976b). Periosteal osteogenic sarcoma. *Cancer*, **37**, 2476–2485.

Vennos, E., *et al*. (1992). Rothmund–Thomson syndrome: review of the world literature. *Journal of the American Academy of Dermatology*, **27**, 750–762.

Volpe, R. and Mazabraud, A. (1982). Hemangioendothelioma (angiosarcoma) of bone: a distinct pathologic entity with an unpredictable course? *Cancer*, **49**, 727–736.

Weiss, S. and Dorfman, H. (1977). Adamantinoma of long bone: an analysis of nine new cases with emphasis on metastisizing lesions and fibrous dysplasia-like changes. *Human Pathology*, **8**, 141–153.

Wojno, K., *et al*. (1992). Chondroid chordomas and low-grade chondrosarcomas of the craniospinal axis: an immunohistochemical analysis of 17 cases. *American Journal of Surgical Pathology*, **16**, 1144–1152.

Wold, L., *et al*. (1982). Hemangioendothelial sarcoma of bone. *American Journal of Surgical Pathology*, **6**, 59–70.

Woodruff, R., *et al*. (1979). Solitary plasmacytoma. II. Solitary plasmacytoma of bone. *Cancer*, **43**, 2344–2347.

FURTHER READING

Dahlin, D. and Unni, K. (1986). *Bone Tumors: General Aspects and Data on 8,542 Cases*, 4th edn (Charles C. Thomas, Springfield, IL).

Dorfman, H. D. and Czerniak, B. (1998). *Bone Tumors* (Mosby, St. Louis, MO).

Fechner, R. and Mills, S. (1993). *Atlas of Tumor Pathology: Tumors of the Bones and Joints*, 3rd Series (Armed Forces Institute of Pathology, Washington, DC).

Huvos, A. G. (1991). *Bone Tumors: Diagnosis, Treatment, Prognosis* (W. B. Saunders, Philadelphia).

Jaffe, H. L. (1972). *Tumors and Tumorous Conditions of the Bones and Joints* (Lea and Febiger, Philadelphia).

Mirra, J. M. (1989). *Bone Tumors: Clinical, Radiologic and Pathologic Correlations* (J.B. Lippincott, Philadelphia).

Unni, K. (1996). *Dahlin's Bone Tumors: General Aspects and Data on 11,087 Cases*, 5th edn (Lippincott-Raven, Philadelphia).

Chapter 48

Soft Tissues

Andre M. Oliveira
Mayo Graduate School of Medicine, Rochester, MN, USA

Antonio G. Nascimento
Mayo Clinic, Mayo Medical School, Rochester, MN, USA

CONTENTS

- Development and Structure of Soft Tissues and Soft Tissue Neoplasms
- Classification
- Epidemiology, Aetiology and Pathogenesis
- Genetics and Molecular Biology
- Grading and Staging Systems
- Diagnosis
- Clinicopathological Features
- Prognostic Factors
- Treatment

DEVELOPMENT AND STRUCTURE OF SOFT TISSUES AND SOFT TISSUE NEOPLASMS

Soft tissue is a broad and poorly defined term that includes distinct tissue types, including connective or supporting tissues, vascular tissue and peripheral nerve tissue. Connective tissues develop from the embryonic mesoderm and are composed of mesenchymal cells. Connective tissues include cartilage, fat, ligaments, tendons, capsules, fasciae, aponeuroses, skeletal muscle, smooth muscle and bone. Bone is excluded from the definition of soft tissue because of its unique histological characteristics. Vascular tissue, which is composed of endothelial cells, is also a mesoderm-derived tissue, originating from primitive mesenchymal cells (angioblasts) during the third week of the embryonic development. Peripheral nerve tissue, in contrast to the other soft tissues, is an ectoderm-derived tissue predominantly originating from the neural crest.

New concepts on soft tissue oncogenesis are being advanced; to avoid misconceptions, we shall discuss some ideas briefly. First, soft tissue tumours currently are considered to be a group of distinct neoplasms showing predominantly mesenchymal differentiation. The word histogenesis will be abandoned, at least for the time being. The erroneous idea that soft tissue tumours are originated from well-developed adult tissue has little clinicopathological support in most cases. Two examples that challenge the concept of histogenesis are liposarcoma and rhabdomyosarcoma. These tumours frequently arise in areas devoid of adipose tissue and skeletal muscle, respectively. Mesenchymal cells are mesoderm-derived motile and nonpolarized embryonic connective tissue cells that can undergo differentiation into distinct cell types. Accordingly, the current hypothesis on soft tissue oncogenesis is that genetic alterations in mesodermal progenitor cells lead to many types of soft tissue tumours. The line of differentiation seems to depend mainly on two factors: the level of commitment of these mesodermal progenitors to development into a certain type of cell when the oncogenic events occur and on the differential pattern of gene expression during the clonal evolution of the tumour. However, it is important to note that even the concepts of mesoderm progenitors or cell of origin will not be blindly adopted or denied. Research has challenged traditional ideas of cell differentiation, dedifferentiation and transdifferentiation, finding a deeper phenotypic cellular plasticity than previously thought (Orkin, 2000). Furthermore, nothing precludes the possibility that oncogenic processes also occur in well-developed cells. Some tumours of smooth muscle and neural differentiation are examples of this point.

Second, soft tissue neoplasms are not necessarily restricted to the same patterns of development and differentiation as are normal tissues. In neoplasms, uncontrolled and cumulative genetic and epigenetic alterations result in cell phenotypes that are distinct from those of normal cells. Searching for counterparts in normal cell types is helpful mainly for classification purposes; this activity should not obligate us to current paradigms of differentiation.

Third, soft tissue neoplasms need not be exclusively restricted to mesenchymal differentiation. The same tumorigenic alterations described above may result in other

cell phenotypes, such as epithelial. There are numerous examples of soft tissue neoplasms with a biphasic appearance, or aberrant differentiation; an example is the biphasic synovial sarcoma. Epithelial–mesenchymal and mesenchymal–epithelial transitions are well-known developmental processes that also can occur during tumorigenesis.

Fourth, the traditional and academic distinction between benign and malignant soft tissue tumours is blurred because many soft tissue tumours behave in a benign fashion despite malignant-like clinicopathological features. The reverse is also true. Examples that do not exactly fit the traditional view include deep fibromatoses (desmoid tumours) and metastasizing benign leiomyomata. Therefore, a spectrum of tumours ranging from the very benign to the fully malignant does exist.

Fifth, soft tissue tumour oncogenesis does not follow the current model of carcinogenesis. At our present level of knowledge, concepts of dysplasia or preneoplastic lesions do not readily apply in most cases of soft tissue sarcoma. Malignant transformation of benign soft tissue tumours into malignant tumours seems very rare; the best example is the malignant transformation of neurofibroma in malignant peripheral nerve sheath tumour in patients with neurofibromatosis.

Sixth, soft tissue sarcomas should not be regarded as rare tumours refractory to any kind of treatment. They comprise various neoplasms with distinct epidemiological, clinicopathological and genetic features that respond differently to various therapeutic modalities. Many advances in cancer genetics and treatment have come from basic and clinical research on soft tissue neoplasms.

In this chapter we discuss the general classification, epidemiology, pathogenesis, clinical features, histological features (including current grading systems), relevant immunohistochemical and ultrastructural findings, cytogenetics, molecular genetics and treatment of soft tissue sarcomas. Emphasis on the pathogenesis and molecular aspects of soft tissue sarcomas will fill the gap present in most textbooks on the subject. New and relevant concepts are discussed as appropriate. Owing to space limitations, clinicopathological features of exceedingly rare sarcomas arising in visceral locations and benign soft tissue tumours, including pseudosarcomatous lesions, are not discussed. The reader is referred to comprehensive reviews on the subject for additional information. For mesothelial tumours, including solitary fibrous tumours, see chapter on *Pleura and Peritoneum.*

CLASSIFICATION

The current World Health Organisation classification of soft tissue tumours is based primarily on clinicopathological features (Weiss and Sobin, 1994). The tumours are classified according to their resemblance to normal tissues or their line of differentiation and clinical behaviour (**Table 1**). The tumours are generally classified as benign or malignant (malignant tumours also are known as sarcomas, from the Greek words *sarx*, flesh, and *oma*, tumour). Many specific entities and variants exist within each category. Despite the efforts of several experts in soft tissue pathology to create a useful and rational classification, the current one, like any other, is not perfect. The field of soft tissue pathology is one of the most dynamic areas in diagnostic and experimental pathology and, because of tremendous technical advances in immunohistochemistry, cytogenetics and molecular genetics, previously unrecognized tumours have been described and fully characterized, and many similarities between apparently disparate entities have been identified. Therefore, future classifications are expected to incorporate this new information.

Table 1 World Health Organisation's histological classification of soft tissue tumours

1. Fibrous tissue tumour
Benign
Fibroma
Keloid[a]
Nodular fasciitis[a]
Proliferative fasciitis[b]
Proliferative myositis[b]
Elastofibroma[b]
Fibrous hamartoma of infancy[c]
Myofibromatosis, solitary and multicentric[c]
Fibromatosis colli[c]
Calcifying aponeurotic fibroma[c]
Hyaline fibromatosis[c]
Fibromatosis
Superficial fibromatosis
Palmar and plantar fibromatosis[a]
Infantile digital fibromatosis (digital fibroma)[c]
Deep fibromatosis
Abdominal fibromatosis (desmoid tumour)[a]
Extra-abdominal fibromatosis (desmoid tumour)[a]
Intra-abdominal and mesenteric fibromatosis[a]
Infantile fibromatosis[c]
Malignant
Fibrosarcoma
Adult fibrosarcoma[b]
Congenital or infantile fibrosarcoma[c]
2. Fibrohistocytic tumours
Benign
Fibrous histiocytoma
Cutaneous histocytoma (deramtofibroma)[a]
Deep histiocytoma[a]
Juvenile xanthogranuloma[c]
Reticulohistiocytoma[b]
Xanthoma[b]
Intermediate
Atypical fibroxanthoma[b]
Dermatofibrosarcoma protuberans[a]
Pigmented dermatofibrosarcoma protuberans (Bednar tumour)[a]

Table 1 (*Continued*)

- Giant cell fibroblastoma[c]
- Plexiform fibrohistiocytic tumour[a]
- Angiomatoid fibrous histiocytoma[a]
- Malignant
  - Malignant fibrous histiocytoma
  - Storiform-pleomorphic[b]
  - Myxoid[b]
  - Giant cell[b]
  - Xanthomatous (inflammatory)[b]

3. Lipomatous tumours

- Benign
  - Lipoma[a]
  - Lipoblastoma (foetal lipoma)[c]
  - Lipomatosis[a]
  - Angiolipoma[a]
  - Spindle cell lipoma[b]
  - Pleomorphic lipoma[b]
  - Angiomyolipoma[a]
  - Myelolipoma[b]
  - Hibernoma[b]
  - Atypical lipoma[b]
- Malignant
  - Well-differentiated liposarcoma[b]
  - Lipoma-like
  - Sclerosing
  - Inflammatory[b]
  - Myxoid liposarcoma[a]
  - Round cell (poorly differentiated myxoid) liposarcoma[b]
  - Pleomorphic liposarcoma[b]
  - Dedifferentiated liposarcoma[b]

4. Smooth muscle tumours

- Benign
  - Leiomyoma[a]
  - Angiomyoma[a]
  - Epithelioid leiomyoma[a]
  - Leiomyomatosis peritoneales disseminata[b]
- Malignant
  - Leiomyosarcoma[b]
  - Epithelioid leiomyosarcoma[b]

5. Skeletal muscle tumours

- Benign
  - Rhabdomyoma
  - Adult[b]
  - Genital[a]
  - Fetal[c]
- Malignant
  - Rhabdomyosarcoma
  - Embryonal rhabdomyosarcoma[c]
  - Botryoid rhabdomyosarcoma[c]
  - Spindle cell rhabdomyosarcoma[c]
  - Alveolar rhabdomyosarcoma[c]
  - Pleomorphic rhabdomyosarcoma[b]
  - Rhabdomyosarcoma with ganglionic differentiation (ectomesenchymoma)[c]

6. Endothelial tumours of blood and lymph vessels

- Benign
  - Papillary endothelial hyperplasia[a]
  - Haemangioma

Table 1 (*Continued*)

  - Capillary haemangioma[a]
  - Cavernous haemangioma[a]
  - Venous haemangioma[a]
  - Epithelioid haemangioma (angiolymphoid hyperplasia, histiocytoid haemangioma)[a]
  - Pyogenic granuloma (granulation tissue type haemangioma)[a]
  - Acquired tufted haemangioma (angioblastoma)[a]
  - Lymphangioma[a]
  - Lymphangiomyoma[a]
  - Lymphangiomyomatosis[a]
  - Angiomatosis[a]
  - Lymphangiomatosis[c]
- Intermediate: haemangioendothelioma
  - Spindle cell haemangioendothelioma[a]
  - Endovascular papillary angioendothelioma (Dabska tumour)[c]
  - Epithelioid hemangioendothelioma[a]
- Malignant
  - Angiosarcoma[b]
  - Lymphangiosarcoma[b]
  - Kaposi's sarcoma[b]

7. Perivascular tumours

- Benign
  - Benign haemangiopericytoma[a]
  - Glomus tumour[a]
- Malignant
  - Malignant haemangiopericytoma[b]
  - Malignant glomus tumour[b]

8. Synovial tumours

- Benign
  - Tenosynovial giant cell tumour
  - Localized[a]
  - Diffuse (extra-articular pigmented villonodular synovitis)[a]
- Malignant
  - Malignant tenosynovial giant cell tumour[b]

9. Mesothelial tumours

- Benign
  - Solitary fibrous tumour of pleura and peritoneum (localized fibrous mesothelioma)[b]
  - Multicystic mesothelioma[a]
  - Adenomatoid tumour[b]
  - Well-differentiated papillary mesothelioma[a]
- Malignant
  - Malignant solitary fibrous tumour of pleura and peritoneum (malignant localized fibrous mesothelioma)
  - Diffuse mesothelioma[a]
  - Epithelial
  - Spindled (sarcomatoid)
  - Biphasic

10. Neural tumours

- Benign
  - Traumatic neuroma[a]
  - Morton neuroma[b]
  - Neuromuscular hamartoma[c]
  - Nerve sheath ganglion[a]
  - Schwannoma (neurilemoma)[a]
  - Plexiform schwannoma[a]

Table 1 (*Continued*)

Cellular schwannoma[a]
Degenerated (ancient) schwannoma[b]
Neurofibroma
Diffuse[a]
Pacinian[a]
Plexiform[a]
Epithelioid[a]
Granular cell tumour[a]
Melanocytic schwannoma[b]
Neurothekeoma (nerve sheath myxoma)[a]
Ectopic meningioma[a]
Ectopic ependymoma[a]
Ganglioneuroma[a]
Pigmented neuroectodermal tumour of infancy (retinal enlage tumour, melanotic progonoma)[c]
Malignant
Malignant peripheral nerve sheath tumour (MPNST) (malignant shwannoma, neurofibrosarcoma)
MPNST with rhabdomyosarcoma (malignant Triton tumour)[a]
MPNST with glandular differentiation[a]
Epithelioid MPNST[a]
Malignant granular cell tumour[a]
Clear cell sarcoma (malignant melanoma of soft parts)[a]
Malignant melanotic schwannoma[a]
Neuroblastoma[c]
Ganglioneuroblastoma[c]
Neuroepithelioma (peripheral neuroectodermal tumour, peripheral neuroblastoma)[a]

11. Paraganglionic tumours
Benign
Paraganglioma[a]
Malignant
Malignant paraganglioma[a]

12. Cartilage and bone tumours
Benign
Panniculitis ossificans[b]
Myositis ossificans[a]
Fibrodysplasia (myositis) ossificans progressiva[c]
Extraskeletal chondroma[a]
Extraskeletal osteochondroma[a]
Extraskeletal osteoma[a]
Malignant
Extraskeletal chondrosarcoma[a]
Well-differentiated chondrosarcoma
Myxoid chondrosarcoma
Mesenchymal chondrosarcoma
Dedifferentiated chondrosarcoma
Extraskeletal osteosarcoma[b]

13. Pluripotential mesenchymal tumours
Benign
Mesenchymoma[a]
Malignant
Malignant mesenchymoma[a]

14. Miscellaneous tumours
Benign
Congenital granular cell tumour[c]
Tumoural calcinosis[a]
Myxoma
Cutaneous[a]
Intramuscular[b]
Angiomyxoma[a]
Amyloid tumour[b]
Parachordoma[a]
Ossifying fibromyxoid tumour[a]
Juvenile angiofibroma[c]
Inflammatory myofibroblastic tumour (inflammatory fibrosarcoma)[a]
Malignant
Alveolar soft part sarcoma[a]
Epithelioid sarcoma[a]
Extraskeletal Ewing sarcoma[a]
Synovial sarcoma[a]
Monophasic fibrous type
Malignant (extrarenal) rhabdoid tumour[c]
Desmoplastic small cell tumour of children and young adults[a]

15. Unclassified tumours

[a]Although this tumour is known to present more often in children than in adults, it is not strictly considered a neoplasm of either age group.
[b]More than 90% of cases present in patients beyond the third decade of life.
[c]Most tumours of this type present before the age of 20 years.
[From Weiss, 1994, *Histological Typing of Soft Tissue Tumours*, 2nd edn (Springer, Berlin), by permission of the publisher.]

EPIDEMIOLOGY, AETIOLOGY AND PATHOGENESIS

Soft tissue tumours are common if one considers the occurrence of benign and malignant tumours, the latter being 100 times less common than the former. Sarcomas represent approximately 1% of all malignant tumours in adults and 15% of all malignant tumours in children. According to a recent report, almost 8000 new cases of soft tissue sarcoma occur and more than 4000 patients die annually of this condition in the USA (Landis *et al.*, 1999). Soft tissue sarcomas are slightly more common in males and can occur anywhere in the body. The most common nonvisceral anatomical sites are the extremities and retroperitoneum.

The aetiology of soft tissue sarcomas is still largely unknown. Environmental and genetic factors have been associated with their development. Ionizing radiation for the treatment of other neoplasms, such as lymphomas and breast and cervical carcinomas, was found to induce high-grade sarcomas, most commonly osteosarcoma and malignant fibrous histiocytoma. Megavoltage and orthovoltage radiation have been implicated. Postradiation sarcomas usually arise 2–3 years after radiation therapy, and a higher risk is observed with cumulative doses greater than 10 Gy. The criteria proposed for the diagnosis of a postradiation sarcoma, which have been constantly modified over the years, require tumour development in a previously

unaffected area within the radiotherapy field and a minimum latency period of 2–3 years. However, the risk of radiation therapy is very low; postradiation sarcoma develops in only 0.1% of patients with cancer who survive more than 5 years after diagnosis.

Many chemicals have been implicated in the aetiology of sarcomas, including thorotrast, poly(vinyl chloride), arsenic, alkylating agents, phenoxyacetic acids, chlorophenols and dioxin (2,3,7,8-tetrachlorodibenzo-*p*-dioxin, or TCDD). The causal role of the three last-mentioned agents in the development of soft tissue sarcomas is controversial. The associations between the development of hepatic angiosarcomas and thorotrast, poly(vinyl chloride) and arsenic are better established.

Viruses have been implicated in the aetiology of a few sarcomas. The best example is the association between Kaposi sarcoma and the human herpes virus 8 (also known as Kaposi sarcoma-associated herpes virus, KSHV). KSHV has been detected in more than 95% of Kaposi sarcomas in all clinical settings. KSHV encodes proteins that disrupt cell cycle and apoptosis control mechanisms. Interference with the normal function of the retinoblastoma and p53 tumour-suppressor proteins as well as derepression of the c-*MYC* oncogene are caused by KSHV-derived proteins (Antman and Chang, 2000). KSHV also has been implicated in the pathogenesis of primary effusion lymphoma and multicentric Castleman disease.

Epstein–Barr virus has also been implicated in the pathogenesis of some sarcomas. Sequences of Epstein–Barr virus have been found in smooth muscle tumours that arise in children with acquired immune deficiency syndrome (AIDS) (McClain *et al*., 1995) and in the rare follicular dendritic cell sarcoma. More recently, the expression of the adenovirus early region 1A gene (*E1A*) in normal human fibroblasts and keratinocytes was found to induce the classical Ewing sarcoma translocation t(11;22)(q24;q12) with the formation of the fusion transcript *EWS/FLI1*. However, these intriguing results are unconfirmed.

A history of trauma is common in patients with soft tissue and bone sarcomas. However, the causal role of trauma is questionable. It is currently believed that trauma near or over the tumour area does not cause but rather leads to the discovery of the tumour. The short time between trauma and the diagnosis of sarcoma supports this view.

Chronic lymphoedema has been correlated with the development of lymphangiosarcoma (currently considered angiosarcoma) in different clinical settings. Stewart–Treves syndrome, the classical example, is characterized by the development of angiosarcoma in a lymphoedematous arm after radical mastectomy. Breast irradiation can also cause chronic arm lymphoedema and secondary angiosarcoma outside the irradiation field. Stewart–Treves syndrome is rarely observed today because of major modifications in the surgical and postsurgical management of patients with breast cancer. Filarial-induced lymphoedema has also been implicated in the development of angiosarcoma.

Several genetic disorders have been associated with an increased risk for the development of soft tissue sarcomas (see Section B). Li–Fraumeni syndrome, neurofibromatosis type I (von Recklinghausen disease), Gardner syndrome, familial retinoblastoma and Beckwith–Wiedemann syndrome are well-known examples. Li–Fraumeni syndrome is a rare familial syndrome characterized by an inherited predisposition to epithelial and nonepithelial tumours, including soft tissue sarcomas, osteosarcomas, breast and lung adenocarcinomas, medulloblastomas, adrenal cortical tumours, leukaemias and others. In 75% of cases, Li–Fraumeni syndrome results from germline mutations of the *TP53* gene on chromosome 17q13. However, germline mutations of the *CHK2* gene (*CH*eckpoint *K*inase *2*) have been described in a subset of patients with Li–Fraumeni syndrome (Bell *et al*., 1999). Neurofibromatosis type I is characterized by germline mutations of the *NF1* (*N*euro*F*ibromin) gene on chromosome 17q11.1. The NF1 protein normally inhibits the protein product of the *RAS* proto-oncogene (*RA*t *S*arcoma virus). Inactivation of *NF1* predisposes to neurofibromas, malignant peripheral nerve sheath tumours and glial tumours. Gardner syndrome is characterized by mutations of the *APC* gene (*A*denomatous *P*olyposis *C*oli) located on 5q21. Multiple intestinal adenomatous polyps develop in patients with Gardner syndrome; these patients have an increased risk for colon cancer and mesenteric fibromatosis (desmoid tumours). Patients with familial retinoblastoma carry a germline mutation in one of the *RB1* gene (*R*etino*B*lastoma *1*) alleles on chromosome 13q24. Loss of the wild-type allele on the homologous chromosome predisposes to retinoblastoma, osteosarcoma, pinealoma and soft tissue sarcomas. Beckwith–Wiedemann syndrome is a sporadic or autosomal-dominant disorder characterized by generalized overgrowth, hemihypertrophy and a tendency for neonatal hypoglycaemia, among other conditions, and an increased risk for Wilms tumour, adrenocortical carcinoma, hepatoblastoma and embryonal rhabdomyosarcoma. The syndrome seems frequently to be caused by loss of imprinting of the *IGFII* (*I*nsulin *G*rowth *F*actor-*II*) and *H19* genes on chromosome 11p15.5. Other disorders associated with an increased incidence of soft tissue sarcomas include tuberous sclerosis, Gorlin syndrome, Werner syndrome and Carney triad. (For additional information, see chapter on *Inherited Predispositions to Cancer*.)

GENETICS AND MOLECULAR BIOLOGY

Genetic alterations in soft tissue sarcomas can be tumour specific or nonspecific. Tumour-specific alterations include genetic alterations highly characteristic of certain soft tissue sarcomas and most commonly represented by chromosomal translocations (**Table 2**). Tumour-nonspecific alterations include genetic abnormalities seen

Table 2 Characteristic chromosomal and genetic abnormalities in sarcomas

Tumour type	Chromosomal abnormality	Fusion transcript or genetic abnormality	Prevalence (%)
Alveolar soft part sarcoma	t(X;17)(p11.2;q25)	*TFE3/ASPL*	>99
Alveolar rhabdomyosarcoma	t(2;13)(q35;q14)	*PAX3/FKHR*	75
	t(1;13)(p36;q14)	*PAX1/FKHR*	10
Angiomatoid fibrous histiocytoma	t(12;16)(q13;p11)	*TLS/ATF1*	?[a]
Clear cell sarcoma (malignant melanoma of soft parts)	t(12;22)(q13;q12)	*EWS/ATF1*	>85
Congenital fibrosarcoma/mesoblastic nephroma	t(12;15)(p13;q25)	*ETV6/NTRK3*	>99
Dermatofibrosarcoma protuberans /giant cell fibroblastoma	t(17;22)(q22;q13)	*COL1A1/PDGFB*	>99
Desmoplastic small round cell tumour	t(11;22)(p13;q12)	*EWS/WT1*	>99
Epithelioid hemangioendothelioma	t(1;3)(p36.3;q25)	*varNOW*	?
Ewing sarcoma/peripheral neuroectodermal tumour[b]	t(11;22)(q24;q12)	*EWS/FLI1*	95
	t(21;22)(q22;q12)	*EWS/ERG*	5
	t(7;22)(p22;q12)	*EWS/ETV1*	<1
	t(17;22)(q12;q12)	*EWS/E1AF*	<1
	t(2;22)(q33;q12)	*EWS/FEV1*	<1
	t(1;22)(p36.1;q12)	*EWS/ZSG*	<1
Extraskeletal myxoid chondrosarcoma	t(9;22)(q22;q12)	*EWS/CHN*	75
	t(9;17)(q22;q11)	*RBP56/CHN*	25
Inflammatory myofibroblastic tumour	2p23 rearrangements	*TPM3/ALK*	?
		TPM4/ALK	?
Malignant rhabdoid tumour	del(22q11.2)	*hSNF/INI1* (deletion or mutation)	?
Myxoid/round cell liposarcoma	t(12;16)(q13;p11)	*TLS/CHOP*	95
	t(12;22)(q13;q12)	*EWS/CHOP*	5
Synovial sarcoma	t(X;18)(p11.2;q11.2)	*SYT/SSX-1*	65
		SYT/SSX-2	35
		SYT/SSX-4	<1

[a]Rearranged in a single case.
[b]Data from bone tumours.
Adapted from Ladanyi and Bridge, 2000, by permission of W. B. Saunders.

in various tumours and often associated with the clonal evolution of the tumour.

Since the discovery of the translocation t(11;22)(q24;q12) in Ewing sarcoma/peripheral primitive neuroectodermal tumour (PNET) in 1983 (Aurias *et al*., 1983; Turc-Carel *et al*., 1983), an explosion of new knowledge has contributed to our understanding of the molecular biology of soft tissue tumours. Many soft tissue sarcomas are characterized by specific, recurrent, balanced chromosomal translocations that result in the fusion of two constitutional genes into a new chimaeric aberrant gene. These chimaeric genes usually function as altered DNA transcription factors and seem to participate in the pathogenesis of these tumours. The fusion between the *EWS* gene (*EW*ing *S*arcoma) on chromosome 22q12 with the *FLI1* gene (*F*riend *L*eukaemia virus *I*ntegration *1*) on chromosome 11q24 is the best-known example (de Alava and Gerald, 2000). *EWS* encodes for a ubiquitously expressed protein involved in mRNA transcription. The structure of *EWS* protein includes an N-terminal domain homologous to the eukaryotic RNA polymerase II and a C-terminal domain that contains an RNA-binding sequence. *FLI1* encodes for a protein member of a large family of DNA-binding transcription factors that contain a highly conserved amino acid sequence, the ETS domain (*E*rythroblastosis virus-*T*ransforming *S*equence), in the C-terminal end. During the translocation process, the 5′ end of the *EWS*, which contains the *EWS* promoter region, fuses with the 3′ end of the *FLI1*, which contains the ETS domain, resulting in a potent DNA transcription factor (Sandberg and Bridge, 2000) (**Figure 1**).

Similar translocation mechanisms occur in other sarcomas and often involve similar genes (**Table 2**). For example, *EWS* fuses with several ETS family members in other translocations observed in Ewing sarcoma/PNET and related tumours: *ERG* (*E*TS-*R*elated *G*ene), *ETV-1* (*E*TS *T*ranslocation *V*ariant-*1*), *E1AF* (*E1A* *F*actor), *FEV* (*F*ifth *E*wing *V*ariant) and *ZSG* (*Z*inc finger *S*arcoma *G*ene). In addition, *EWS* is fused with other gene partners in clear cell sarcoma (melanoma of soft parts), desmoplastic small round cell tumour, myxoid liposarcoma and extraskeletal myxoid chondrosarcoma (**Table 2**). Other genes involved in sarcoma and leukaemia chromosomal translocations show striking similarities to *EWS*. *FUS/TLS*, *RPB56/hTAFII68*, and *EWS* genes, which contribute their 5′ end to many chimaeric

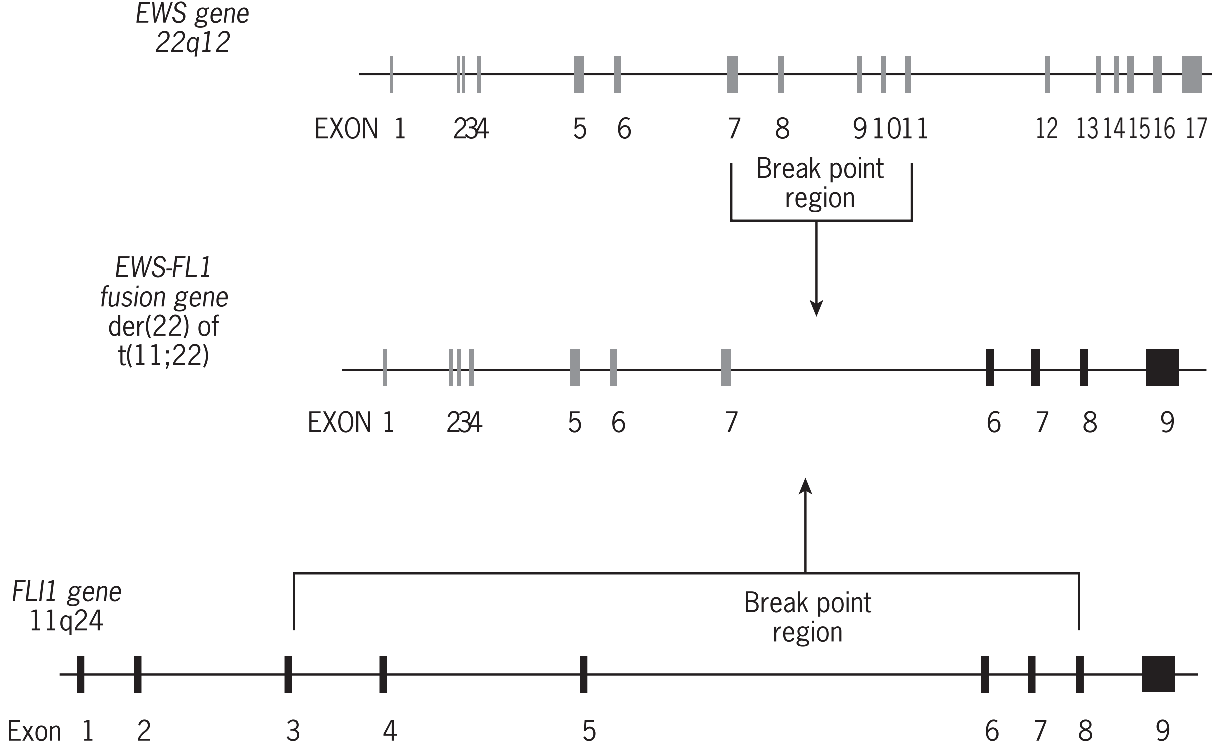


Figure 1 Schematic diagram representing the fusion of the *EWS* gene on chromosome 22q12 with the *FLI1* gene on chromosome 11q24. Fusions of exon 7 of *EWS* with exon 6 (type I) and exon 7 (type II) of *FLI1* occur in 70–75 and 15–25% of cases, respectively. *EWS/FLI* type I gene has been correlated with a low proliferative rate and prolonged survival in Ewing sarcoma. (Adapted from de Alava and Gerald, 2000.)

genes, have similar structures and seem to be originated from the same ancestor gene (**Figure 2**). *FUS/TLS* or *EWS* fuses with the *CHOP* gene in myxoid liposarcoma, and *EWS* or *RPB56/hTAFII68* fuses with the *CHN* gene in extraskeletal myxoid chondrosarcoma. *FUS/TLS* also fuses with the *ERG* gene in acute myeloid leukaemia. Other chromosomal translocations generate chimaeric proteins that function as growth factors (see Dermatofibrosarcoma Protuberans and Giant Cell Fibroblastoma, below) and tyrosine kinase proteins (see Congenital Fibrosarcoma, below).

The underlying mechanism responsible for the genesis of these translocations is still unknown. However, some clues have been found. DNA sequences near the breakpoints of the *FUS/TLS* and *CHOP* genes in myxoid liposarcoma and in the *PAX3* and *FKHR* in alveolar rhabdomyosarcoma are homologous to translin– and topoisomerase II–DNA binding sequences, suggesting the involvement of these proteins or their DNA binding sequences for the genesis of the translocations. In addition, a LINE-1 DNA sequence was found to be present in the translocation t(11;22)(p13:q24) breakpoint in desmoplastic small round cell tumour. It has been proposed that a series of constraint factors collaborate at the level of chromosomal rearrangements, gene expression and gene product function to determine the acquisition and maintenance of a specific translocation (Barr, 1998).

Inactivation of the tumour-suppressor gene *TP53* is commonly observed in soft tissue sarcomas. Loss of heterozygosity of 17p and missense mutations of the *TP53* are the most common mechanisms, occurring in 30–50% of cases. Overexpression of the p53 protein occurs after DNA damage and results in G1/S cell cycle arrest, induction of DNA repair systems, and apoptosis activation (see the chapters *Regulation of the Cell Cycle* and *Genetic Instability and DNA Repair*). Missense mutations of *TP53* result in an abnormal p53 protein that has a prolonged half-life and accumulates above the levels for immunohistochemical detection. Immunoreactivity for mutant p53 protein was found in approximately 27% of cases in a series of 211 soft tissue sarcomas and correlated with decreased survival (Cordon-Cardo, 1994). In the same series, mutant p53 protein expression did not always correlate with *TP53* mutation. Other studies have found a correlation between mutant p53 protein and high-grade histology or poor survival.

The tumour-suppressor gene *RB1* is commonly inactivated in soft tissue sarcomas. Normally, the hypophosphorylated Rb1 protein forms a complex with the transcription factor E2F, arresting the cell cycle in the G1/S checkpoint. Complexes of cyclins and cyclin-dependent kinases phosphorylate Rb1, releasing E2F and promoting cell cycle progression (see the chapter *Regulation of the Cell Cycle*). Loss of Rb1 protein

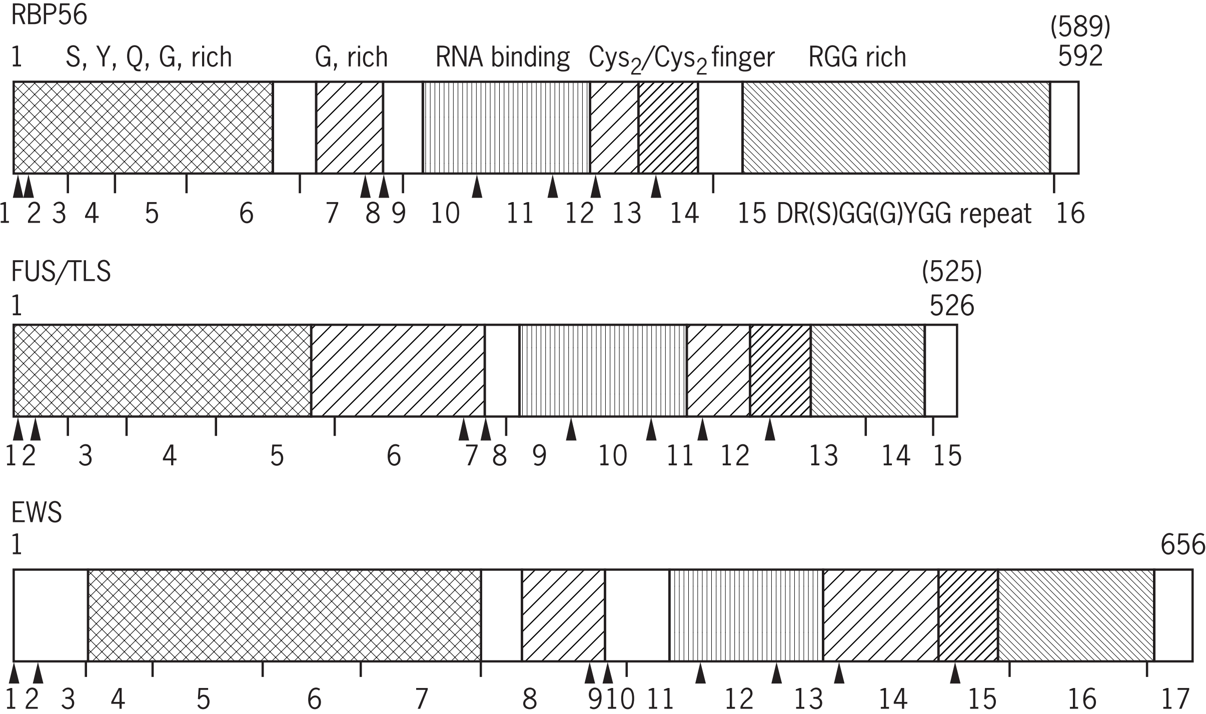


Figure 2 Structural similarities of RBP56, FUS/TLS and EWS proteins suggest that the respective genes originated from the same ancestral gene. *RBP56(hTAFII68)* is rearranged in extraskeletal myxoid chondrosarcoma. *TLS* is rearranged in myxoid liposarcoma and some acute leukaemias. A single case of angiomatoid malignant fibrous histiocytoma also showed a *TLS* rearrangement. *EWS* is rearranged in Ewing sarcoma, clear cell sarcoma, myxoid liposarcoma, extraskeletal myxoid chondrosarcoma and desmoplastic small round cell tumour. (From Morohoshi *et al*., 1998, Genomic structure of the human RBP56/hTAFII68 and FUS/TLS genes. *Gene*, **221**, 191–198, copyright 1998, with permission from Elsevier Science.)

expression has been correlated with a worse prognosis in some (Cance *et al*., 1990) but not all series of soft tissue sarcomas (see Prognostic Factors, below).

MDM2 (*M*urine *D*ouble *M*inute *2*) gene amplification and protein overexpression are commonly observed in soft tissue sarcomas (Oliner *et al*., 1992). *MDM2* is mapped on 12q13 and has its transcription regulated by the p53 protein in a negative feedback loop. Mdm2 protein seems to form a trimeric complex with both p53 and Rb1. Mdm2 protein inhibits apoptosis and stimulates cell cycle progression by inactivating the p53 protein. *MDM2* gene amplification is not always correlated with Mdm2 protein overexpression detected by immunohistochemical analysis. Overexpression of Mdm2 protein was correlated with a worse prognosis in large series of patients with soft tissue sarcomas.

Additional genetic alterations have been observed in soft tissue sarcomas. On the same region of *MDM2* (12q13-15), other genes, including *HMGI-C*, *SAS*, *GLI* and *CDK4*, were found to be amplified. The *HMGI-C* gene (*H*igh *M*obility *G*roup Protein *I-C*) encodes for a DNA transcription factor. *HMGI-C* has been found to be rearranged or amplified in lipomas and atypical lipomatous tumours (well-differentiated liposarcomas). The *SAS* gene (*S*arcoma *A*mplified *S*equence), a member of the transmembrane 4 superfamily of proteins, is involved in cell growth regulation and was found to be amplified in malignant fibrous histiocytomas and liposarcomas. The *GLI* gene (*GLI*oma) encodes a transcription factor and is amplified in gliomas and high-grade childhood sarcomas. *CDK4* (*C*yclin *D*ependent *K*inase *4*) encodes a cell cycle protein kinase, and its amplification may favour cell proliferation. Deletions or rearrangements, but not point mutations, of the genes *INK4A* (*IN*hibitor of Cyclin Dependent *K*inase *4A*, which encodes for the cell cycle inhibitors p16 and p19ARF) and *INK4B* (which encodes for the cell cycle inhibitor p15) on the chromosome 9p21 were found in high-grade sarcomas and correlated with poor survival after adjustment for size and grade. Mutations of K-*RAS* (*K*irsten) and H-*RAS* (*H*arvey) oncogenes were found in a few series of soft tissue sarcomas and were more common in malignant fibrous histiocytoma. Interestingly, these mutations are common in Korean patients and rare in American patients.

HER2/neu (*H*uman *E*pidermal growth factor *R*eceptor *2*) gene amplification or protein overexpression as well as microsatellite instability seem uncommon in soft tissue sarcomas, but the data are limited. (See Synovial Sarcoma later in this chapter for additional information on HER2/neu).

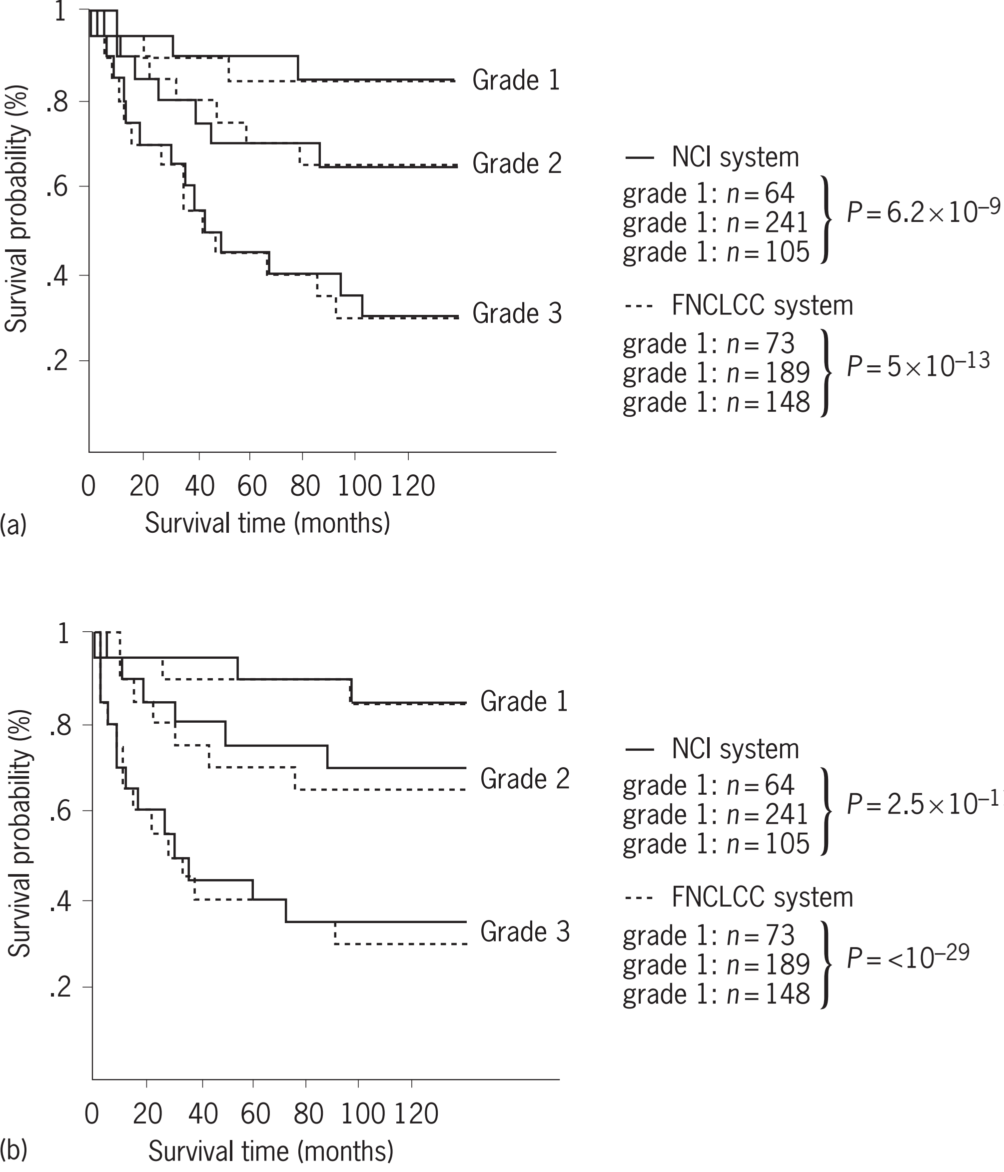


Figure 3 Survival curves according to histologic grade using the National Cancer Institute (NCI) and French Federation of Cancer Centres (FNCLCC) systems. (a) Overall survival; (b) metastasis-free survival. (From Guillou *et al.*, 1997, by permission of the American Society of Clinical Oncology.)

GRADING AND STAGING SYSTEMS

Many systems have been proposed for the grading and staging of soft tissue sarcomas. Their main goal is to identify patient groups with comparable clinicopathologic features for treatment optimization, prognostication and data comparison among cancer centres.

Histological grading has been considered the most powerful independent prognostic factor for soft tissue sarcomas and is based on the assessment of morphological features. Degree of tumour differentiation (resemblance to normal tissue counterparts), mitotic activity, necrosis and cellularity are histological variables commonly evaluated during the grading process. Currently, the French Federation of Cancer Centres (Trojani *et al.*, 1984) and the National Cancer Institute (Costa *et al.*, 1984) grading systems are the most commonly used and provide similar results (**Figure 3**) (Guillou *et al.*, 1997). The French Federation of Cancer Centres uses a 3-grade system that combines tumour differentiation, mitotic count, and necrosis to create a score (**Table 3**). The NCI also uses a three-grade system based on histological subtype and the amount of necrosis. The prognostic information obtained with each system has been validated in many reports but problems remain. Subjectivity on the assessment of some histologic variables, uneven representation of specific

Table 3 French Federation of Cancer Centres system for sarcoma grading

A. Tumour differentiation score[a]	B. Tumour necrosis score[b]
1 Well-differentiated liposarcoma	0 No necrosis
Well-differentiated fibrosarcoma	1 ≤50% tumour surface examined
Well-differentiated MPNST	2 >50% tumour surface necrosis
Well-differentiated leiomyosarcoma	
Well-differentiated chondrosarcoma	
2 Myxoid liposarcoma	C. Mitotic activity score
Conventional fibrosarcoma	1 0–9 mitoses/10 HPF
Conventional MPNST	2 10–19 mitoses/10 HPF
Well-differentiated malignant haemangiopericytoma	3 ≥;20 mitoses/10 HPF
Myxoid MFH	
Typical storiform/pleomorphic MFH	
Conventional leiomyosarcoma	
Myxoid chondrosarcoma	
Conventional angiosarcoma	
3 Round cell liposarcoma	Final grade
Pleomorphic liposarcoma	Score A + score B + score C
Dedifferentiated liposarcoma	
Poorly differentiated fibrosarcoma	
Poorly differentiated MPNST	Score
Epithelioid malignant schwannoma	2–3 Grade 1 sarcoma
Malignant Triton tumour	4–5 Grade 2 sarcoma
Conventional malignant haemangiopericytoma	6–8 Grade 3 sarcoma
Giant cell and inflammatory MFH	
Poorly differentiated/pleomorphic/epithelioid leiomyosarcoma	
Synovial sarcoma (any subtype)	
Rhabdomyosarcoma (any subtype)	
Mesenchymal chondrosarcoma	
Poorly differentiated/epithelioid angiosarcoma	
Extraskeletal osteosarcoma	
Extraskeletal Ewing sarcoma/PNET	
Alveolar soft part sarcoma	
Malignant rhabdoid tumour	
Clear cell sarcoma	
Undifferentiated sarcoma	

[a]Tumours showing close similarity to normal adult tissue are scored 1, sarcomas of certain histological sybtype are scored 2 and sarcomas of uncertain histological subtype are scored 3.
[b]Estimated by microscopic review only.
HPF, high-power field (1 HPF = 0.174 mm^2); MFH, malignant fibrous histiocytoma; MPNST, malignant peripheral nerve sheath tumour; PNET, peripheral neurectodermal tumour.
Data from Trojani *et al.* (1984).
Adapted from Guillou *et al.*, 1997, by permission of the American Society of Clinical Oncology.

sarcomas and lack of standardization of cutoff points among the reported series are factors that preclude universal acceptance of a particular grading system. Because soft tissue sarcomas are rare, prospective multicentre studies are needed to determine useful histological variables for each type of sarcoma.

The two most common staging systems are the American Joint Committee on Cancer/Union Internationale Contre le Cancer (AJCC/UICC) staging system (Fleming *et al.*, 1997) (**Table 4**) and the Enneking system (Enneking *et al.*, 1980). These systems are chiefly designed for the staging of adult soft tissue sarcomas. The AJCC/UICC system is a four-stage system based on tumour size, presence or absence of lymph node involvement and distant metastases and histological grade. Approximate 5-year overall survival rates for patients with soft tissue sarcomas in stages I, II and III are 95, 80, and 45%, respectively (**Figure 4**) (Pisters and Brennan, 2000). Important limitations of the AJCC/UICC system are the inclusion of lymph node status and the noninclusion of the anatomic site or the adequacy of the surgical margins. The first is a limitation in that lymph node metastases occur infrequently (2.6–5.9%) in sarcomas, in contrast to epithelial tumours. However, epithelioid sarcoma, angiosarcoma and rhabdomyosarcoma often do metastasize to lymph nodes. Most sarcomas metastasize hematogenously, and distant metastases are often present when lymph node involvement is discovered. Anatomical site is omitted in the AJCC/UICC system. Retroperitoneal, visceral and head and neck sarcomas are less likely to be completely resected without affecting vital structures; these

Table 4 American Joint Committee on Cancer/Union Internationale Contre le Cancer Staging System (Fleming *et al.*, 1997)

Tumour (T)	
T1 tumours: ≤5 cm (T1a, T1b)	a, superficial to superficial fascia
T2 tumours: >5 cm (T2a, T2b)	b, deep to superficial fascia
Grade (G)	Stage I
G1: well-differentiated	IA: G1/G2 T1a/b N0 M0
G2: moderately differentiated	IB: G1/G2 T2a N0 M0
G3: poorly differentiated	
G4: undifferentiated	Stage II
	IIA: G1/G2 T2b N0 M0
Lymph nodes (N)	IIB: G3/G4 T1a/b N0 M0
N0: no involvement	IIC: G3/G4 T2a N0 M0
N1: involved	
	Stage III
Metastasis (M)	G3/4 T2b N0 M0
M0: no metastasis	
M1: metastasis present	Stage IV
	IVA: Any G Any T N1 M0
	IVB: Any G Any T Any N M1

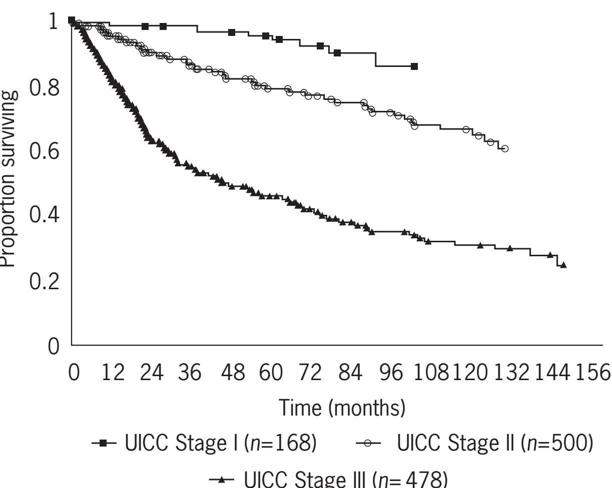


Figure 4 Overall survival for 1146 patients with primary soft tissue sarcoma according to the Union Internationale Contre le Cancer (UICC) staging system. Patients were treated at Memorial Sloan-Kettering Cancer Center, New York. (From Pisters and Brennan, 2000, by permission of Churchill Livingstone.)

sarcomas tend to be associated with a worse prognosis (see Treatment, below).

The Enneking system is surgical staging based on anatomical compartments and histological grade (Enneking *et al.*, 1980). The primary tumour is considered intracompartmental or extracompartmental according to its confinement within well-defined anatomic boundaries. Each of the system's three stages carries its own prognosis.

DIAGNOSIS

The diagnosis of soft tissue sarcoma requires a multidisciplinary approach. Physical examination, imaging studies, histological evaluation and ancillary techniques such as immunohistochemical analysis, electron microscopy, and cytogenetics molecular studies are important steps toward a correct diagnosis.

Clinical History and Physical Examination

Clinical findings alone are not specific for the diagnosis of a soft tissue tumour, but some clues point toward the possibility of a sarcoma. Information on age, sex, tumour location, the presence of genetic syndromes and the duration and quality of symptoms narrow the differential diagnosis substantially. In general, a soft tissue sarcoma presents as a painless mass of a few months' duration. Short duration of symptoms suggests a reactive process, such as that observed in nodular fasciitis (a pseudosarcomatous lesion commonly seen in the upper extremities of young patients). In contrast, long duration of symptoms suggests a benign tumour. Pain, paresthesias and paresis are observed when there is direct involvement or compression of major neurovascular structures and can occur with benign and malignant tumours. Organ-specific dysfunction also may occur due to the same causes. For example, large retroperitoneal liposarcomas can lead to renal dysfunction due to ureter compression. Paraneoplastic syndromes are seen rarely; examples include fever, weight loss, anaemia and paraproteinaemia occasionally seen in angiomatoid fibrous histiocytoma (see Angiomatoid Fibrous Histiocytoma, below); increased hepatic transaminases in the absence of liver metastases in patients with malignant peripheral nerve sheath tumour (Stauffer syndrome); osteomalacia in patients with phosphaturic mesenchymal tumour; and hypoglycemia in patients with solitary fibrous tumour. Patients who present with metastatic disease have symptoms and signs according to the location of the metastases.

Imaging Studies

After the initial clinical assessment, imaging studies often are needed for better delineation of soft tissue masses and their relation to surrounding normal structures. Simple radiographs have a limited role in the evaluation of soft tissue masses. However, simple radiographs can detect a soft tissue calcification or ossification, involvement of adjacent bone, or a possible bone origin of a soft tissue tumour (Kransdorf and Murphey, 1997). Computed tomography (CT) has been largely replaced by magnetic resonance imaging (MRI) in the evaluation of soft tissue tumours. Nonetheless, CT remains the best imaging technique for detecting pulmonary metastases, providing essential information for staging and

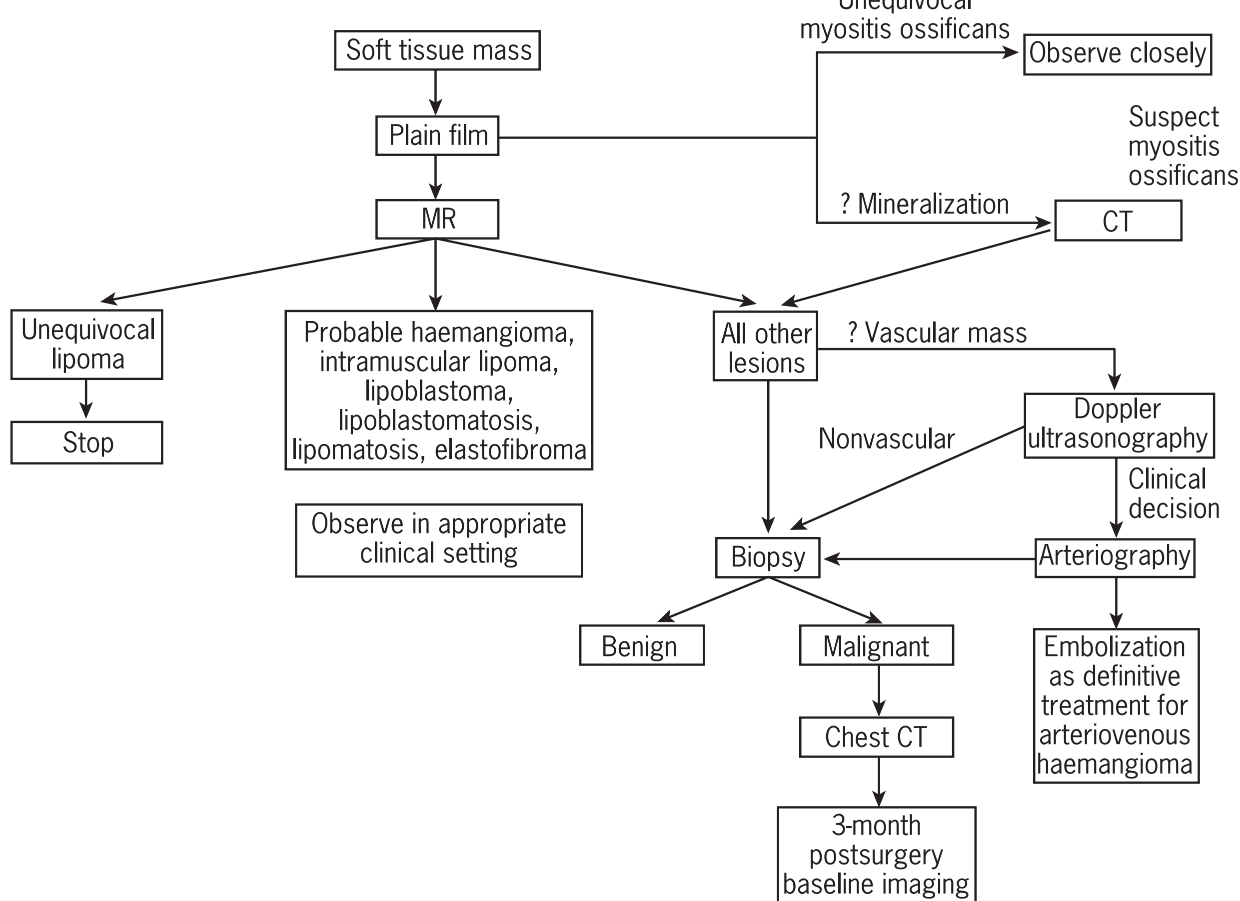


Figure 5 Algorithm for imaging evaluation of soft tissue masses proposed by Kransdorf and Murphey. CT, computed tomography; MR, magnetic resonance. (From Kransdorf and Murphey, 1997, by permission of WB Saunders.)

management. CT is superior in detecting calcification or ossification in soft tissue masses, such as zonal ossification in myositis ossificans, and has higher sensitivity than simple radiographs in detecting subtle bone involvement. Ultrasonography is useful for distinguishing between a solid and a cystic lesion and between soft tissue oedema and a soft tissue mass; it is also useful for guiding percutaneous biopsy. MRI is considered the gold standard because it provides precise information regarding tumour size, depth, local extension and relationship to normal structures such as vessels, bone and nerves. However, MRI is not usually diagnostic. Soft tissue lesions for which MRI can allow diagnosis include haemangioma, myositis ossificans, elastofibroma, fibromatosis, lipoma and liposarcoma, ganglion, haematoma, neurofibroma and malignant peripheral nerve sheath tumour (MPNST). Scintigraphy usually is not needed except to differentiate bone metastases from soft tissue tumours. Other imaging techniques are seldom used at present. An algorithm for imaging evaluation of soft tissue masses is shown in **Figure 5** (Kransdorf and Murphey, 1997). (Additional information can be found in the section on Diagnostic Imaging and Image-Guided Intervention.)

Biopsy and Basic Histological Assessment

Biopsy is usually the final step in the evaluation of soft tissue masses (see section on *Diagnostic Imaging*). It is generally advisable to perform a biopsy on any enlarging soft tissue mass (larger than 4–5 cm) with a median duration of 4–8 weeks in an adult patient (Pisters and Brennan, 2000). The biopsy can be closed or open. Closed procedures include core-needle biopsy and fine-needle aspiration (FNA). Open procedures include incisional and excisional biopsies. In incisional biopsies only part of the tumour is resected; in excisional biopsies the entire lesion is removed. Closed procedures, particularly core-needle biopsies, are the preferred method in most centres. FNA also is routinely used in many institutions and usually provides satisfactory results. In addition, ancillary techniques can be performed on FNA-obtained material. However, the most important role of FNA is documentation of tumour recurrence or metastasis. The major disadvantages of core biopsies and FNA are that tumour sampling is inadequate for precise histological analysis and, when

preoperative treatment with radiotherapy or chemotherapy is contemplated, a limited amount of tumour material is available for future studies. In these latter cases, the tumour is often largely necrotic after preoperative treatment and has few or no viable cells. Open procedures, such as excisional biopsies, are sometimes indicated for small superficial lesions, usually smaller than 4–5 cm. Imaging techniques such as CT and ultrasonography are important for guiding biopsy procedures.

Assessment of tumour type and histological grade is performed on paraffin-embedded, formalin-fixed material stained with haematoxylin and eosin. However, frozen material stained with haematoxylin and eosin or toluidine blue is commonly used in some institutions for a fast preliminary diagnosis.

Immunohistochemical Analysis

Immunohistochemical analysis has been extensively used in the study of soft tissue tumours and is primarily based on enzymatic methods, which work well on formalin-fixed material. The avidin–biotin–enzyme complex is one of the most commonly used methods. More than 30 markers are routinely used to evaluate soft tissue tumours, mainly for diagnostic purposes. Vimentin is a 57-kDa cytoplasmic intermediate filament present in almost all soft tissue tumours but also in some carcinomas and melanomas. In the analysis of soft tissue tumours, vimentin primarily is used as an internal control marker; its presence indicates preservation of tissue antigenicity. It is important to emphasize that vimentin is not useful to distinguish between sarcomas and other tumour types.

Cytokeratins are much more useful than vimentin for differentiating soft tissue sarcomas from carcinomas. However, cytokeratin expression is found in up to 5% of soft tissue sarcomas and is considered diagnostic for certain sarcomas. Well-known examples of sarcomas that express cytokeratins are synovial sarcoma, epithelioid sarcoma, desmoplastic small round cell tumour, MPNST, angiosarcoma, leiomyosarcoma, epithelioid haemangioendothelioma and rhabdoid tumour.

Epithelial membrane antigen (EMA) is a complex of proteins isolated from the human milk fat globule membrane. EMA is commonly expressed in synovial sarcoma, epithelioid sarcoma, MPNST, and other soft tissue tumours.

Muscle markers include desmin, muscle-specific actin (MSA HHF35), alpha smooth muscle actin (SMA 1A4), MyoD1 (also known as Myogenic Factor 3), myoglobulin, myogenin, calponin and caldesmon. Desmin is a 52-kDa intermediate filament expressed in leiomyosarcoma, rhabdomyosarcoma and other tumours of muscle differentiation. MSA comprises alpha (skeletal, smooth and cardiac) and gamma smooth muscle actin. MSA is expressed in more than 90% of leiomyosarcomas and 95% of rhabdomyosarcomas. SMA also is expressed in smooth muscle neoplasms such as leiomyosarcoma but is seldom expressed in rhabdomyosarcoma (1–2%). SMA expression frequently indicates the presence of myofibroblastic or myoid differentiation in nonmuscle tumours. MyoD1 is a 45-kDa protein member of a group of myogenic regulatory nuclear proteins that also includes myogenin, myf5 and mrf4-herculin/myf-6. MyoD1 is expressed in an early stage of skeletal muscle differentiation and is found primarily in rhabdomyosarcoma. Myoglobulin is a highly specific but not sensitive marker for rhabdomyoblastic differentiation, considering that it is primarily expressed in more differentiated rhabdomyosarcomas.

CD31, Factor VIII-related antigen (Factor VIIIRa) and CD34 are commonly used markers for endothelial differentiation. CD31 (glycoprotein gpIIa) is a membrane protein from the immunoglobulin supergene family. CD31 shows the highest sensitivity and specificity for endothelial differentiation, being detected in more than 80% of angiosarcomas. Factor VIIIRa is less sensitive than CD31 but is a useful marker for endothelial differentiation. CD34 (also known as haematopoietic progenitor cell antigen) is a 110-kDa transmembrane protein primarily expressed by human haematopoietic stem cells and endothelial cells. CD34 is sensitive but not specific for endothelial differentiation. CD34 expression is also of diagnostic importance in nonendothelial tumours such as DFSP, giant cell fibroblastoma, solitary fibrous tumour, and gastrointestinal stromal tumour.

S-100 protein, which is principally expressed by melanocytic tumours, also is expressed by soft tissue tumours, such as clear cell sarcoma (malignant melanoma of soft parts; 90% of tumours) and MPNST (50–70% of tumours). Nonetheless, S-100 expression has also been identified in other soft tissue sarcomas. CD68 is a 110-kDa lysosomal glycoprotein found mainly in monocytes and macrophages. In the past, CD68 and Factor XIIIa were considered markers of histiocytic differentiation. However, studies have shown the lack of specificity of both markers for this aim, limiting their diagnostic applicability. In addition, expression of CD68 and factor XIIIa by tumour-infiltrating histiocytes can lead to false-positive results.

p30/32 MIC2 (*M*onoclonal *I*mperial *C*ancer Research Fund *2*) or CD99 is a 32-kDa cell membrane glycoprotein encoded by the *MIC2* pseudoautosomal gene located on chromosome Xpter-p22.32. MIC2 is expressed by more than 95% of Ewing sarcoma/PNET. However, MIC2 also is expressed by other neoplasms, such as lymphomas, rhabdomyosarcomas, synovial sarcoma and mesenchymal chondrosarcoma.

Recently, immunohistochemical techniques have been developed to detect abnormal chimaeric proteins or amino acid sequences associated with specific sarcomas. It is expected that they will be used commonly in the future. Prognostic markers detected by immunohistochemical analysis include Ki-67, p53, Mdm2, Rb1 and p-glycoprotein (see Prognostic Factors, below).

Ultrastructural Analysis

With the advent of immunohistochemical and molecular techniques, the role of ultrastructural analysis in the diagnosis of soft tissue tumours is more limited but still important. Generally, ultrastructural analysis is required when histological, immunohistochemical and cytogenetic and molecular studies are inconclusive. Ultrastructural studies are useful in the identification of poorly differentiated rhabdomyosarcoma, the differentiation of sarcomatoid carcinoma or spindle cell melanoma from spindle cell sarcomas such as MPNST and monophasic synovial sarcoma and the identification of atypical forms of alveolar soft part sarcoma. Electron microscopy remains an important research tool, however, providing ultrastructural support for new morphological and immunohistochemical findings in many of the previously and most of the recently recognized tumours. For example, neuroendocrine differentiation in extraskeletal myxoid chondrosarcoma was confirmed through detection of neurosecretory granules by electron microscopy.

Cytogenetics, Molecular Cytogenetics and Molecular Genetics

Cytogenetic and molecular techniques are important adjuvant tools in the diagnosis of soft tissue sarcomas. Traditional cytogenetic evaluation has provided fundamental information regarding chromosomal abnormalities in soft tissue sarcomas. However, this technique depends on the availability of fresh viable tumour material and is time- and labour-intensive. Therefore, it is of limited clinical use in many centres.

Fluorescence *in situ* hybridization (FISH) was a considerable technical advance for chromosomal analysis. FISH relies on the use of fluorochrome-labelled complementary DNA (cDNA) probes that hybridize to specific DNA sequences. FISH can be performed not only in fresh tumour material but also in frozen and paraffin-embedded tissue. Both metaphase chromosomal spreads and interphase cells can be studied. In addition, FISH provides direct cell visualization, allowing correlation between histological and cytogenetic features. *In situ* hybridization can also be performed using enzymatic methods, such as the colorimetric *in situ* histochemistry (CISH).

Another important technique in the diagnosis of soft tissue sarcomas is the reverse transcriptase–polymerase chain reaction (RT-PCR), a variant of PCR. RT-PCR is used for detecting abnormal chimaeric mRNA transcripts on frozen and paraffin-embedded material. RT-PCR is preferred to standard PCR because chimeric mRNAs have more structural consistency than native chimeric genes owing to splicing of intronic sequences. More recently, RT-PCR has been tested to detect minimal disease in the peripheral blood in patients with Ewing sarcoma/PNET, rhabdomyosarcoma and myxoid liposarcoma. Southern blot is occasionally used but requires large amounts of frozen material (Ladanyi and Bridge, 2000). Because these methods have advantages and disadvantages, they should be viewed as complementary diagnostic tools. **Table 2** lists some chromosomal and genetic abnormalities that can be investigated with the above techniques.

CLINICOPATHOLOGICAL FEATURES

In this section, we discuss clinicopathological, immunohistochemical, ultrastructural and genetic features of specific sarcomas following the same schema (with some modifications) as used by the World Health Organisation for classification of soft tissue tumours. Rare sarcomas are not discussed; the reader is referred to comprehensive references on the subject.

Tumours of Fibroblastic and Myofibroblastic Differentiation

Fibromatoses

Fibromatoses are locally aggressive lesions without metastatic potential. They are broadly divided into superficial and deep (desmoid type). The main types of superficial fibromatosis are palmar fibromatosis (Dupuytren disease), plantar fibromatosis (Ledderhose disease) and penile fibromatosis (Peyronie disease). Fibromatoses usually occur in adults and can be bilateral. They are characterized by slow nodular fibroblastic and myofibroblastic proliferations that arise in fascias and aponeuroses. Immunohistochemical analysis shows the expression of vimentin and actin. A weak association of Dupuytren disease with alcoholism and epilepsy has been suggested. Cytogenetic studies often show trisomies 7, 8 and 14.

Deep fibromatosis or desmoid tumours are divided anatomically into three main types: extra-abdominal, abdominal and intra-abdominal. They frequently occur in young adults, and there is a female predominance. These tumours can be sporadic or familial. Intra-abdominal (mesenteric) fibromatosis occurs in 5–15% of patients with Gardner syndrome. Abdominal desmoid tumours are more common in females and tend to occur during or after pregnancy. Desmoid tumours commonly present as large tumours, usually exceeding 5 cm. They are characterized by a fascicular fibroblastic and myofibroblastic proliferation with infiltrative borders. Desmoid tumours are considered clonal lesions. Trisomies 8 and 20 and deletion of 5q (region of the *APC* gene) are common cytogenetic abnormalities. Overexpression of β-catenin has been implicated in the pathogenesis of desmoid tumours and seems related to mutations in the *APC* gene. In addition, β-catenin gene mutations have been found in sporadic desmoid tumours.

Clinically, desmoid tumours are characterized by high rates of recurrence (40%; 90% in some old series). Surgical excision followed or not by radiotherapy has been

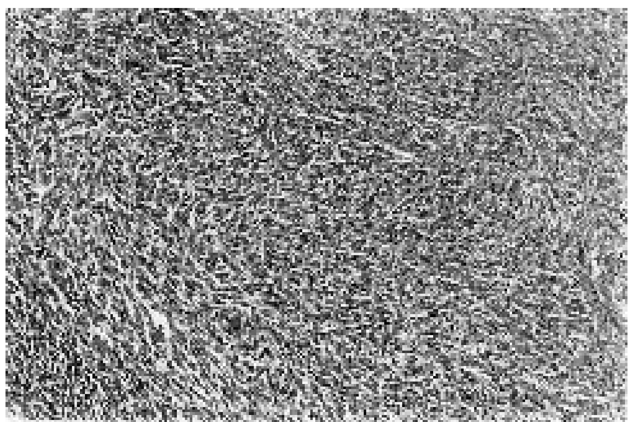


Figure 6 Fascicular spindle cell proliferation in a herringbone pattern is characteristic of but not specific for fibrosarcoma.

considered the therapy of choice. In a series including almost 200 patients, the overall 5- and 10-year relapse rates were 30 and 33%, respectively. Wide local excision with negative surgical margins provided the best results (Ballo *et al*., 1999). However, the need for negative surgical margins to prevent recurrence is not universally accepted. Because desmoid tumours express oestrogen receptors, the use of tamoxifen and anti-inflammatory drugs has been suggested, but the efficacy of these treatments is questionable. Chemotherapy has been attempted in patients with progressive disease.

Fibrosarcoma

Fibrosarcoma was a commonly diagnosed soft tissue sarcoma before the 1980s. However, major advances in immunohistochemistry, electron microscopy, cytogenetics and molecular genetics resulted in the reclassification of most cases as monophasic synovial sarcoma or MPNST. Currently, fibrosarcoma represents less than 5% of soft tissue sarcomas and is considered a diagnosis of exclusion. Fibrosarcoma usually occurs between the fourth and sixth decades of life, and there is a slight male predominance. The lower extremities and the trunk are the most common locations. Histological analysis shows fascicular spindle cell proliferation in a herringbone pattern, which is characteristic but not specific (**Figure 6**). Collagen deposition is usually small, and the cellularity is higher than that observed in deep fibromatoses. However, a low-grade fibrosarcoma can be misdiagnosed as a desmoid tumour. Fibrosarcoma shows immunoreactivity for vimentin only. Cytokeratin, EMA, actins, desmin and S-100 protein must be absent to rule out other differential diagnoses. No specific cytogenetic or molecular findings have been described for the adult type of fibrosarcoma. A study of more than 100 patients from the late 1980s found a 5-year overall survival rate of 40% (Scott *et al*., 1989).

Congenital Fibrosarcoma

Congenital or infantile fibrosarcoma (CF) occurs mainly during the first 2 years of life, has a male predominance and usually arises in the distal extremities. Histologically, it is characterized by a proliferation of oval to round cells with prominent mitotic activity arranged in shorter fascicles than the adult fibrosarcoma. A vascular pattern resembling that of haemangiopericytoma and a rich lymphocytic infiltrate are other frequently observed histological findings. The immunohistochemical profile of CF is similar to that of adult fibrosarcoma. The nonrandom chromosomal translocation t(12;15)(p13;q25) fusing the *ETV6* (*TEL*) gene (*ETS V*ariant *6*) on chromosome 12p13 and the *NTRK3*(*TRKC*) gene (*N*eurotrophin *T*y*R*osine *K*inase type-*3* receptor) on chromosome 15q25 has been found in CF and the cellular variant of congenital mesoblastic nephroma, supporting the view that they have a common pathogenesis. Trisomy 11 is also common in both tumours. Interestingly, the same chimeric gene *ETV6/NTRK3* is found in some cases of acute myelogenous leukaemia. Clinically, CF is characterized by its excellent prognosis (5-year survival rate >80%) and low metastatic potential (10%).

Fibrosarcoma Variants and Related Tumours

Some variants of fibrosarcoma or related lesions with distinct clinicopathologic features have been recognized. They include the low-grade fibromyxoid sarcoma, which should not be confused with myxofibrosarcoma (see Malignant Fibrous Histiocytoma below), the closely related hyalinizing spindle cell tumour with giant rosettes, sclerosing epithelioid fibrosarcoma, acral myxoinflammatory fibroblastic sarcoma and low-grade myofibroblastic sarcoma (myofibrosarcoma).

Inflammatory Myofibroblastic Tumour

Inflammatory myofibroblastic tumour (IMT) is also known as inflammatory pseudotumour, plasma cell granuloma, inflammatory myofibrohistiocytic proliferation, intra-abdominal myxoid hamartoma and inflammatory fibrosarcoma. IMT is a clonal disorder characterized by fascicles of deceptively bland myofibroblastic cells admixed with a prominent inflammatory infiltrate rich in plasma cells, lymphocytes and eosinophils. IMT occurs predominantly during childhood and is commonly associated with systemic symptoms. The abdominal cavity, retroperitoneum and lungs are commonly affected sites. IMT shows immunoreactivity for actin, which supports its myofibroblastic nature. Cytogenetically, recurrent rearrangements of the *ALK* gene (*A*naplastic *L*ymphoma *K*inase) on chromosome 2p23 with the tropomyosin genes *TPM3* and *TPM4* have been found in some cases of IMT (Lawrence *et al*., 2000). Interestingly, the chimaeric *TPM3/ALK* gene also has been described in anaplastic

large cell lymphoma. Expression of Alk protein can also be detected by immunohistochemical analysis. Clinically, IMT is characterized by local recurrence in up to 25% of cases and very low metastatic potential.

Tumours of Fibrohistiocytic Differentiation

The idea of fibrohistiocytic differentiation was introduced in the 1980s and was based on erroneous conclusions drawn from morphological, immunohistochemical and cell culture studies of poorly differentiated sarcomas. Currently, it is largely believed that most of these tumours show some degree of fibroblastic or myofibroblastic differentiation. In this chapter, the term fibrohistiocytic is used in accordance with the current literature and for purposes of classification only.

Atypical Fibroxanthoma

Atypical fibroxanthoma (AF) is a superficial fibrohistiocytic tumour with a bimodal age incidence. In older individuals, AF occurs in areas of solar exposure, such as the head and neck. In young patients, it commonly affects the trunk and extremities. However, it has been suggested that the pleomorphic variant of benign fibrous histiocytoma accounts for most cases of AF in young patients. Clinically, AF presents as small cutaneous nodules that sometimes ulcerate. Histologically, it is characterized by a storiform or fascicular proliferation of spindle and epithelioid cells exhibiting striking pleomorphism and high mitotic activity. Cellular lipidization is common. AF shows immunoreactivity for vimentin and other nonspecific markers. Actin has been detected in some cases, suggesting myofibroblastic differentiation. In a study using digital image analysis, the pleomorphic cells were shown to be aneuploid. In contrast, the spindle cells had a diploid DNA content. The prognosis is excellent, with very low recurrence rates (5%) and an exceedingly small metastatic potential. Local surgical excision using micrographic surgery (Mohs technique) seems to be the treatment of choice.

Dermatofibrosarcoma Protuberans and Giant Cell Fibroblastoma

Dermatofibrosarcoma protuberans (DFSP) and giant cell fibroblastoma (GCF) seem to represent two ends of the same spectrum of tumours. DFSP is a superficial sarcoma that occurs mainly during the third and fourth decades of life. The trunk and upper extremities are the most common locations. Histologically, DFSP is characterized by a storiform proliferation (resembling a rush mat) of spindle cells with minimal mitotic activity and infiltrative margins. The pigmented variant of DFSP, also known as Bednar tumour, is characterized by the presence of abundant melanin-pigmented dendritic spindle cells.

In contrast, GCF frequently occurs in male patients during the first decade of life. The thorax and lower extremities are the most common sites. Histologically, GCF is characterized by pseudovascular (angiectoid) spaces and solid areas where giant cell fibroblasts similar to florete cells are found in variable amounts. Mitotic activity is minimal. Immunohistochemistry shows that DFSP and GCF express vimentin and CD34. Neural, muscular and endothelial markers are usually absent. Ultrastructural analysis shows the presence of modified fibroblasts known as veil cells in both GCF and DFSP. Traditionally, cytogenetic studies have shown ring chromosome 17 in more than 75% of cases of DFSP. It has been further shown that these ring chromosomes harbour a nonrandom translocation t(17;22)(q22;q13) involving the *PDGFβ* gene (*P*latelet-*D*erived *G*rowth *F*actor β-chain) on chromosome 22q13 and the *COL1A1* gene (*COL*lagen type *I* *A*lpha *1*) on chromosome 17q22 in DFSP and GCF (Simon *et al.*, 1997). PDGFβ stimulates cell proliferation and transformation and COL1A1 is an important connective tissue matrix constituent. The chimaeric protein was shown to transform NIH3T3 fibroblast cell lines via an uncontrolled autocrine mechanism, reflecting its oncogenic ability. Clinically, DFSP is characterized by high recurrence rates (up to 75%) and low metastatic potential (1%). The presence of fibrosarcoma or malignant fibrous histiocytoma areas has been associated with higher metastatic rates (15%). This last impression, however, was questioned in one report. GCF also presents a high recurrence rate (up to 50%) but metastases have not been reported so far.

Plexiform Fibrohistiocytic Tumour

Plexiform fibrohistiocytic tumour (PFT) is a rare neoplasm that occurs predominantly in the dermis and subcutaneous tissue of the upper extremities during the first two decades of life. There is a strong female predilection. Histologically, PFT is composed of nodules of histiocyte-like and osteoclast-like giant cells (plexiform areas) separated by fascicles of spindle cells. Haemorrhage and haemosiderin deposition are common in the plexiform areas. Immunohistochemical studies have shown the presence of CD68 in the histiocyte-like and osteoclast-like giant cells and actin in the spindle cells. PFT shows a diploid DNA content. Cytogenetic studies are rare. PFT recurs locally in up to one-third of cases and metastasizes in one-fifth. Because of the tendency of PFT to recur locally, wide local surgical excision seems to be the treatment of choice.

Angiomatoid Fibrous Histiocytoma

Angiomatoid fibrous histiocytoma (AFH), previously known as angiomatoid malignant fibrous histiocytoma, is a rare subcutaneous tumour that arises predominantly in the extremities. AFH commonly occurs in the first three decades of life and does not show sex predilection. AFH may be associated with systemic symptoms, including fever, weight loss, anaemia and paraproteinaemia. These

seem to be caused by cytokines released by the tumour or by the associated inflammatory cells. Histologically, AFH is characterized by a nodular proliferation of oval to spindle cells associated with pseudovascular spaces. A rich lymphoplasmacytic infiltrate often surrounds the lesion. Immunohistochemical studies have shown the presence of vimentin, CD68, actin and desmin, which suggests that AFH has a line of differentiation similar to that of myoid fibroblastic reticulum cells found in the connective tissue of lymph nodes. Ploidy studies have shown a diploid DNA content in most cases. The chromosomal translocation t(12;16)(q13:p11) involving the *ATF-1* gene (*A*ctivating *T*ranscription *F*actor-*1*; see Clear Cell Sarcoma, below) on chromosome 12q13 and the *FUS* gene (*FUS*ion; see Myxoid and Round Cell Liposarcomas, below) on chromosome 16p11 was found in a case of AFH. Despite the need for confirmatory studies, this finding opened new avenues for the understanding of the pathogenesis of this tumour. AFH is characterized by low rates of local recurrence (10–20%) and limited metastatic potential (<2%). Surgical excision is considered the treatment of choice.

Malignant Fibrous Hystiocytoma

Malignant fibrous histiocytoma (MFH) comprises a heterogeneous group of undifferentiated neoplasms characterized by high pleomorphism and cellular atypia. In the past, a fibrohistiocytic origin was attributed to these neoplasms on the basis of morphological and functional similarities to fibroblasts and histiocytes found in studies of cell cultures. This misinterpretation was supported by the detection of nonspecific histiocytic markers on MFH. With advances in the immunohistochemical techniques and careful tissue sampling, many cases diagnosed as MFH have been reclassified as other high-grade tumours, such as pleomorphic liposarcoma, dedifferentiated liposarcoma, pleomorphic rhabdomyosarcoma, pleomorphic leiomyosarcoma, undifferentiated carcinoma and anaplastic large cell lymphoma. However, there is still a group of tumours that lack specific features and thus remain unclassifiable. Therefore, MFH is currently regarded as a group of tumours with fibroblastic characteristics but undetermined or aberrant lines of differentiation as compared with normal cell types. MFH is considered a diagnosis of exclusion.

MFH was once the most common soft tissue sarcoma. MFH most commonly occurs between the fifth and eighth decades of life and usually arises in the lower extremities (particularly thigh), upper extremities and retroperitoneum (Weiss and Enzinger, 1978). Males are affected slightly more frequently than are females. Histologically, MFH is divided into four variants: storiform/pleomorphic, myxoid, giant cell and inflammatory. Storiform/pleomorphic MFH is characterized by a proliferation (in a pattern resembling a rush mat) of spindle to oval cells and bizarre multinucleated cells showing prominent mitotic activity (**Figure 7**) (Weiss and Enzinger, 1978). Myxoid MFH, also known as myxofibrosarcoma, has a multinodular and mucinous appearance and tends to be more superficial than the storiform/pleomorphic variant. Myxoid MFH is unusual in the retroperitoneum. Extensive myxoid areas characterize this histological variant; these areas are rich in hyaluronic acid, have high vascularization and contain lipoblast-like cells (pseudolipoblasts). It is currently believed that the giant cell variant of MFH probably encompasses a group of distinct neoplasms rich in osteoclast-like giant cells; this variant includes giant cell-rich extraskeletal osteosarcoma, giant cell-rich pleomorphic leiomyosarcoma, giant cell-rich undifferentiated tumours and the soft tissue counterpart of the giant cell tumour of bone. Inflammatory MFH, the least common variant (<5%), most commonly is found in the retroperitoneum and often is associated with peripheral leucocytosis and systemic symptoms. Histologically, it is characterized by the presence of an inflammatory infiltrate rich in neutrophils, xanthomatous histiocytes, lymphocytes and eosinophils among bizarre Reed–Sternberg-like cells. MFH has no specific immunohistochemical profile; cytogenetic studies have shown very complex karyotypes.

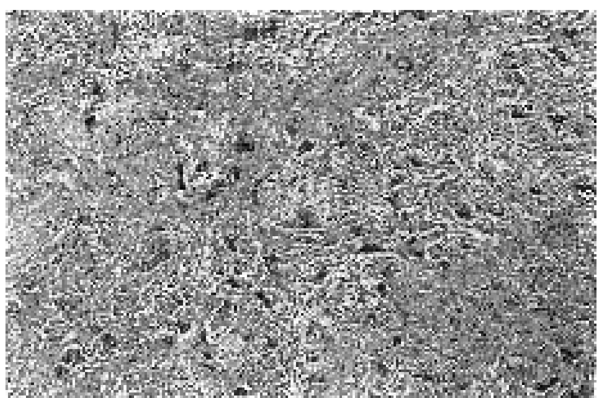


Figure 7 Malignant fibrous histiocytoma showing bizarre giant cells in a vague storiform pattern.

The prognosis of MFH is dictated primarily by its grade and stage. However, other clinicopathological features have been correlated with tumour aggressiveness. A retroperitoneal location and the inflammatory variant have been associated with a poor prognosis. In contrast, myxoid MFH tends to run a less aggressive clinical course. In a series of 216 patients with localized MFH, 5-year disease specific-free, metastasis-free and local recurrence-free survival rates were 70, 63 and 63%, respectively. The presence of high-grade histological features was considered the most important prognostic factor for metastasis-free and disease specific-free survival (Le Doussal *et al.*, 1996).

Sarcomas of Adipose Tissue Differentiation

Sarcomas of adipose tissue differentiation, or liposarcomas, are currently considered the most common soft

tissue sarcoma. This group comprises four major variants with distinct clinicopathological and genetic features: well-differentiated, myxoid/round cell, pleomorphic and dedifferentiated.

Well-differentiated Liposarcoma

Well-differentiated liposarcoma arises most commonly in the extremities and retroperitoneum. Histologically, well-differentiated liposarcoma is divided into three major subvariants: lipoma-like (adipocytic), sclerosing and inflammatory. A spindle cell variant also has been recognized. The lipoma-like variant is the most common and is predominantly composed of uniloculated adipocytes and scattered atypical multivacuolated lipoblasts. The sclerosing variant, which occurs most frequently in the retroperitoneum, has extensive fibrillary fibrotic areas where scattered multivacuolated lipoblasts are found. The inflammatory (or lymphocyte-rich) variant is characterized by the presence of lymphoplasmacytic nodular aggregates and may be misdiagnosed as inflammatory myofibroblastic tumour or Castleman disease. This variant often arises in the retroperitoneum. Well-differentiated liposarcomas arising in superficial locations also are known as atypical lipomas because they have an excellent prognosis and no metastatic potential. However, controversies regarding this nomenclature remain. Cytogenetic studies frequently show rings or giant rods derived from chromosome 12 in all subtypes of well-differentiated liposarcomas. These structures contain amplified 12q13-15 sequences, which include the *MDM2*, *SAS*, *HMGI-C* and *CDK4* genes. The clinical course of well-differentiated liposarcomas is usually excellent, with overall 5-year survival rates >90%. However, a worse prognosis has been found in retroperitoneal and dedifferentiated tumours. Local recurrence seems to be prevented by wide local excision and is more common in the retroperitoneum.

Myxoid and Round Cell Liposarcomas

Myxoid liposarcoma is the most common subtype of liposarcoma, representing 35–45% of all liposarcomas. It occurs mainly during the fifth decade of life, has a male predilection, and most commonly arises in the lower extremities, particularly the thigh (Kilpatrick *et al*., 1996). Histologically, myxoid liposarcoma is composed of spindle and stellated cells immersed in a myxoid matrix rich in hyaluronic acid. A chicken-wire pattern of vascularization is characteristic (**Figure 8**). Lipoblasts occasionally are found, and cystic degeneration of the matrix (resembling lymphangioma) is common.

Round cell liposarcoma is the cellular and poorly differentiated form of myxoid liposarcoma. Histologically, it is composed of round and larger cells with high mitotic and apoptotic activity. Cytogenetic and molecular studies have shown that myxoid and round cell liposarcomas share the same chromosomal translocations. The translocation

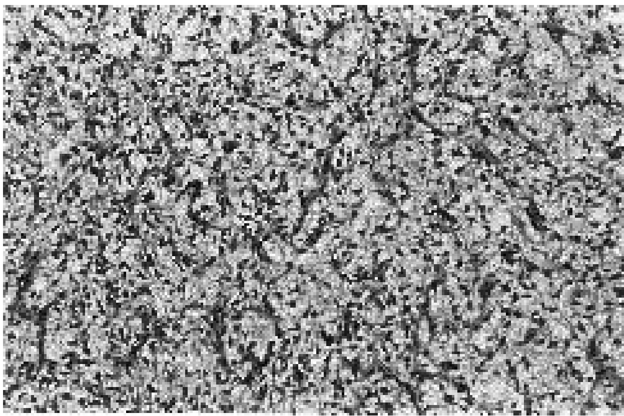


Figure 8 A chicken-wire pattern of vascularization is characteristic and almost diagnostic of myxoid liposarcoma.

t(12;16)(q13;p11) occurs in 95% of cases and involves the *TLS* gene (*T*ranslocated in *L*iposarcoma) or *FUS* gene (*FUS*ion) on chromosome 16p11 and the *CHOP* gene (*C*CAAT/enhancer binding protein *HO*mologous *P*rotein) or *DDIT3* (*D*NA-*D*amage-*I*nducible *T*ranscript *3*) on chromosome 12q13 (**Figure 9**) (Crozat *et al*., 1993; Rabbitts *et al*., 1993). *TLS* encodes for RNA-binding protein similar to the EWS protein (**Figure 2**). *CHOP* encodes for DNA-damage-inducible negative transcription regulator, which also is involved in adipocyte differentiation. In the second chromosomal translocation t(12;22)(q13;q12), which occurs in <5% of cases, the *EWS* gene on chromosome 22q12 fuses with *CHOP* (**Figure 9**). Both chimaeric proteins TLS/CHOP and EWS/CHOP can transform NIH3T3 fibroblasts, reflecting their oncogenic properties.

Myxoid liposarcoma has been associated with a relatively good prognosis, with a 5-year overall survival rate of 75%. Round cell differentiation has been considered the single most important adverse prognostic factor in myxoid liposarcoma (Kilpatrick *et al*., 1996). More recently, expression of the cell cycle inhibitor p27^{kip1} was associated with a prolonged metastasis-free and overall survival in myxoid/round cell liposarcoma (Oliveira *et al*., 2000).

Dedifferentiated Liposarcoma

Dedifferentiated liposarcoma is traditionally defined as a well-differentiated liposarcoma showing an abrupt transition to areas of high-grade pleomorphic histological features, similar to storiform/pleomorphic MFH or fibrosarcoma. However, low-grade dedifferentiated areas also have been recognized. Dedifferentiated liposarcoma is predominantly found in the retroperitoneum and appears to occur as a *de novo* phenomenon in >90% of cases. Divergent differentiation to osteosarcoma, chondrosarcoma, rhabdomyosarcoma, leiomyosarcoma and neural-like areas have been described.

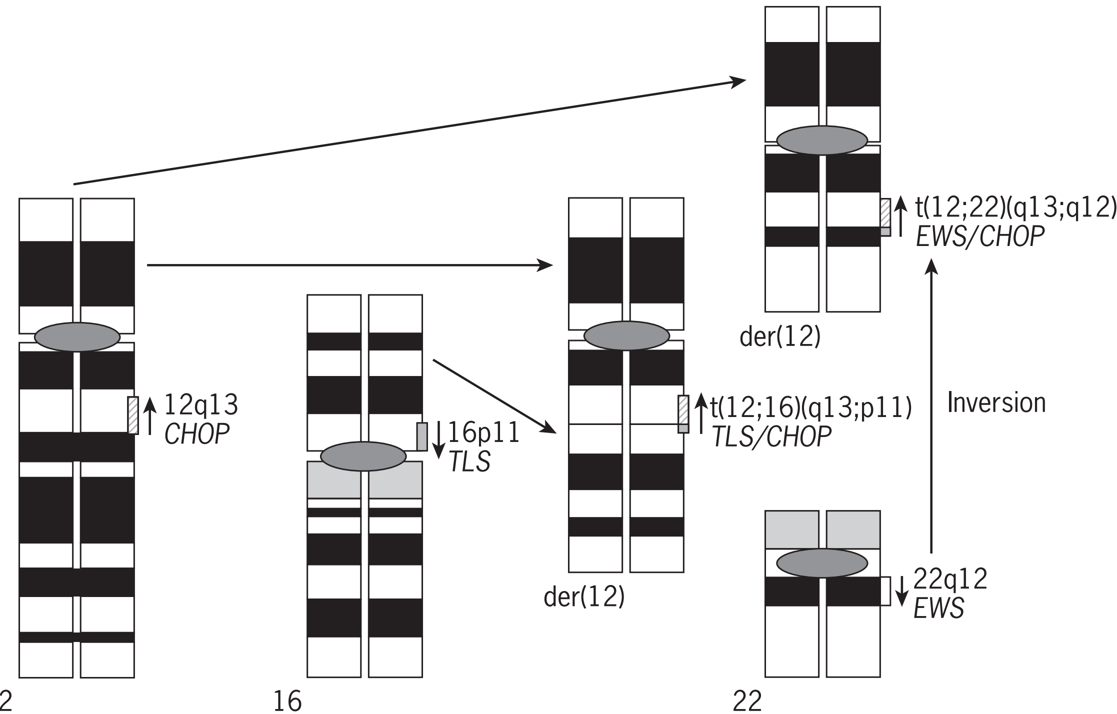


Figure 9 Schematic diagram showing two chromosomal translocations in myxoid liposarcoma: t(12;16)(q13;p11) and t(12;22)(q13;q12). The most common translocation (95% of cases) involves the fusion of the *CHOP* gene on chromosome 12q13 with the *TLS* gene on chromosome 16p11. Observe that the 5′ end of *TLS* gene on chromosome 16 fuses with the downstream sequences of the *CHOP* gene on the derivative chromosome 12. In the least common translocation (5% of cases), the 5′ end of *EWS* gene fuses with downstream sequences of *CHOP*. Observe that the *EWS* gene is read from centromere to telomere (c→t) on chromosome 22 and that the *CHOP* gene is read from telomere to centromere (t→c) on chromosome 16. This suggests that an inversion involving the *EWS* gene must occur before its fusion with the *CHOP* gene, which may explain the lower frequency of this second chromosomal translocation in myxoid liposarcoma. The derivative chromosomes 16 and 22 are not shown (arrowhead, 3′ end; arrow tail, 5′ end).

Immunoreactivity for S-100 protein, frequently found in liposarcomas, is not observed in the dedifferentiated areas. Mdm2 protein overexpression is common, and *TP53* gene mutations are rare. The amount and histological grade of the dedifferentiated component have not been correlated with outcome. In a large series of patients with dedifferentiated liposarcomas, the local recurrence, metastatic and disease-related mortality rates were 41, 17 and 28%, respectively. Overall 5-year survival rates of 50–70% have been estimated.

Pleomorphic Liposarcoma

Pleomorphic liposarcoma comprises <5% of the cases of liposarcoma and most frequently occurs in the extremities. Pleomorphic liposarcoma is a high-grade sarcoma showing MFH-like histological features, including the finding of scattered multiloculated lipoblasts. The incidence of pleomorphic liposarcoma is expected to increase because of ongoing reclassification of cases of MFH. The prognosis of pleomorphic liposarcoma parallels that of other high-grade sarcomas.

Sarcomas of Smooth Muscle Differentiation

Leiomyosarcoma is the single example of this group and probably represents 3–5% of soft tissue sarcomas. Leiomyosarcomas arising in soft tissues are clinically divided into four main types: cutaneous, subcutaneous and intramuscular, vascular and intra-abdominal. All types are morphologically and immunohistochemically indistinguishable. They are characterized by a fascicular proliferation of eosinophilic spindle cells with cigar-shaped nuclei and perinuclear vacuolization. The tumour fascicles tend to cross each other obliquely, a feature not readily appreciated in pleomorphic variants. Immunohistochemical studies show smooth muscle actin (SMA) in >90% of cases and desmin in 50–70% of cases. Approximately 40% of leiomyosarcomas express cytokeratin and epithelial membrane antigen. Cytogenetic studies frequently show complex karyotypes, and no consistent recurrent chromosomal abnormality has been identified. It has been suggested that 1p36 and 8p21p-ter chromosomal losses

may be associated with a more aggressive clinical course, but studies are needed to confirm these initial impressions.

Cutaneous Leiomyosarcoma

Cutaneous leiomyosarcoma occurs most frequently in the lower extremities of young adults. Clinically, they are painful tumours with almost no metastatic potential. Local recurrences occur in up to one-third of cases and seem strongly related to the adequacy of the surgical excision.

Subcutaneous and Intramuscular Leiomyosarcomas

Subcutaneous and intramuscular leiomyosarcomas occur more frequently in the thigh but the retroperitoneum also is a common site. In contrast to the cutaneous form, subcutaneous and soft tissue leiomyosarcomas frequently metastasize to the lungs. A 5-year overall survival rate of 64% was found in a large series of cases (Gustafson *et al*., 1992).

Vascular Leiomyosarcoma

Vascular leiomyosarcoma represents <5% of all cases of leiomyosarcoma. It arises most commonly in the inferior vena cava and in veins of the lower extremities (Enzinger and Weiss, 1995). There is no sex predilection except for those arising in the inferior vena cava, where a striking female predilection is observed. The pulmonary artery is the single most commonly affected arterial site. In a review of 144 cases of inferior vena cava leiomyosarcoma, radical tumour excision was associated with a better overall survival, despite high recurrence rates (50%) (Mingoli *et al*., 1991).

Intra-abdominal Leiomyosarcoma

Intra-abdominal leiomyosarcoma most commonly occurs in the retroperitoneum, omentum and mesentery and seems more common in females. These tumours can reach large sizes and often follow an aggressive clinical course. In a series of 44 cases, the overall 5-year survival rate was 21% (Hashimoto *et al*., 1985). However, with the advent of immunohistochemical analysis, some intra-abdominal leiomyosarcomas have been reclassified as extragastrointestinal stromal tumours (see the chapters *Upper Gastrointestinal Tract (Oesophagus, Stomach)* and *Lower Gastrointestinal Tract (Small Intestine, Appendix, Colon, Rectum, Anus*).

Sarcomas of Skeletal Muscle Differentiation

Rhabdomyosarcoma and its subtypes comprise this group of sarcomas. However, rhabdomyoblastic differentiation may be observed in other soft tissue sarcomas, such as malignant peripheral nerve sheath tumour with rhabdomyoblastic differentiation (Triton tumour), dedifferentiated liposarcoma and others.

Rhabdomyosarcoma

Rhabdomyosarcoma is the most common peadiatric soft tissue sarcoma (Qualman *et al*., 1998). Classically, it is divided into the embryonal, botryoid, alveolar and pleomorphic types. However, modifications of this traditional classification are taking place and no schema is universally accepted. Currently, three histological variants are recognized: embryonal (50–60), alveolar (25–30) and pleomorphic (5%).

Embryonal rhabdomyosarcoma occurs most commonly in the first decade of life and is most commonly found in the head and neck area and in the genitourinary tract. It is characterized by a proliferation of small basophilic spindle cells associated with a variable number of rhabdomyoblasts (strap or tadpole cells) in a myxoid stroma. Variants of embryonal rhabdomyosarcoma include the spindle cell and botryoid. Spindle cell rhabdomyosarcoma occurs primarily in the paratesticular area and is associated with an excellent prognosis (5-year overall survival >85%). Botryoid rhabdomyosarcoma has a grape-like macroscopic appearance and most commonly occurs in the genitourinary tract. The characteristic histological finding is the condensation of the rhabdomyoblasts underneath the epithelial lining (cambium layer).

Alveolar rhabdomyosarcoma occurs more commonly in adolescents and frequently arises in the extremities, particularly in the distal segments of the upper extremities, and the trunk. It is characterized by nests of round basophilic cells separated by fibrous septa, simulating the pulmonary parenchyma. Osteoclast-like giant cells frequently are seen in the fibrous septa, and rhabdomyoblasts are scarce.

Pleomorphic rhabdomyosarcoma arises more commonly in the extremities of adults after the fourth decade of life and is associated with an aggressive clinical course. Histologically, it is similar to storiform/pleomorphic MFH and occasionally has large eosinophilic rhabdomyoblasts.

Two other rhabdomyosarcoma variants have been recognized: anaplastic and undifferentiated sarcoma. These terms are in the International Classification of Rhabdomyosarcoma (ICR). Anaplastic rhabdomyosarcoma is considered the pleomorphic variant of rhabdomyosarcoma that occurs in children. Undifferentiated sarcoma, which does not show any rhabdomyoblastic differentiation, is included in the ICR because it has a similar response to the chemotherapy used for other rhabdomyosarcomas (Qualman *et al*., 1998).

Immunohistochemical studies are important in the differential diagnosis of other tumours, such as Ewing sarcoma/PNET and desmoplastic small round cell tumour. Rhabdomyosarcoma cells express MSA (94%), desmin (80–95% of tumours), MyoD1 (90–95%) and myogenin (75–99%). Myoglobin is expressed in more differentiated rhabdomyoblasts and consequently is less sensitive (60–78%). SMA is present in a minority of cases (<5%).

Ultrastructural studies show the presence of thick myosin filaments associated with ribosomes and Z bands; this finding indicates muscle differentiation. Ultrastructural studies are particularly useful for the diagnosis of poorly differentiated rhabdomyosarcomas, such as pleomorphic and anaplastic rhabdomyosarcoma.

Cytogenetic and molecular studies have shown that alveolar rhabdomyosarcoma is associated with two chromosomal translocations. The translocation t(2;13)(q35;q14) occurs in 75% of cases and involves the fusion of the transcription factors *PAX3* gene (*PA*ired bo*X 3*) on chromosome 2q35 and *FKHR* gene (*F*or*KH*ead *R*elated) on chromosome 13q14 (Galili *et al*., 1993). In the translocation t(1;13)(p36;q14), which occurs in approximately 10% of cases, *FKHR* is fused with another partner, *PAX7*. *PAX3* and *PAX7* are specifically expressed during the development of the dorsal neural tube and somites. *PAX3* is also important for the migration of myoblasts to the limbs. Gene expression analyses using cDNA microarrays in alveolar rhabdomyosarcoma have shown that the *PAX3/FKHR* fusion transcript induces the expression of several genes involved in myogenic differentiation, including *MyoD1* and myogenin. In addition, PAX3 protein and the chimaeric protein FKHR/PAX3 stimulate the expression of the antiapoptotic protein BCL-X_L and the *MET* proto-oncogene. PAX7/FKHR fusion transcript in alveolar rhabdomyosarcoma was associated with a younger age at presentation, extremity location, prolonged event-free survival and a trend for a prolonged overall survival in a small series of cases (Kelly *et al*., 1997).

Embryonal rhabdomyosarcoma frequently shows loss of heterozygosity on 11p15, and this may contribute to the commonly observed overexpression of insulin-like growth factor-II (IGFII). The mechanism underlying this effect seems to result from loss of imprinting or loss of heterozygosity of the normally silent *IGFII* maternal-derived gene allele (see Epidemiology, Aetiology and Pathogenesis, above). Furthermore, loss of heterozygosity of the 11p15 locus can lead to loss of the tumour-suppressor gene *GOK*.

Three staging systems for rhabdomyosarcoma have been proposed: the Clinical Group System (CGS), the TNM system from the Union Internationale Contre le Cancer (UICC) and the TNM system from the American Joint Committee on Cancer (AJCC). The CGS is a surgico-pathological system based on the surgical resectability of the tumour. The TNM system proposed by the UICC is based on the invasiveness of the tumour according to the presence or absence of tumour confinement to a specific organ or muscular group. The TNM system proposed by the AJCC relies primarily on tumour size (Fleming *et al*., 1997). In the Intergroup Rhabdomyosarcoma Study-V, the information provided by these three systems has been combined to establish prognostic groups (**Table 5**).

Table 5 Prognostic groups for rhabdomyosarcoma

Stage	Group[a]	Site[b]	Size[c]	Age (years)	Histology[d]	Lymph nodes[e]	Distant metastasis[e]	Disease-free survival
1	I	Favourable	a or b	<21	Embryonal	N0	M0	
1	II	Favourable	a or b	<21	Embryonal	N0	M0	
1	III	Orbit only	a or b	<21	Embryonal	N0	M0	Excellent (>85%)
2	I	Unfavourable	a	<21	Embryonal	N0 or NX	M0	
1	II	Favourable	a or b	<21	Embryonal	N1	M0	
1	III	Orbit only	a or b	<21	Embryonal	N1	M0	
1	III	Favourable (no orbit)	a or b	<21	Embryonal	Any	M0	
2	II	Unfavourable	a	<21	Embryonal	N0 or NX	M0	Very good (70–85%)
3	I or II	Unfavourable	a	<21	Embryonal	N1	M0	
3	I or II	Unfavourable	b	<21	Embryonal	Any	M0	
2	III	Unfavourable	a	<21	Embryonal	N0 or NX	M0	
3	III	Unfavourable	a	<21	Embryonal	N1	M0	
3	III	Unfavourable	a	<21	Embryonal	Any	M0	Good (50–70%)
1 or 2 or 3	I or II or III	Any	a or b	<21	Alveolar	Any	M0	
4	Any	Any	a or b	<10	Embryonal	N0 or N1	M1	
4	IV	Any	a or b	>10	Embryonal	N0 or N1	M1	Poor (<30%)
4	IV	Any	a or b	<21	Alveolar	N0 or N1	M1	

[a]Group I, localized tumour, completely excised; group II, total gross tumour excision with evidence of regional spread; group III, incomplete excision with gross residual tumour; group IV, distant metastasis present at onset.
[b]Favourable sites are orbit and eyelid, nonparameningeal head and neck and nonbladder and nonprostate genitourinary tract; unfavourable sites are bladder, prostate, extremity, parameningeal and others.
[c]a, tumour size $\leq$5 cm; b, tumour size >5 cm.
[d]Embryonal, including also botryoid and spindle cell variants; alveolar, including also solid alveolar, anaplastic, and undifferentiated variants.
[e]N0, regional lymph nodes not clinically involved; N1, regional lymph nodes clinically involved; NX, lymph node status unknown; M0, no distant metastases; M1, distant metastases present.
Adapted from Wexler *et al*., 2000, Soft tissue sarcomas of childhood. In: Bust, R. C., Jr, *et al*. (eds), *Cancer Medicine*, 5th edn. 2198–2203. (B. C. Decker, Hamilton, ON), by permission of the publisher.

Treatment of rhabdomyosarcoma commonly entails chemotherapy, radiotherapy and surgical excision. Chemotherapy is indicated for unresectable tumours, to allow subsequent surgical excision, for elimination of microscopic disease after primary surgical excision and for metastatic disease. Radiation therapy is used to provide local tumour control after surgical excision or for metastatic disease. Vincristine, D-actinomycin, and cyclophosphamide are commonly used drugs (Arndt and Crist, 1999).

Clinical features and the histological subtypes of rhabdomyosarcoma seem the most important factors in predicting the overall prognosis (**Table 5**). The International Classification of Rhabdomyosarcoma divides rhabdomyosarcomas into those associated with a superior prognosis (botryoid and spindle cell variants of embryonal rhabdomyosarcoma), intermediate prognosis (embryonal rhabdomyosarcoma) and poor prognosis (alveolar and anaplastic or undifferentiated) (Newton *et al*., 1995; Qualman *et al*., 1998). In a large series of cases from the Intergroup Rhabdomyosarcoma Study, 5-year overall survival rates of 95, 88, 64 and 53% for botryoid, spindle cell, embryonal and alveolar subtypes, respectively, were found. By multivariate analysis, the histological categories retained their prognostic significance after adjustment for other prognostic factors in rhabdomyosarcoma, such as tumour sizes primary site, and clinical group (Newton *et al*., 1995).

Sarcomas of Endothelial and Pericytic Differentiation

The three most important sarcomas showing endothelial differentiation are epithelioid haemangioendothelioma, Kaposi sarcoma and angiosarcoma. These sarcomas and haemangiopericytoma are discussed below.

Epithelioid Haemangioendothelioma

Epithelioid haemangioendothelioma is a rare vascular tumour that occurs most commonly in the superficial and deep soft tissues, liver and bone. The tumour affects mainly adults in the fourth to sixth decades of life. Approximately 60% of cases arise in a vessel, usually a vein of medium size or larger. Histologically, it is characterized by cords and nests of round to spindle eosinophilic cells arranged in a myxohyaline stroma. The cells commonly exhibit an intracytoplasmic vacuolization (signet ring-like cells) in which erythrocytes occasionally are found. Mitotic activity is usually very low. Immunohistochemical studies show the expression of vascular markers, such as factor VIIIRa, CD31 and CD34. Expression of cytokeratin and actin is frequently found. Ultrastructural studies support its endothelial differentiation. Cytogenetic and molecular studies are almost nonexistent, but the balanced chromosomal translocation t(1;3)(p36.3;q25) was recently found in two cases. Epithelioid haemangioendothelioma is less aggressive than angiosarcoma, but it metastasizes in 20% of cases and has a mortality rate of 17% when it occurs in soft tissues. Higher mortality rates are seen in hepatic and pulmonary tumours.

Kaposi Sarcoma

Kaposi sarcoma is a unique tumour showing features of lymphatic endothelial differentiation. Whether Kaposi sarcoma is a reactive process or a true neoplasm remains to be settled because monoclonality has been shown by some but not all investigators (Gill *et al*., 1998). A multicentric clonal evolution has been proposed to explain the pathogenesis of Kaposi sarcoma. In this elegant model, Kaposi sarcoma arises from independent cells and acquires clonal characteristics during tumour progession (Gill *et al*., 1998).

Kaposi sarcoma traditionally has been divided into four clinical variants: classical, endemic, post-transplant-associated or immunosuppression-associated and epidemic or acquired immunodeficiency syndrome (AIDS)-associated (Antman and Chang, 2000). The classical form is prevalent in Eastern Europe and the Mediterranean and typically presents as indolent purple–blue cutaneous nodules on the lower extremities of elderly men. An increased risk of haematological disorders, particularly Hodgkin lymphoma, has been attributed to this variant. The endemic form occurs in human immunodeficiency virus (HIV)-negative individuals and commonly manifests as indolent nodular lesions associated with localized or generalized lymphadenopathy. A more aggressive form is commonly found in children (Antman and Chang, 2000). The transplant-associated or immunosuppression-associated Kaposi sarcoma is an aggressive form that occurs months or years after transplantation or the beginning of immunosuppression. It often is characterized by lymphatic or visceral manifestations (Antman and Chang, 2000). The epidemic or AIDS-associated Kaposi sarcoma is a very aggressive form that involves the skin, mucosas, viscera and lymph nodes. It is more commonly seen in homosexual AIDS patients.

Independent of clinical presentation, all Kaposi sarcoma are histologically similar. The relationship between Kaposi sarcoma and Kaposi sarcoma-associated herpes virus is discussed above (see Epidemiology, Aetiology and Pathogenesis, above). Histologically, Kaposi sarcoma is characterized by a spindle cell nodular proliferation separated by slit-like thin-walled vascular channels. Extravasated erythrocytes, haemosiderin deposition and hyaline intracytoplasmic and extracytoplasmic PAS-positive globules are commonly seen. Immunohistochemical analysis shows CD31 and CD34 in most cases and variable expression for factor VIIIRa, consistent with Kaposi sarcoma endothelial differentiation. Treatment includes surgical excision (simple excisions), radiation therapy and single or multiple agent chemotherapy, including interferon-α (Antman and Chang, 2000).

Angiosarcoma

Known in the past as malignant haemangioendothelioma or lymphangiosarcoma, angiosarcoma is a highly malignant neoplasm of endothelial differentiation. The term angiosarcoma encompasses blood and lymphatic vascular tumours because no current technique reliably differentiates the two. However, new markers specific for the lymphatic endothelium, such as M2A oncofoetal antigen may shed some light on this issue. Aetiological factors associated with the development of angiosarcoma are discussed above (see Epidemiology, Aetiology and Pathogenesis).

Clinically, angiosarcomas are divided into four groups: cutaneous, soft tissue, postirradiation and visceral. Cutaneous angiosarcoma not associated with lymphoedema shows a predilection for elderly men and most frequently occurs in the head and neck region, particularly in the scalp. It is characterized by a haemorrhagic appearance and multinodularity. The prognosis is reserved, with 5-year overall survival rates <35%. Cutaneous angiosarcoma associated with chronic lymphoedema is discussed above (see Epidemiology, Aetiology and Pathogenesis).

Soft tissue angiosarcoma occurs predominantly in the abdominal cavity, including the retroperitoneum, and lower extremities of older men. As in the cutaneous form, the prognosis is very poor, with more than 50% of patients dying before 1 year after the diagnosis. In a large series of cases, older age, the presence of retroperitoneal tumours and high expression of the proliferative marker Ki-67 were associated with a more aggressive clinical course (Meis-Kindblom and Kindblom, 1998).

Postirradiation angiosarcoma is a highly aggressive tumour that tends to have a shorter latent period than other irradiation-induced sarcomas. Postirradiation angiosarcoma should not be confused with the cutaneous angiosarcoma associated with chronic lymphoedema. Visceral angiosarcoma often occurs in the liver, breast and spleen and is discussed in the respective chapters.

Histologically, all clinical types of angiosarcoma are characterized by an infiltrative proliferation of anastomosing irregular vascular spaces lined by atypical and pleomorphic endothelial cells. Immunohistochemical studies show the expression of Factor VIIIRa, CD31 and CD34 in most cases and the expression of cytokeratin in 35%. However, immunoreactivity for some endothelial markers tends to be lost in more undifferentiated areas. Cytogenetic and molecular studies are limited, and no specific abnormality has been found. *TP53* and K-*RAS2* gene mutations are common. A proliferative autocrine loop mediated by vascular endothelial growth factor (VEGF) and its receptor FLT has been proposed for the pathogenesis of angiosarcoma.

Haemangiopericytoma

Haemangiopericytoma has been considered a controversial entity since it was first described more than 50 years ago. Its putative pericytic nature was suggested by its morphological similarities to normal pericytes. However, this analogy has not been confirmed by immunohistochemical analysis since most haemangiopericytomas do not express actin, which is normally expressed by pericytes. Most cases diagnosed as haemangiopericytoma have been reclassified as entities such as monophasic synovial sarcoma, mesenchymal chondrosarcoma and solitary fibrous tumour. Currently, the most prevalent opinion is that haemangiopericytoma is a morphological pattern shared by distinct entities, and only a few unclassifiable cases should be termed haemangiopericytoma. Moreover, the old subdivision of haemangiopericytoma into adult and infantile forms no longer holds because the latter has been reclassified as myofibroma. Cases that still deserve the label haemangiopericytoma are some soft tissue tumours that arise in the retroperitoneum and pelvis, meningeal haemangiopericytoma (previously known as angioblastic meningioma), and sinonasal haemangiopericytoma.

Histologically, haemangiopericytomas are composed of a vague nodular proliferation of basophilic oval cells arranged around dilated thin-walled vessels with a characteristic staghorn appearance. Nuclear atypia, high mitotic activity, high cellularity and necrosis and haemorrhage tend to be associated with a more aggressive clinical course. Immunohistochemical analyses show the expression of vimentin and CD34. Actin is expressed more consistently by sinonasal haemangiopericytomas. Clinically, some haemangiopericytomas are associated with hypoglycaemia, which seems to be caused by IGF and IGF receptor expression by the tumour. Meningeal haemangiopericytomas often are associated with an aggressive clinical course characterized by local recurrences and metastases. Sinonasal haemangiopericytomas tend to recur locally but do not mestastasize.

Sarcomas of Peripheral Neuroectodermic Differentiation

Extraskeletal Ewing Sarcoma and Peripheral Primitive Neuroectodermal Tumour (PNET)

Extraskeletal and skeletal Ewing sarcoma and peripheral primitive neuroectodermal tumour (PNET) (also known as neuroepithelioma) represent opposite ends of a continuum of primitive tumours showing various degrees of neuroectodermic differentiation. Ewing sarcoma is the more undifferentiated tumour and PNET the more differentiated tumour. In extraskeletal sites, they tend to occur in the paravertebral areas, thoracopulmonary region (Askin tumour) and lower extremities of adolescents and young adults. There is no sex predilection but African Americans are rarely affected.

Histologically, sheets or lobules of small round cells separated by strands of fibrous tissue are characteristic

but not specific. Homer Wright rosettes frequently are found in PNET. Brisk mitotic activity and areas of confluent necrosis (filigree pattern) are found in both Ewing sarcoma and PNET. Immunohistochemical studies show the expression of the MIC2 protein in both tumours. Expression of neuron-specific enolase, synaptophysin, Leu-7 (CD57), neurofilament, S-100 protein and the protein gene product 9.5 (PGP 9.5) as well as the presence of Homer Wright rosettes have been suggested as diagnostic criteria for PNET because they indicate neuroectodermal differentiation. (See also the chapter *Bones*.)

Cytogenetic studies done mostly in skeletal tumours have shown the balanced translocations t(11;22)(q24;q12) and t(21;22)(q22;q12) in 95% and 5% of cases, respectively (Sandberg and Bridge, 2000). In these translocations, the *EWS* gene on chromosome 22q12 fuses with *FLI1* gene on chromosome 11q24 or *ERG* gene on chromosome 21q22, respectively. Other translocations have been identified involving the *EWS* gene in a minority of cases (**Table 2**). *EWS/FLI1* has been shown to transform mouse NIH3T3 fibroblasts in transfection experiments, which reflects its oncogenic properties. Among several types of *EWS/FLI1* transcripts, the fusion of the *EWS* exon 7 with the *FLI1* exon 6 is the most common (type I) (**Figure 1**) and has been associated with a more aggressive disease independent of stage, tumour location and age (de Alava *et al*., 1998). However, the clinical significance of individual transcripts other than those of type I remains unknown. Other common cytogenetic abnormalities are trisomies 8 and 12, der(16)t(1;16)(q12;q11.2) and deletion 1p36. The first two appear to represent independent events during tumour evolution. The inactivation by distinct mechanisms of the tumour suppressor gene $p16^{INK4}$ has been correlated with a worse prognosis in Ewing sarcoma.

Many other prognostic factors have been suggested for Ewing sarcoma/PNET but detection of distant metastases at presentation, which occurs in approximately 25% of patients, is the most adverse (Terrier *et al*., 1996). The value of neuroectodermal differentiation as a prognostic factor in Ewing sarcoma/PNET remains controversial. Treatment of localized skeletal Ewing sarcoma/PNET consists of a combination of surgical excision, radiotherapy and chemotherapy. The most commonly used agents are vincristine, doxorubicin, cyclophosphamide, D-actinomycin, etoposide and ifosfamide. The 5-year overall survival rate can reach 70% for patients with localized disease and 20–30% for those who presented with distant metastasis (Arndt and Crist, 1999). Treatment and survival data on extraskeletal Ewing sarcoma/PNET are limited, and the clinical course seems more aggressive. The overall 5-year survival rate appears to be <40%. Potential new treatments for Ewing sarcoma/PNET have been reported. The use of monoclonal anti-CD99 and nanocapsules with antisense *EWS/FLI1* were shown to inhibit Ewing sarcoma cells in both *in vitro* and *in vivo* experiments.

Malignant Peripheral Nerve Sheath Tumour

Malignant peripheral nerve sheath tumour (MPNST), formerly known as malignant schwannoma and neurofibrosarcoma, is divided into three clinical groups: neurofibromatosis type I-associated (40–50%), sporadic (40–50%) and postirradiation (5–10%). The lifetime risk for the development of MPNST in patients with neurofibromatosis type I (NF1) is estimated to be 2%. The anatomical distribution is wide, and there is a male predominance. The incidence of NF1-associated MPNST peaks during the fourth decade of life. The sporadic form occurs more often in the lower extremities, shows no sex predilection and peaks during the fifth decade of life. The postirradiation form is the least common and occurs up to 30 decades after the radiation therapy.

Histologically, MPNST is characterized by a fascicular spindle cell proliferation in which there is an abrupt transition from more cellular areas to more myxoid areas. Whorling of tumour cells around vascular spaces is another characteristic feature. In 10% of cases, rhabdomyoblastic differentiation is observed (Triton tumour), which portends a more aggressive clinical course. MPNST shows epithelioid phenotype in 5% of cases. This variant tends to be superficially located and is associated with a more favourable outcome. A pigmented variant also has been described. The traditional criteria for the diagnosis of MPNST relied on the presence of a spindle cell tumour with ultrastructural evidence of Schwann cell differentiation arising in association with a nerve or a benign peripheral nerve tumour (i.e. neurofibroma) in a patient with neurofibromatosis. However, better recognition of MPNST's morphological features and the advent of immunohistochemical analysis has changed this view. MPNST expresses S-100 protein in approximately 50% of cases.

Cytogenetic studies have shown complex karyotypes in most cases, but structural abnormalities involving chromosomes 17 and 22 are occasionally found. Accumulation of mutant p53 protein was found to be higher in MPNSTs than in neurofibroma, suggesting a role for *TP53* mutations in neurofibroma transformation to MPNST. In addition, the proliferative marker Ki-67 tends to be higher and the cell cycle inhibitor $p27^{kip1}$ lower in MPNST than in neurofibroma.

The overall 5-year survival rates for patients with MPNST are approximately 50% for those with the sporadic form, 25% for those with NF1 and 10% for those with postirradiation MPNST. In two large series of cases, tumour size larger than 5 cm, history of NF1 and incomplete surgical excision were associated with a worse prognosis.

Clear Cell Sarcoma (Malignant Melanoma of Soft Parts)

Clear cell sarcoma of tendons and aponeuroses, also known as malignant melanoma of soft parts, is a rare soft

tissue sarcoma that tends to occur in adolescents and young adults. The foot and ankle, followed by the knee, are the most commonly affected sites, despite a wider anatomical distribution. Histologically, clear cell carcinoma is characterized by nests of polygonal or spindle cells separated by delicate fibrous septa. The cells have a clear or eosinophilic cytoplasm and prominent nucleolus. Mitoses are uncommon. Giant cells and melanin pigmentation are seen with some frequency. Ultrastructural studies show the presence of premelanosomes in approximately 50% of cases. Immunohistochemical analyses show expression of S-100 protein, HMB-45 and Melan A.

Cytogenetically, clear cell carcinoma is characterized by the chromosomal translocation t(12;22)(q13;q12), which results in the fusion of the DNA-binding domain of the *ATF-1* gene (*A*ctivating *T*ranscription *F*actor-*1*) located on chromosome 12q13 with the $5'$-end of the *EWS* gene. The chimaeric protein EWS/ATF-1 binds to ATF-DNA binding sites and leads to altered induction of cAMP-inducible promoters. Studies have shown that antibodies against the EWS/ATF-1 protein trigger tumour apoptosis. In addition, the somatostatin gene promoter may be a potential target for drug therapy.

Clear cell sarcoma is characterized by the development of metastasis in 50% of patients, most commonly to lungs and lymph nodes. The overall 5-year survival rate is 65% according to the results in one series. Factors associated with a more aggressive clinical course include larger tumour size, presence of necrosis and nondiploid DNA content. Local radiation therapy was associated with prolonged survival by univariate analysis in a series of 30 cases.

Sarcomas of Osseous or Cartilagenous Differentiation

Extraskeletal Osteosarcoma

Extraskeletal osteosarcoma (ESO) represents <1% of soft tissue sarcomas and occurs mainly after the sixth decade of life. Males are more commonly affected than females. The lower extremities, particularly the thigh, are more frequently involved. Histologically, ESO is similar to malignant fibrous histiocytoma, with a high degree of pleomorphism and brisk mitotic activity. Osteoid formation, present in ESO, differentiates it from malignant fibrous histiocytoma. Extensive areas of necrosis are also common. Immunohistochemical studies are largely nonspecific. However, the expression of osteocalcin and osteonectin, common bone proteins, by tumour cells and osteoid matrix can be used to differentiate ESO from other pleomorphic soft tissue sarcomas. Cytogenetic and molecular data are almost nonexistent, in contrast to data on bone osteosarcoma. Clinically, ESO follows an aggressive clinical course characterized by multiple local recurrences and metastases in >60% of patients. The 5-year overall survival rate is approximately 20–35%. Tumour size smaller than 5 cm, chondroblastic differentiation and a Ki-67 index <24% seem to be associated with a better survival.

Mesenchymal Chondrosarcoma

Mesenchymal chondrosarcoma occurs most often in bone (65%) and soft tissues (35%) of adolescents and young adults. In soft tissues, the most common anatomical sites are the head and neck region and the lower extremities. Histologically, it has a biphasic appearance in which islands of well-developed cartilage are intermixed with areas composed of primitive small ovoid cells. Haemangiopericytoma-like areas are also common. Immunohistochemical studies show expression of S-100 protein in the cartilage areas and MIC2 in the small round cells. Cytogenetic and molecular studies are rare. Investigation of the differential expression of collagen subtypes has suggested that mesenchymal chondrosarcoma has a line of differentiation consistent with premesenchymal chondroprogenitor cells. Mesenchymal chondrosarcoma is characterized by an aggressive clinical course with frequent lung metastasis. The overall 5-year survival rate is approximately 50%.

Extraskeletal Myxoid Chondrosarcoma

Extraskeletal myxoid chondrosarcoma (EMC) occurs predominantly in the lower extremities of individuals in the fifth to eighth decades of life. Histologically, EMC is a multilobular neoplasm composed of cords of small eosinophilic cells immersed in a myxoid matrix. Intracisternal bundles of parallel microtubules and scattered neurosecretory granules are characteristic but not specific ultrastructural findings. Immunohistochemical analysis shows the expression of vimentin and synaptophysin in most cases, indicating neuroendocrine differentiation, and S-100 protein in a minority of cases. Cytogenetic and molecular studies have shown the presence of the chromosomal translocations t(9;22)(q22;q12) and t(9;17) (q22;q11) in 75 and 25% of cases, respectively. In these translocations the *CHN* gene, a member of the steroid/thyroid receptor gene superfamily located on chromosome 9q22 is fused with either *EWS* on chromosome 22q12 or $RPB56/hTAF_{II}68$ on chromosome 17q11, respectively. Clinically, EMC is characterized by an indolent clinical course in most but not all patients. Metastases, most commonly to the lungs, occur in up to 50% of patients. The 5 and 10-year overall survival rates are 90% and 70–80%, respectively.

Sarcomas of Uncertain Line of Differentiation

This group of sarcomas is mainly represented by synovial sarcoma, epithelioid sarcoma, alveolar soft part sarcoma, desmoplastic small round cell tumour and extrarenal rhabdoid tumour.

Synovial Sarcoma

Synovial sarcoma was placed in the category of sarcomas of uncertain line of differentiation because its putative synovial differentiation has not been confirmed. It represents approximately 10% of all soft tissue sarcomas, but this figure is expected to increase owing to better recognition of monophasic and poorly differentiated variants using immunohistochemical, cytogenetic and molecular techniques. Synovial sarcoma has a slight predilection for males and occurs at any age, particularly affecting adolescents and young adults. It can arise in any site but is more common in the lower extremities. Despite its common periarticular location, intra-articular synovial sarcoma is exceedingly rare.

Histologically, synovial sarcoma is classified into biphasic, monophasic and poorly differentiated variants. The biphasic variant is composed of epithelial and spindle cell components. The epithelial component has a glandular appearance; the spindle cell component comprises fascicles of spindle cells frequently arranged in a herringbone pattern similar to that in fibrosarcoma (**Figure 7**). Monophasic synovial sarcoma can be purely epithelial, which is very rare, or purely spindled. Areas with a haemangiopericytoma-like vascular pattern as well as mast cell infiltration are common. Calcification and metaplastic ossification are also frequent, being detected in plain films in up to 30% of cases. The undifferentiated form of synovial sarcoma has a round cell shape that may simulate Ewing sarcoma/PNET.

Immunohistochemical analysis shows expression for vimentin, epithelial membrane antigen (EMA) and cytokeratin. EMA and cytokeratin are commonly expressed by the epithelial component and only focally expressed by the spindle cell component. In addition, it has been shown that synovial sarcoma commonly expresses cytokeratins 7, 13 and 19 but not cytokeratins 8 and 18. S-100, MIC2, and the antiapoptotic protein Bcl-2 also are expressed in many cases. More recently, *Her-2/neu* gene amplification and protein overexpression were found in cases of synovial sarcoma.

With the use of modern cytogenetic techniques, the characteristic translocation t(X;18)(p11.2;q11.2) is found in most cases (>99%), including cases of poorly differentiated tumours (**Table 2**). This translocation results in the fusion of the *SYT* gene (*SY*novial sarcoma *T*ranslocation) on chromosome 18q11.2 with members of the *SSX* family of genes (*S*ynovial *S*arcoma *X* chromosome breakpoint) on chromosome Xp11.2 (*SSX-1*, *SSX-2*, and, more rarely, *SSX-4*; the other members, *SSX-3*, *SSX-5*, and *SSX-6* have not been detected in synovial sarcoma translocations) (Clark *et al*., 1994). Whereas *SYT* encodes for a ubiquitously expressed nuclear protein that seems to function as a transcriptional coactivator, *SSX* encodes for transcription repressor proteins. The mechanism of action of the chimaeric protein remains unknown, but it has been suggested that it works through recruitment of the PcG complex of proteins involved in transcriptional repression.

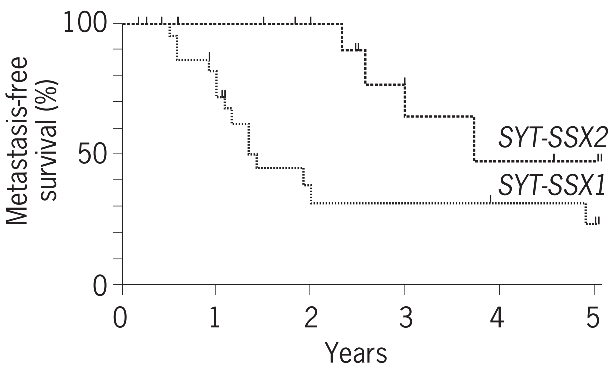


Figure 10 In synovial sarcoma, patients with the *SYT/SSX-1* fusion transcript had shorter metastasis-free survival than patients with the *SYT/SSX-2* fusion transcript (hazard ratio = 3; 95% confidence interval = 1.1–8; $P = 0.03$; Cox regression model). (From Kawai *et al.*, 1998, SYT-SSX gene fusion as a determinant of morphology and prognosis in synovial sarcoma. *New England Journal of Medicine*, **338**, 153–160, by permission of the Massachusetts Medical Society.)

It has also been shown that the *SYT/SSX-1* fusion transcript, which is detected in 65% of cases, is associated with biphasic histological findings, a high proliferation rate, and shorter metastasis-free survival (**Figure 10**). The specificity of the t(X;18) for synovial sarcomas has been questioned because it also has been found in some cases of MPNST, suggesting that MPNST and synovial sarcoma may be part of the same spectrum of tumours. However, these intriguing results need further confirmation.

Many adverse prognostic factors have been proposed for synovial sarcoma, including older age, larger tumour size, poorly differentiated histological features, nuclear grade, necrosis, mitotic activity, low mast cell infiltration, Ki-67 expression, *TP53* mutations, bone or neurovascular invasion, *SYT/SSX-1* chimaeric gene and hepatocyte growth factor. However, the clinical utility of most of these prognostic factors remains to be settled. Synovial sarcoma metastasizes in 50% of cases, most commonly to lungs and lymph nodes, and 5-year overall survival rates between 24 and 64% have been reported. Treatment involves surgical excision, radiotherapy and chemotherapy. Chemotherapy often includes ifosfamide because synovial sarcoma seems to be particularly sensitive to this drug.

Epithelioid Sarcoma

Epithelioid sarcoma is a rare soft tissue sarcoma that occurs predominantly in adolescents and young adults. It presents most commonly as dermal or subcutaneous nodules in the distal portions of the upper extremities, particularly the hands and wrists. However, its anatomical distribution is wider. Epithelioid sarcoma occurring

around the pelvic region also has been termed proximal type and seems to follow a more aggressive clinical course. Histologically, epithelioid sarcoma has a multinodular architecture with a granuloma-like appearance. The nodules are composed of epithelioid or spindle cells often surrounded by an inflammatory infiltrate rich in lymphocytes. Necrotic areas are common, often assuming a map-like pattern. Immunohistochemical studies show the expression of vimentin, EMA and cytokeratin in most cases and CD34 in approximately half.

Cytogenetic studies show frequent rearrangements or deletions involving chromosome 22, particularly the band 22q11. Interestingly, this chromosomal band often is deleted or rearranged in rhabdoid tumour (see Extrarenal Malignant Rhabdoid Tumour, below), an entity that shares some degree of morphological overlap with epithelioid sarcoma. Epithelioid sarcoma has a high local recurrence rate, up to 80% in some series, owing to its extensively infiltrative margins along tendons and fascia, often well beyond the apparent clinical limits of the tumour. Metastases are very common, mostly to regional lymph nodes, lungs and scalp. The estimated 5-year survival rate is approximately 50–70%. Some prognostic factors associated with a worse outcome include tumour size larger than 5 cm, necrosis, vascular invasion, lymph node metastasis and proximal location.

Alveolar Soft Part Sarcoma

Alveolar soft part sarcoma (ASPS), also known as Christopherson tumour, represents <1% of soft tissue sarcomas and commonly occurs in lower extremities in patients between the second and third decades of life. In children, the head and neck are preferentially involved. Female patients are slightly more affected. Histologically, ASPS is characterized by a pseudoalveolar architecture composed of richly vascularized fibrous septa lined by polygonal cells with low mitotic activity and minimal pleomorphism. PAS-positive intracytoplasmic rod-like inclusions are common and are seen to have a characteristic crystalline rhomboid shape on ultrastructural analysis. The line of differentiation of ASPS is still subject to debate, and a putative skeletal muscle differentiation has been suggested. However, the expression of muscle markers, such as desmin and MyoD1, is inconsistently detected.

Cytogenetic and molecular studies have shown that the nonreciprocal translocation t(X;17)(p11.2;q25) is characteristic of ASPS (100% of cases) and results in the fusion of the transcription factor *TFE3* gene (*T*ranscription *F*actor binding to IGHM *E*nhancer *3*) on chromosome Xp11.2 with the *ASPL* gene (*A*lveolar *S*oft *P*art sarcoma *L*ocus) on chromosome 17q25. Interestingly, the *TFE3* gene commonly is rearranged in some cases of renal papillary carcinoma. ASPS is associated with a poor long-term prognosis, with 5- and 10-year overall survival rates of 46–67% and 38%, respectively. Metastases to lungs, brain and bone are common. Younger age and tumours smaller than 10 cm in diameter tend to be associated with a better prognosis.

Desmoplastic Small Round Cell Tumour

Desmoplastic small round cell tumour (DSRCT) is a highly malignant sarcoma that most commonly occurs in the peritoneal cavity of adolescents and young adults of the male sex. DSRCT usually forms large intra-abdominal whitish masses that are histologically characterized by nests or strands of small round cells surrounded by a dense desmoplastic stroma. Mitoses, apoptotic bodies, necrosis and lymphatic invasion are common. Immunohistochemical studies show the expression of several proteins, including cytokeratin, EMA, neuron-specific enolase and desmin, supporting its divergent differentiation. Desmin expression is usually paranuclear and of globular appearance. MIC2 expression is not uncommon and may cause problems in the differential diagnosis with Ewing sarcoma/PNET.

Cytogenetic analyses show the chromosomal translocation t(11;22)(p13;q12) in almost all cases of DSRCT. This results in the fusion of the tumour-suppressor gene *WT1* on chromosome 11p13 to the *EWS* gene on chromosome 22q12. The causal mechanism of this translocation is unknown but insertion of LINE-1 DNA sequences at the genomic breakpoint of DSRCT has been identified. The chimaeric product has oncogenic attributes and seems to activate the *IGF1* gene promoter with higher affinity that the WT1 protein alone. DSRCT follows a highly aggressive clinical course, often characterized by multiple local recurrences. Distant metastases are uncommon. Despite intensive chemotherapy regimens, radical surgical procedures and radiotherapy, most patients die less than 3 years after diagnosis.

Extrarenal Malignant Rhabdoid Tumour

Malignant rhabdoid tumour (MRT) used to be a highly controversial entity because the so-called rhabdoid phenotype can be found in several other tumours. However, recent findings have shed some light on this issue. Currently, it seems that renal MRT and extrarenal MRT are characterized by deletions or mutations of the chromatin remodelling gene *hSNF5/INI1* located on chromosome 22q11.2. In addition, germ-line mutations of the *hSNF5/INI1* gene also have been described, predisposing the carriers to renal MRT, extrarenal MRT and other tumours in the so-called rhabdoid predisposition syndrome. In soft tissues, MRT tends to occur in young patients and in a wide anatomical distribution. Histologically, MRT is characterized by sheets of polygonal cells with hyaline globular cytoplasmic inclusions and vesicular nuclei. By immunohistochemical analysis, the cells consistently express vimentin, cytokeratin and EMA. Ultrastructural analysis demonstrates that the cytoplasmic inclusions are composed

of whorls of intermediate filaments. Clinically, soft tissue MRTs follow a very aggressive clinical course with higher rates of metastasis. Death usually occurs less than 2 years after diagnosis.

PROGNOSTIC FACTORS

Several prognostic factors have been proposed for soft tissue sarcomas, but only a few are universally accepted and routinely used. The large void between the identification and the clinical application of prognostic factors for soft tissue sarcoma exists for many reasons. First, the prognostic factors proposed so far have been incompletely validated. Theoretically, the validation of proposed prognostic factors is carried out in three major steps: exploratory studies, retrospective clinical studies and prospective clinical trials. The first step is the study of a plausible prognostic factor in relationship to outcomes and the predictive data available. This step is fundamental to hypothesis formulation. The second step confirms the initial observations through large studies in which multivariate analysis establishes the independence of the investigated prognostic factor in relation to known predictors. Definition and validation of cutoff points in two or more distinct data sets are tasks involved in the second step. The third step consists of well-designed multi-institutional prospective studies with a large number of patients using pre-established cutoff points (Hall and Going, 1999). Unfortunately, almost all studies of prognostic factors in soft tissue sarcomas have not gone beyond the second step.

Other important problems include (1) the uneven representation of specific sarcomas in heterogeneous series, (2) the number and type of variables included in multivariate analysis and how they are selected and modelled, including pre-establishment of cutoffs, (3) the statistical power of the study for detecting differences between groups of patients, (4) standardization of specific outcomes (local recurrence, metastasis, disease-free and overall survival) and (5) appropriateness of the statistical methods. In addition, the identification of a prognostic factor does not necessarily mean that it has clinical utility. A meaningful prognostic factor should also influence clinical decisions.

Despite numerous problems associated with histological grading systems, histological grade has been considered the most powerful prognostic factor in heterogeneous series of patients with soft tissue sarcoma (**Figure 3**) and frequently incorporates other factors such as mitotic count, necrosis and degree of differentiation (see Grading and Staging Systems, above). In a retrospective series involving more than 1000 cases of soft tissue sarcomas, the presence of high-grade histological features was the most powerful independent prognostic factor for metastasis-free survival ($P = 0.0001$; relative risk (RR) = 4.3; confidence interval (CI) = 2.6–6.9) and disease-specific overall survival ($P = 0.0001$; RR = 4.0; CI = 2.5–6.6), but not for local recurrence-free survival (Pisters *et al.*, 1996). Tumour size has been considered an adverse prognostic factor in many but not all series including more than 100 patients. Standardization of cutoffs for tumour size and other problems already described partially explain these divergent results.

Several other prognostic factors have been evaluated. DNA ploidy, S-phase fraction, AgNORs (silver staining for nucleolar organizer regions; segments of DNA containing ribosomal genes whose expression reflects the proliferation activity of the tumour), Ki-67, Mdm2, p53, *RB1* gene product, c-Myc, PDGFβ (platelet-derived growth factor β) and Bcl-2 have been correlated with histological grade and decreased metastasis-free or overall survival in many but not all studies. Overexpression of the transmembrane p-glycoprotein, product of the multiple resistance gene 1 (*MDR-1*), has been correlated with decreased sensitivity to chemotherapy and survival. Specific chimaeric mRNA transcripts have been shown to have a prognostic role in synovial sarcoma and Ewing sarcoma (**Figure 9**) (see Extraskeletal Ewing Sarcoma and Peripheral Primitive Neuroectodermal Tumour, above, and Synovial Sarcoma, above).

TREATMENT

The treatment of soft tissue sarcomas commonly combines combination of surgical excision, radiation therapy and chemotherapy. The therapeutic approach is chosen after adequate imaging and pathological studies and determination of whether the patient has localized, locally recurrent or metastatic disease. With this information, a specific therapeutic plan is best chosen based on individual characteristics of the patient and results obtained in clinical trials.

Overview of Sarcoma Treatment

Surgical excision is used primarily for the treatment of localized or locally recurrent sarcoma and is the only potentially curative treatment at present. However, surgical excision also can be used for the treatment of metastatic sarcoma, particularly pulmonary metastasis (metastasectomies). Radiation therapy is used primarily as an adjuvant treatment for the control of localized or locally recurrent sarcoma. Chemotherapy is used for localized and metastatic sarcomas.

Treatment of Clinical Groups

Localized Sarcoma

Local disease is defined when a soft tissue sarcoma is confined to a certain anatomical site or compartment without clinicoradiological evidence of systemic disease

(metastasis). In this situation, surgical excision is considered the mainstay therapy because it is the only potentially curative modality at present.

Surgical excision can be divided into five basic types: intralesional, marginal, wide, radical and amputation. In an intralesional excision (also known as curettage, debulking or incisional biopsy), the tumour is incompletely resected and gross disease is left behind. This surgical procedure is performed when broader excisions cannot be carried out without affecting vital structures. In marginal excisions, the tumour is excised through its pseudocapsule or perilesional reactive fibroinflammatory zone. Because there is a substantial risk for microscopic positive margins after this procedure, marginal excision is commonly indicated for benign soft tissue tumours or after successful preoperative radiotherapy or chemotherapy for sarcomas. Local recurrence rates of 60% or more have been reported with this type of resection alone.

Wide excision, which is currently the most common type of surgical excision performed, includes a rim of normal tissue surrounding the tumour. Wide excision is used for low-grade and high-grade sarcomas with or without adjuvant treatment. Local recurrent rates of up to 30% have been reported when wide excision is performed alone. Because of this, there is a general arbitrary recommendation that a surgical margin of at least 2 cm around the tumour be achieved. Radical excision involves the removal of the entire anatomical compartment, i.e. all tissues within the natural anatomical boundaries. Radical excision often is indicated for recurrent sarcomas or when imaging studies were inconclusive for defining the tumour's anatomical limits. Amputation, often used in the past, is performed uncommonly now because a limb-sparing operation followed by postoperative radiotherapy provide similar results. Amputation is indicated mainly (1) when it is impossible to obtain adequate margins without compromising vital structures in primary or locally recurrent disease, (2) when the dose and volume of adjuvant radiotherapy are likely to cause important local complications and (3) in certain situations when the use of a prosthesis will provide better postoperative results than a radical limb-sparing operation (O'Sullivan *et al.*, 1999).

Currently, radiotherapy commonly is used in combination with surgical excision for the management of localized soft tissue sarcoma. Radiotherapy can be delivered in the form of external-beam radiation or brachytherapy (from the Greek *brachys*, short). In the latter, the source of radiation is placed close to or in the tumour bed. Several retrospective and few prospective clinical trials have shown that the combination of radiotherapy and surgical excision provides better results than surgical excision alone in the management of localized soft tissue sarcoma. However, there is no clear evidence that local control of the disease automatically confers prolonged survival. Preoperative radiation therapy has the advantage of using smaller radiation fields and doses but has been associated with a higher risk of wound complications. Postoperative radiation therapy decreases the latter risk. Advantages of brachytherapy include cost-effectiveness and a short interval between the surgical procedure and the beginning of radiation therapy. However, brachytherapy seems to provide results that are inferior to those of external-beam radiation for low-grade tumours.

Several other types of adjuvant radiotherapy have been used for local disease control, including a combination of brachytherapy and external beam radiation, intraoperative radiotherapy, hyperfractionation, the use of radiosensitizers, and others. However, none has been shown to be superior to external-beam radiation.

The role of adjuvant chemotherapy for localized soft tissue sarcoma remains controversial despite convincing evidence of its clear though modest benefit (Benjamin, 1999). Despite the lack of statistical significance, most randomized trials have shown a clear trend for prolonged disease-free survival and overall survival in patients who received adjuvant chemotherapy with multiple regimens. Moreover, meta-analysis of some reported series has confirmed these impressions (**Figure 11**) (Sarcoma Meta-analysis Collaboration, 1997). Reasons accounting for some of the persistent controversies include large variation in the criteria for patient enrollment (tumour location, histological subtype and grade, size and depth), follow-up periods and chemotherapy regimens used.

Neoadjuvant or preoperative chemotherapy has been advocated for treating localized soft tissue sarcomas based on three theoretical assumptions: (1) elimination of occult micrometastases present at diagnosis, before the development of chemoresistant tumour clones, (2) reduction of tumour volumes, to allow less radical surgical procedures and (3) clinical and pathological assessment of tumour chemoresistance by an *in vivo* test (Pisters and Brennan, 2000). Nonrandomized studies have shown complete and partial response rates between 3 and 64% with the use of neoadjuvant chemotherapy. However, these results did not translate into a better outcome (metastasis-free or overall survival) for responders as compared with nonresponders or patients treated with adjuvant chemotherapy. In spite of these findings, patients with large high-grade tumours should be considered for neoadjuvant chemotherapy because of a high risk of development of metastasis.

Locally Recurrent Sarcoma

Soft tissue sarcomas recur in up to 90% of cases according to the adequacy of the surgical margins. Because head and neck and retroperitoneal sarcomas are difficult to resect without affecting vital structures, local recurrence rates are much higher than those observed in extremity sarcomas. The treatment of locally recurrent soft tissue sarcoma should be highly individualized, but surgical re-excision of the recurrent tumour should be attempted when feasible because good long-term results have been obtained with this approach. Amputation is indicated especially when dealing with

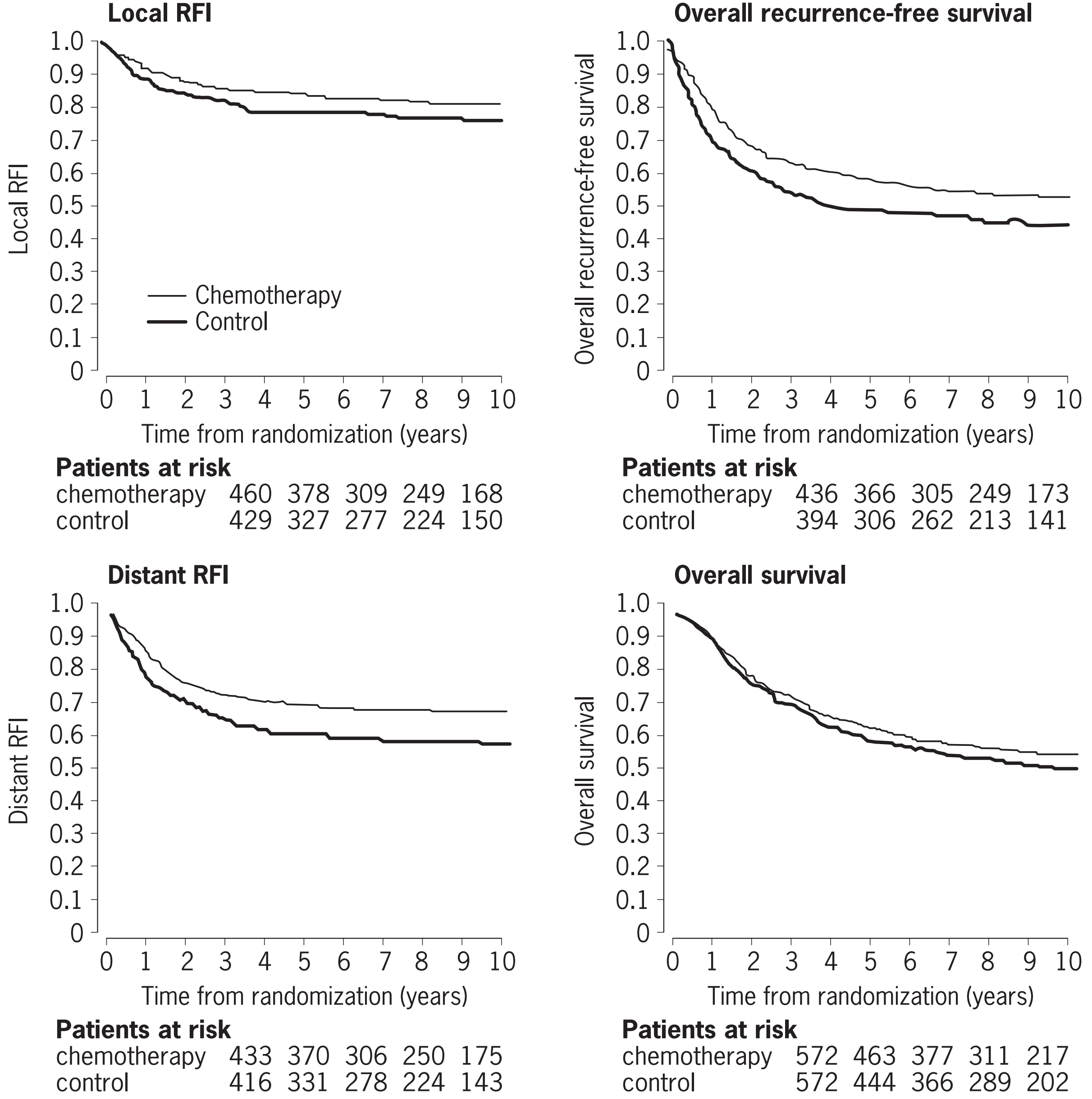


Figure 11 Kaplan–Meier curves obtained with a quantitative meta-analysis of data on 1568 patients from 14 clinical trials comparing doxorubicin-based adjuvant chemotherapy with control in localized soft tissue sarcoma. Prolonged local recurrence-free (hazard ratio (HR) = 0.73; 95% confidence interval (CI) = 0.56–0.94; $P = 0.016$), metastasis-free (HR = 0.70; 95% CI = 0.57–0.85; $P = 0.0003$), disease-free (HR = 0.75; 95% CI = 0.64-0.87; $P = 0.0001$) and overall survival (HR = 0.89; 95% CI = 0.76-1.03; $P = 0.12$) were observed. However, no statistical significance was reached for overall survival. RFI, recurrence-free interval. (From Sarcoma Meta-analysis Collaboration, 1999, Adjuvant chemotherapy for localized resectable soft-tissue sarcoma of adults: meta-analysis of individual data. *Lancet*, **350**, 1647–1654, by permission of The Lancet, Ltd.)

multiple local recurrences or tumour involvement of important neurovascular structures. The use of adjuvant radiation often is indicated if the tumour was not previously irradiated. Using this approach, local control was obtained in up to 80% of patients in some series. Additional radiation therapy in previously irradiated recurrent tumours should be evaluated on an individual basis, but local control can be attained in most patients. The role of chemotherapy for locally recurrent disease remains to be established.

Metastatic Sarcoma

Metastatic soft tissue sarcomas are primarily managed with chemotherapy or surgical excision (metastasectomy).

Chemotherapy is the main therapeutic modality and is indicated for nonresectable metastases. The most effective single agents are doxorubicin and ifosfamide. They induce objective response in more than 20% of patients and exhibit a dose–response relationship (O'Byrne and Steward, 1999). Doxorubicin-based combination regimens such as CYVADIC (cyclophosphamide, vincristine, doxorubicin (Adriamycin) and dacarbazine) and MAID (mesna, doxorubicin, ifosfamide and dacarbazine) have led to response rates in up to 70% of patients, but no improvement in overall survival has been demonstrated. Furthermore, no difference in terms of response rate, remission duration or overall survival was seen in a large prospective randomized phase III trial comparing CYVADIC, doxorubicin–ifosfamide, and doxorubicin alone for treatment of advanced soft tissue sarcoma. The concomitant use of recombinant human granulocyte macrophage colony-stimulating factor (hrGM-CSF) has been used by some investigators to allow more intensive chemotherapy regimens. Other treatments include hyperthermic isolated limb perfusion (HILP) using an antiangiogenic factor, such as recombinant tumour necrosis factor alpha (rTNFα), in combination with cytotoxic drugs and whole-body hyperthermia.

Surgical excision of metastatic disease has been investigated in several studies; the benefits are marginal. In a series of 135 patients who were treated at a single institution for metastatic sarcoma to the lungs, complete excision, incomplete excision and no excision of the metastatic sarcoma were associated with median survivals of 19, 10 and 8 months, respectively. However, the 3-year overall survival rate was only 23% for those in whom complete metastasectomy was achieved (Gadd *et al.*, 1993). A series evaluating 230 patients with pulmonary and nonpulmonary metastasis found that unresectability of distant metastases correlated with a shorter survival by multivariate analysis ($P = 0.0001$; HR = 2.3; CI = 1.2–3.7; **Figure 12**) (Billingsley *et al.*, 1999). Other investigations found similar numbers, with median survivals after development of pulmonary metastases ranging from 18 to 27 months. Comparison of surgical excision alone with the combination of surgical excision and chemotherapy has provided discordant results among studies. A prognostic system for use with patients who have pulmonary metastases has been proposed. It is based on the presence of metastases, number of metastases, metastasis-free period less than 18 months and size of the metastases larger than 2 cm (Choong *et al.*, 1995). It seems that complete metastasectomy is the single most important factor for prolonged postmetastasis survival and that careful patient selection is critical for achieving good results. Four criteria for pulmonary metastasectomy have been suggested: ability to control the primary tumour, absence of extra-thoracic metastases, clinical condition satisfactory for surgical excision and absence of a better treatment option (McCormack, 1990).

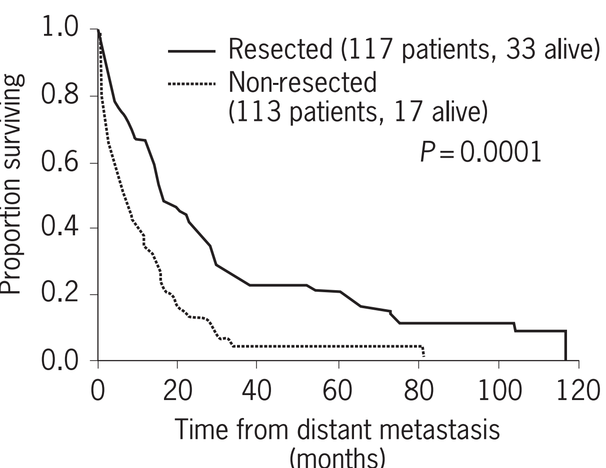


Figure 12 Overall survival curves comparing the effect of complete resection of metastatic soft tissue sarcoma with incomplete or no resection. Shorter overall survival was seen in patients with nonresected metastases (hazard ratio = 2.3; 95% confidence interval = 1.2–3.7; $P = 0.0001$). (From Billingsley *et al.*, 1999, by permission of the American Cancer Society.)

Future Treatments

Alternative treatments for soft tissue sarcomas have been investigated with promising results. These treatments include use of the peroxisome proliferator-activated receptor-gamma ligand troglitazone to induce terminal adipocytic differentiation in liposarcomas, antisense *FLI/EWS* and anti-MIC2 (CD99) for Ewing sarcoma/PNET, tyrosine kinase inhibitors (STI-576) for c-*KIT* proto-oncogene product expressing tumours, angiogenesis inhibitors and others. Additional information on current clinical trials for the treatment of sarcomas can be found at http://clinicaltrials.gov.

Websites

http://www.cancerlinks.org/sarcoma.html.

http://www.cancernet.nci.nih.gov.

REFERENCES

Antman, K. and Chang, Y. (2000). Kaposi's sarcoma. *New England Journal of Medicine*, **342**, 1027–1038.

Arndt, C. A. and Crist, W. M. (1999). Common musculoskeletal tumors of childhood and adolescence. *New England Journal of Medicine*, **341**, 342–352.

Aurias, A., *et al.* (1983). Chromosomal translocations in Ewing's sarcoma (letter to the editor). *New England Journal of Medicine*, **309**, 496–497.

Ballo, M. T., *et al*. (1999). Desmoid tumor: prognostic factors and outcome after surgery, radiation therapy, or combined surgery and radiation therapy. *Journal of Clinical Oncology*, **17**, 158–167.

Barr, F. G. (1998). Translocations, cancer and the puzzle of specificity. *Nature Genetics*, **19**, 121–124.

Bell, D. W., *et al*. (1999). Heterozygous germ line hCHK2 mutations in Li–Fraumeni syndrome. *Science*, **286**, 2528–2531.

Benjamin, R. S. (1999). Evidence for using adjuvant chemotherapy as standard treatment of soft tissue sarcoma. *Seminars in Radiative Oncology*, **9**, 349–351.

Billingsley, K. G., *et al*. (1999). Multifactorial analysis of the survival of patients with distant metastasis arising from primary extremity sarcoma. *Cancer*, **85**, 389–395.

Cance, W. G., *et al*. (1990). Altered expression of the retinoblastoma gene product in human sarcomas. *New England Journal of Medicine*, **323**, 1457–1462.

Choong, P. F., *et al*. (1995). Survival after pulmonary metastasectomy in soft tissue sarcoma. Prognostic factors in 214 patients. *Acta Orthopaedic Scandinavica*, **66**, 561–568.

Clark, J., *et al*. (1994). Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nature Genetics*, **7**, 502–508.

Cordon-Cardo, C., *et al*. (1994). Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas. *Cancer Research*, **54**, 794–799.

Costa, J., *et al*. (1984). The grading of soft tissue sarcomas. Results of a clinicohistopathologic correlation in a series of 163 cases. *Cancer*, **53**, 530–541.

Crozat, A., *et al*. (1993). Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature*, **363**, 640–644.

de Alava, E. and Gerald, W. L. (2000). Molecular biology of the Ewing's sarcoma/primitive neuroectodermal tumour family. *Journal of Clinical Oncology*, **18**, 204–213.

de Alava, E., *et al*. (1998). EWS-FLI1 fusion transcript structure is an independent determinant of prognosis in Ewing's sarcoma. *Journal of Clinical Oncology*, **16**, 1248–1255.

Enneking, W. F., *et al*. (1980). A system for the surgical staging of musculoskeletal sarcoma. *Clinical Orthopaedics*, **18**, 106–120.

Enzinger, F. M. and Weiss, S. W. (1995). *Soft Tissue Tumours*, 3rd edn (Mosby, St. Louis).

Fleming, I. D., *et al*. (1997). *AJCC Cancer Staging Manual*. *American Joint Committee on Cancer*, 5th edn (Lippincott-Raven, Philadelphia).

Gadd, M. A., *et al*. (1993). Development and treatment of pulmonary metastases in adult patients with extremity soft tissue sarcoma. *Annals of Surgery*, **218**, 705–712.

Galili, N., *et al*. (1993). Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. *Nature Genetics*, **5**, 230–235.

Gill, P. S., *et al*. (1998). Evidence for multiclonality in multicentric Kaposi's sarcoma. *Proceedings of the National Academy Sciences of the USA*, **95**, 8257–8261.

Guillou, L., *et al*. (1997). Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. *Journal of Clinical Oncology*, **15**, 350–362.

Gustafson, P., *et al*. (1992). Soft tissue leiomyosarcoma. A population-based epidemiologic and prognostic study of 48 patients, including cellular DNA content. *Cancer*, **70**, 114–119.

Hall, P. A. and Going, J. J. (1999). Predicting the future: a critical appraisal of cancer prognosis studies. *Histopathology*, **35**, 489–494.

Hashimoto, H., *et al*. (1985). Malignant smooth muscle tumours of the retroperitoneum and mesentery: a clinicopathologic analysis of 44 cases. *Journal of Surgical Oncology*, **28**, 177–186.

Kilpatrick, S. E., *et al*. (1996). The clinicopathologic spectrum of myxoid and round cell liposarcoma. A study of 95 cases. *Cancer*, **77**, 1450–1458.

Kelly, K. M., *et al*. (1997). Common and variant gene fusions predict distinct clinical phenotypes in rhabdomyosarcoma. *Journal of Clinical Oncology*, **15**, 1831–1836.

Kransdorf, M. J. and Murphey, M. D. (1997). *Imaging of Soft Tissue Tumours*. (W. B. Saunders, Philadelphia).

Ladanyi, M. and Bridge, J. A. (2000). Contribution of molecular genetic data to the classification of sarcomas. *Human Pathology*, **31**, 532–538.

Landis, S. H., *et al*. (1999). Cancer statistics. *CA Cancer Journal for Clinicians*, **49**, 8–31.

Lawrence, B., *et al*. (2000). TPM3-ALK and TPM4-ALK oncogenes in inflammatory myofibroblastic tumours. *American Journal of Pathology*, **157**, 377–384.

Le Doussal, V., *et al*. (1996) Prognostic factors for patients with localized primary malignant fibrous histiocytoma: a multicenter study of 216 patients with multivariate analysis. *Cancer*, **77**, 1823–1830.

McClain, K. L., *et al*. (1995). Association of Epstein–Barr virus with leiomyosarcomas in children with AIDS. *New England Journal of Medicine*, **332**, 12–18.

McCormack, P. (1990). Surgical resection of pulmonary metastases. *Seminars in Surgical Oncology*, **6**, 297–302.

Meis-Kindblom, J. M. and Kindblom, L. G. (1998). Angiosarcoma of soft tissue: a study of 80 cases. *American Journal of Surgical Pathology*, **22**, 683–697.

Mingoli, A., *et al*. (1991). Leiomyosarcoma of the inferior vena cava: analysis and search of world literature on 141 patients and report of three new cases. *Journal of Vascular Surgery*, **14**, 688–699.

Newton, W. A. Jr, *et al*. (1995). Classification of rhabdomyosarcomas and related sarcomas. Pathologic aspects and proposal for a new classification – an Intergroup Rhabdomyosarcoma Study. *Cancer*, **76**, 1073–1085.

O'Byrne, K. and Steward, W. P. (1999). The role of chemotherapy in the treatment of adult soft tissue sarcomas. *Oncology*, **56**, 13–23.

Oliner, J. D., *et al*. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature*, **358**, 80–83.

Oliveira, A. H., *et al*. (2000). p27[kipl] protein expression correlates with survival in myxoid and round-cell liposarcoma. *Journal of Clinical Oncology*, **18**, 2888–2893.

Orkin, S. H. (2000). Stem cell alchemy. *Nature Medicine*, **6**, 1212–1213.

O'Sullivan, B., *et al*. (1999). The local management of soft tissue sarcoma. *Seminars in Radiative Oncology*, **9**, 328–348.

Pisters, P. W., *et al*. (1996). Analysis of prognostic factors in 1,041 patients with localized soft tissue sarcomas of the extremities. *Journal of Clinical Oncology*, **14**, 1679–1689.

Pisters, P. W. T. and Brennan, M. F. (2000). Sarcomas of soft tissue. In: Abeloff, M. D., *et al*. (eds), *Clinical Oncology*, 2nd edn. 2273–2313 (Churchill Livingstone, New York).

Qualman, S. J., *et al*. (1998). Intergroup Rhabdomyosarcoma Study: update for pathologists. *Pediatric Development Pathology*, **1**, 550–561.

Rabbitts, T. H., *et al*. (1993). Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. *Nature Genetics*, **4**, 175–180.

Sandberg, A. A. and Bridge, J. A. (2000). Updates on cytogenetics and molecular genetics of bone and soft tissue tumours: Ewing sarcoma and peripheral primitive neuroectodermal tumors. *Cancer Genetics and Cytogenetics*, **123**, 1–26.

Sarcoma Meta-analysis Collaboration (1997). Adjuvant chemotherapy for localised resectable soft-tissue sarcoma of adults: meta-analysis of individual data. *Lancet*, **350**, 1647–1654.

Scott, S. M., *et al*. (1989) Soft tissue fibrosarcoma. A clinicopathologic study of 132 cases. *Cancer*, **64**, 925–931.

Simon, M. P., *et al*. (1997). Deregulation of the platelet-derived growth factor B-chain gene via fusion with collagen gene COL1A1 in dermatofibrosarcoma protuberans and giant-cell fibroblastoma. *Nature Genetics*, **15**, 95–98.

Terrier, P., *et al*. (1996). Small round blue cell tumors in bone: prognostic factors correlated to Ewing's sarcoma and neuroectodermal tumours. *Seminars in Diagnostic Pathology*, **13**, 250–257.

Trojani, M., *et al*. (1984). Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *International Journal of Cancer*, **33**, 37–42.

Turc-Carel, C., *et al*. (1983). Chromosomal translocations in Ewing's sarcoma (letter to the editor). *New England Journal of Medicine*, **309**, 497–498.

Weiss, S. W. and Enzinger, F. M. (1978). Malignant fibrous histiocytoma: an analysis of 200 cases. *Cancer*, **41**, 2250–2266.

Weiss, S. W. and Sobin, L. H. (1994). *Histological Typing of Soft Tissue Tumours* (Springer, Berlin).

FURTHER READING

Bridge, J. A. and Sandberg, A. A. (2000). Cytogenetic and molecular genetic techniques as adjunctive approaches in the diagnosis of bone and soft tissue tumors. *Skeletal Radiology*, **29**, 249–258.

Delattre, O., *et al*. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature*, **359**, 62–165.

Drobnjak, M., *et al*. (1994). Prognostic implications of p53 nuclear overexpression and high proliferation index of Ki-67 in adult soft-tissue sarcomas. *Journal of the National Cancer Institute*, **86**, 549–554.

Khan, J., *et al*. (1999). cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. *Proceedings of the National Academy Sciences of the USA*, **96**, 13264–13269.

Knezevich, S. R., *et al*. (1998). A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nature Genetics*, **18**, 184–187.

May, W. A., *et al*. (1997). EWS/FLI1-induced manic fringe renders NIH 3T3 cells tumorigenic. *Nature Genetics*, **17**, 495–497.

Ordonez, N. (1997). Application of immunohistochemistry in the diagnosis of soft tissue sarcomas: a review and update. *Advances in Anatomical Pathology*, **5**, 67–85.

Pisters, P. W. and Pollock, R. E. (1999). Staging and prognostic factors in soft tissue sarcoma. *Seminars in Radiative Oncology*, **9**, 307–314.

Rubin, B. P., *et al*. (1998). Congenital mesoblastic nephroma t(12;15) is associated with ETV6-NTRK3 gene fusion: cytogenetic and molecular relationship to congenital (infantile) fibrosarcoma. *American Journal of Pathology*, **153**, 1451–1458.

Somerhausen, N. S. A. and Fletcher, C. D. M. (1999). Soft-tissue sarcomas: An update. *European Journal of Surgical Oncology*, **25**, 215–220.

Shapiro, D. N., *et al*. (1993). Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. *Cancer Research*, **53**, 5108–5112.

Zucman, J., *et al*. (1993). EWS and ATF-1 gene fusion induced by t(12;22) translocation in malignant melanoma of soft parts. *Nature Genetics*, **4**, 341–345.

Chapter 49

Pleura and Peritoneum

Jonathan A. Fletcher
Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA

Christopher N. Otis
Tufts University School of Medicine, Baystate Medical Center, Springfield, MA, USA

CONTENTS

- Introduction
- Normal Development and Structure
- Aetiology
- Screening and Prevention
- Clinical Presentation and Pathology
- Cytogenetic and Molecular Genetic Findings
- Prognostic Factors
- Overview of Present Clinical Management

INTRODUCTION

Primary cancers of the pleura and peritoneum are uncommon in the general population. Nonetheless, such cancers – particularly those which arise from mesothelial cells and are therefore known as malignant mesothelioma – have commanded substantial attention in recent years. There are several reasons why mesotheliomas are subject to active discussion in the scientific and lay press. First, they often arise from exposure to asbestos fibres, and are thereby a public and occupational health concern. Second, they are a particularly deadly form of cancer. Third, diagnosis of mesothelioma is often difficult.

Malignant mesotheliomas are locally aggressive neoplasms in which the neoplastic proliferation originates from pleural, peritoneal or, rarely, pericardial mesothelial cells (Craighead and Mossman, 1982; Craighead, 1987; Antman, 1993). Mesotheliomas are linked epidemiologically to asbestos exposure and from that standpoint alone have been a major public health concern. Although mesotheliomas are often associated with extensive occupational asbestos exposure, a substantial 'bystander' risk has been documented, and as many as 50% of American patients have no history of known asbestos exposure (Craighead and Mossman, 1982). Advances in clinical recognition of mesothelioma have allowed the evaluation of additional risk factors. One known risk factor is inhalation of the airborne mineral dust erionite (Selcuk *et al.*, 1992), which is responsible for endemic mesothelioma in central Turkey. Another potential risk factor – albeit less convincingly implicated than asbestos – is exposure to SV40 virus. Accordingly, accurate diagnosis of mesothelioma is a global concern, and this concern will certainly persist after the present wave of asbestos-associated cases peaks and subsides.

Many other types of cancer can involve the pleura or peritoneum, but discussion of these entities, individually, is beyond the scope of this chapter. Such cancers are occasionally primary and can be multifocal and/or associated with familial cancer syndromes, as in the case of papillary serous carcinoma of the peritoneum (Schorge *et al.*, 1998). However, nonmesothelial pleural and peritoneal tumours are more often metastatic from another tissue site. Pleural or peritoneal metastases are generally manifested by shortness of breath or abdominal swelling/discomfort due to fluid accumulation, and they are sometimes the only evident metastases in carcinomas arising from organs such as lung, large bowel, ovary or breast. Patients can come to medical attention because of symptomatic pleural or peritoneal metastases at a time when the primary carcinoma is not yet evident. These scenarios emphasize the importance of distinguishing histologically and immunohistochemically between epithelial-type mesothelioma and carcinoma.

NORMAL DEVELOPMENT AND STRUCTURE

Mesothelial cells derive from the mesoderm, and they form a single cell layer lining the pleural, peritoneal and pericardial spaces. They allow relatively frictionless

movement of opposed surfaces, a function which is important for the organs they encompass. For example, they ease the motion of pleura and pericardial surfaces which is important during respiration and cardiac contraction/relaxation, respectively. Mesothelial cells also play vital roles in regulating transport of fluids and molecules across the pleural, peritoneal and pericardial spaces. The pleura and peritoneum proper include not only mesothelial cells but also fibrovascular connective tissue which provides support for the mesothelial cells. Cancers can arise from any of the cell types in the pleura and peritoneum, but this chapter will focus on those – malignant mesotheliomas – which originate from the mesothelial cells.

AETIOLOGY

Malignant mesothelioma was diagnosed infrequently prior to 1960, at which time Wagner *et al*. (1960) described a dramatically increased incidence in asbestos miners. Thereafter, mesothelioma was highly publicized owing, in part, to public health concerns. Subsequent epidemiological studies confirmed the striking association between asbestos exposure and mesothelioma. Mesothelioma was found to be common in groups as diverse as native Americans (occupational exposure during production of silver jewellery); carpenters, plumbers, and electricians (occupational exposure to asbestos insulation materials); and villagers of central Turkey (inhalation of the endogenous mineral erionite, a structural mimic of asbestos) (Craighead and Mossman, 1982; Craighead, 1987; Selcuk *et al*., 1992; Antman, 1993). The latency period between asbestos exposure and presentation with mesothelioma is typically 30–45 years. Mesothelioma incidence will probably increase over the next several decades because asbestos use rose in most countries after the Second World War and was not regulated until 1970. A registry of mesothelioma deaths has been maintained in England, Wales and Scotland since 1968, and annual deaths from mesothelioma in these populations rose from 154 in 1968 to 1009 in 1991 (Peto *et al*., 1995). Based on known occupational asbestos exposure in these countries, it is estimated that mesothelioma incidence will peak in year 2020 and will account for 1% of all deaths in the most highly exposed cohorts (men born during the 1940s) (Peto *et al*., 1995).

Asbestos belongs to a family of magnesium and calcium silicates that withstand extremely high temperatures. The thermoresistant properties of asbestos have been appreciated for many millenia, and were evident to the ancient Greeks, who used asbestos in wraps for cremation. In recent times, asbestos has found widespread use as insulation material in the walls of buildings and as heat-resistant material in engine parts and brake linings, among other applications. Asbestos fibres can be grouped into those which are wavy (serpentines) and straight (amphiboles), the latter being more often implicated in oncogenesis (Walker *et al*., 1992). As discussed below, it is likely that the oncogenic potential of asbestos is multifactorial.

Several *in vitro* studies provide fascinating clues to the pathogenesis of tumorigenic chromosomal rearrangements in mesothelioma. Addition of asbestos fibres to non-neoplastic mesothelial cell cultures results in pronounced chromosome damage that is mediated by physical interactions between asbestos fibres and chromosomes (Lechner *et al*., 1985; Ault *et al*., 1995). Asbestos fibres 'snare' whole chromosomes, or break off pieces of chromosomes, during mitosis, and these chromosomes and chromosome fragments are separated from the normal mitotic spindle apparatus (Ault *et al*., 1995). The end result is aberrant chromosome segregation with loss of a chromosome, or chromosome region, from one daughter cell. Therefore, these studies establish a scenario in which asbestos fibres are directly responsible for random chromosomal damage. Amongst the asbestos-exposed mesothelial cells are presumably those occasional cells in which the chromosomal damage results in loss of key tumour-supressor loci. These cells might represent the starting point for neoplastic mesothelial proliferation. *In vitro* models suggest that asbestos fibres also have a direct role in promoting mesothelial cell proliferation via nonmutational mechanisms (Pache *et al*., 1998; Timblin *et al*., 1998). Consequently, asbestos may inflict initial genotoxic damage, then promote proliferation of the damaged mesothelial cells, enabling them to acquire additional oncogenic mutations.

Considerable controversy attended recent reports that SV40 virus is found in many human mesotheliomas. The SV40 promoter is commonly used to drive expression of various genes which are introduced into mammalian cells as plasmid constructs for *in vitro* analyses. However, SV40 itself is generally nontransforming when expressed in human non-neoplastic cells. Nonetheless, there is substantial, although disputed, evidence that SV40 participates in the genesis of human mesotheliomas. Evaluation of human mesothelioma was prompted by the discovery, by Carbone *et al*., (1994) that Syrian hamsters developed mesothelioma following injection of wild-type SV40 virus into the pleural space. The same group noted that SV40-like sequences were demonstrable by polymerase chain reaction (PCR) in primary human mesotheliomas (Carbone *et al*. (1994). Other groups did not confirm this finding, or noted that the ability to identify SV40-like sequences depended on the specific SV40 oligonucleotide primers used for the PCR (Strickler *et al*., 1996). In addition, SV40 protein was not identified in human mesothelioma cell lines (Modi *et al*., 2000), and immunohistochemical studies of primary mesothelioma tumours revealed a weak, non-nuclear, expression pattern of putative SV40, which was

interpreted as inconsistent with a transforming role. These contrary findings raised the question of whether the apparent SV40 genomic sequences might represent trivial PCR contaminants, 'bystander' SV40 in non-neoplastic cells, or a closely related virus distinct from SV40. A follow-up, multi-instutional study again supported the presence of true SV40 DNA within primary human mesotheliomas, and recent studies show that SV40 synergizes with asbestos in transforming mesothelial cells *in vitro* (Bocchetta *et al*., 2000). Nonetheless, it is unclear whether the neoplastic cells in primary human mesotheliomas express SV40 protein. It is important that these questions be resolved inasmuch as the possibility of an SV40 role is a substantial public health concern. Millions of individuals, in the United States and in other countries, were potentially exposed to SV40 in poliovirus vaccines administered between 1955 and 1963.

SCREENING AND PREVENTION

Malignant mesothelioma is an asbestos-associated neoplasm that has been diagnosed with increasing frequency over the past two decades. Contemporary regulations on asbestos constitute one key measure in preventing mesotheliomas, but these measures are not expected to eradicate the disease. Many mesotheliomas develop in persons with no evident exposure to asbestos, and whose lung tissue does not contain elevated levels of asbestos fibres. Furthermore, several non-asbestos risk factors have been identified, including radiation therapy to the chest or abdomen, inhalation of certain environmental dusts and, potentially, exposure to the SV40 virus.

One strategy for mesothelioma prevention would involve screening at-risk individuals so that mesotheliomas could be detected at their earliest, *in situ*, stages and be cured surgically. Indeed, it would seem that the unique epidemiology of mesothelioma, particularly its relationship to asbestos exposure, should permit the development of effective screening measures. However, that has not been the case, and timely diagnosis remains a challenge. Only a small subset of individuals, presumably less than 5%, develop mesothelioma even after occupational exposures to large amounts of asbestos (Ribak *et al*., 1998). In addition, individuals exposed to asbestos develop benign pleural and peritoneal masses, known as 'plaques,' more often than they do malignant mesothelioma (Kannerstein *et al*., 1977). The challenges for screening are compounded because the latency period between asbestos exposure and clinical manifestations of disease is long, commonly more than 30 years, and conventional radiological screening methods, such as plain radiographs and CAT scans, do not distinguish early mesothelioma from benign asbestos-mediated pleural and peritoneal thickening or effusions. Further, patients often present with nonspecific symptoms, leading to delays in diagnosis, and mesothelioma progresses rapidly once clinically apparent. Given these challenges, it is perhaps more realistic to expect that improved survival in mesothelioma patients will result from better therapies rather than earlier detection. Most mesotheliomas are probably beyond the reach of surgical cure long before they first manifest clinically and radiologically. This is because mesothelioma cell shedding, and resultant pleural or peritoneal dissemination, occur early in the course of the disease.

CLINICAL PRESENTATION AND PATHOLOGY

Although the accuracy of mesothelioma diagnosis has improved dramatically over the past 20 years, many potential pitfalls remain. Most patients with malignant mesothelioma present with discomfort and/or respiratory distress due to mesothelioma bulk tumour, pleural effusion or ascites (Craighead, 1987; Antman, 1993). However, some patients present only with nonspecific symptoms, such as weight loss, fevers and fatigue. The most common radiographic sign is that of pleural effusion with or without pleural plaques, and cytological examination of effusion or ascites fluid is often the initial diagnostic procedure. Cytological diagnosis is challenging as the morphological similarities between malignant mesothelioma and reactive mesothelial hyperplasia may preclude differentiating between the two in fluid cytology. When overt cytological features of malignancy are present, distinction from adenocarcinoma (primary to the lung or from a distant site) may be impossible without employing immunocytochemical studies. Cytological diagnosis of malignancy was achieved in 22–66% of mesothelioma patients in several large series, whereas the remaining cases could not be distinguished from reactive mesothelial hyperplasia (Scherman and Mark, 1990; Boutin *et al*., 1993). Reactive mesothelial hyperplasia is a relatively nonspecific finding which can occur in response to almost any type of pleural or peritoneal insult. For example, pronounced mesothelial hyperplasia can be seen in association with pleural or peritoneal effusions secondary to infection, cirrhosis or rheumatoid arthritis. Although pleural or peritoneal thickening are demonstrated by computed tomography (CT) or magnetic resonance (MR) scanning in some patients with mesothelioma, these findings are also not conclusive for malignancy. Asbestos is well known to induce pleural or peritoneal plaques and/or diffuse thickening even in the absence of actual mesothelioma (Kannerstein *et al*., 1977).

Only 50% of mesotheliomas are associated with asbestos exposure. Based on this statistic, it is obvious that the recent surge in diagnosed mesothelioma cannot be due entirely to changing trends in asbestos use. A minor factor

is the growing number of cases occurring as second neoplasms after radiation therapy. A more significant variable, however, is the increasing sophistication of mesothelioma diagnosis. There is little question that many or most epithelial-type mesotheliomas were misdiagnosed as adenocarcinoma prior to 1970, whereas many or most spindle-cell mesotheliomas were misdiagnosed as sarcoma. Increased awareness of histological and immunohistochemical subtleties has allowed more accurate (and more frequent) diagnosis of mesothelioma over the past two decades. Accurate diagnosis of mesothelioma is of more than academic interest, because survival for patients with early-stage mesothelioma has improved with a combined approach of pleural or peritoneal stripping, multiagent chemotherapy and radiation to sites of bulky disease (Sugarbaker *et al*., 1999).

Histopathology

The gross appearance of mesothelioma often features numerous small nodules and plaques and/or generalized thickening of the pleura or peritoneum (**Figure 1; see colour plate section**) (Corson, 1997). As discussed above, there is invariably some degree of associated effusion. Invasion of the lungs proper (pulmonary parenchyma) is unusual in the early stages of pleural mesothelioma, whereas involvement of local lymph nodes is seen occasionally (Sugarbaker *et al*., 1993, 1999). Nonetheless, pulmonary function is often compromised by the combined restrictive effects of malignant pleural effusion and pleural encasement by the tumour (**Figure 1**). The histological appearance of malignant mesothelioma is variable. There are three major morphological presentations, which may be categorized into epithelial, biphasic and sarcomatoid types.

Most mesotheliomas are composed of round epithelial-like cells (**Figure 2; see colour plate section**), and such cases can be misdiagnosed as adenocarcinoma. Other mesotheliomas are composed of spindled cells (**Figure 3; see colour plate section**), and these tumours can be mistaken for various sarcomas, including monophasic synovial sarcoma, malignant peripheral nerve sheath tumour, solitary fibrous tumour and fibrosarcoma. Spindle-cell mesotheliomas can also be mistaken for epithelial cancers, which occasionally have a spindle-cell morphology. Furthermore, mesothelioma spindle cells can be innocent-looking in some cases, despite their invariably lethal behaviour, and can be mistaken for a non-neoplastic – or desmoplastic – spindle-cell proliferation. Immunohistochemistry and electron microscopy have been helpful in resolving these diagnostic considerations. A number of monoclonal antibodies facilitate differentiation of epithelial-type mesothelioma from adenocarcinoma, and other antibodies enable distinction of spindle-cell mesothelioma from sarcoma (Cagle *et al*., 1989; Brown *et al*., 1993; Weiss and Battifora, 1993; Ordonez, 1999). Cytogenetic and molecular analyses are also helpful in some cases. For example, synovial sarcoma, malignant peripheral nerve sheath tumour and solitary fibrous tumour can generally be distinguished based on their respective expression of *SYT–SSX* fusion oncogenes, S100 protein and CD34 and their lack of expression for the calretinin mesothelioma marker. On the other hand, neither immunohistochemistry nor electron microscopy has been helpful in distinguishing malignant mesothelioma from reactive mesothelial hyperplasia (Crotty *et al*., 1992).

Immunohistochemistry

Although all aspects of the clinicopathological evaluation must be considered in arriving at a diagnosis of mesothelioma, most would agree that immunohistochemistry is the key adjunct to routine histology. Indeed, the immunohistochemical evaluation of mesothelioma receives considerable attention in the pathology literature, and a full account of this dynamic field is beyond the realm of this chapter. Only representative highlights will be mentioned, such as the recent discovery that calretinin, a calcium-binding protein, is expressed in virtually all mesothelial cells (**Figures 2** and **4**). Calretinin is expressed in both normal and malignant mesothelial cells, and therefore does not serve to distinguish mesothelioma from reactive mesothelial hyperplasia. However, calretinin is seldom expressed in mesothelioma histological mimics (Doglioni *et al*., 1996), including adenocarcinomas metastatic to pleura or peritoneum, or spindle-cell sarcomas and sarcomatoid carcinomas (Doglioni *et al*., 1996; Attanoos *et al*., 2000). There are many other antibodies which discriminate mesothelioma from histological mimics, and which are suitable for paraffin section immunohistochemistry. None of these reagents provides perfect specificity or sensitivity, and they are therefore applied as panels in which complementary antibodies provide a more convincing final answer. Examples of antibody targets useful in distinguishing mesothelioma from adenocarcinoma include (1) Ber-EP4, which is expressed diffusely and strongly in most bronchogenic adenocarcinomas but absent or only focal in most mesotheliomas, (2) CEA, which is expressed in most adenocarcinomas but rarely in mesothelioma (**Figure 4; see colour plate section**), and (3) Leu-M1, which is also expressed generally in adenocarcinoma but not in mesothelioma. Similarly, TAG-72 (recognized by the antibody B72.3) and thyroid transcription factor-1 (TTF-1) are observed in a large percentage of pulmonary adenocarcinomas, but not in mesotheliomas.

It is important to emphasize that specificity and sensitivity of immunohistochemical detection is influenced by the epitope against which the antibody was raised and the method – particularly monoclonal versus polyclonal – of antibody preparation. Not surprisingly, there

are conflicting reports as to whether the above-mentioned antibodies are suitable for evaluation of mesothelioma. These discrepant results undoubtedly reflect both the use of different antibody preparations as well as the different techniques for immunohistochemical staining. CEA provides an illustrative example. It is well known that monoclonal antibodies to CEA vary in their sensitivity for detecting CEA in bronchogenic adenocarcinomas. Conversely, polyclonal antisera to CEA can recognize a nonspecific cross-reactive antigen which decreases their specificity. Preabsorption with mouse spleen powder is reported to decrease this nonspecific immunoreactivity but, nevertheless, a small percentage of malignant mesotheliomas are immunoreactive to antibodies directed against CEA. Further, immunohistochemical staining for CEA is less useful in evaluation of peritoneal mesothelioma, because a substantial number of ovarian carcinomas do not stain for this glycoprotein.

Keratin expression (**Figure 3**) is a near-universal finding in mesothelioma whereas it is uncommon and/or focal in some of the spindle-cell sarcomas which mimic sarcomatoid mesothelioma. Keratin staining is less useful in distinguishing epithelial-type mesothelioma from adenocarcinoma, although some observers report that the keratin staining pattern is diffuse and perinuclear in most epithelial-type mesotheliomas whereas it stains in a peripheral pattern, accentuating the cell membrane, in many adenocarcinomas (Corson, 1997).

Ultrastructure

Electron microscopy evaluation is less critical than immunohistochemistry in the diagnostic work-up of malignant mesothelioma. Nonetheless, there are characteristic ultrastructural differences – particularly between epithelial-type mesothelioma and adenocarcinoma – which can be useful in arriving at a diagnosis. Notably, epithelial-type mesotheliomas, and the epithelial-type components in mixed histology mesotheliomas, may often be distinguished from adenocarcinomas based on the size and number of microvilli which project from their cell surfaces (**Figure 5**). Epithelial-type mesotheliomas generally have numerous microvilli which are long, thin and branching, whereas adenocarcinomas typically have microvilli which are short and sparse (Corson, 1997).

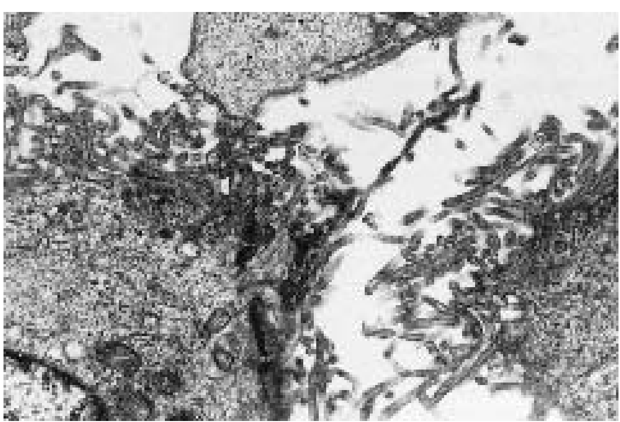


Figure 5 An ultrastructural hallmark of mesothelial cell differentiation is the finding of long branching microvilli at the cell surface (commonly referred to as bushy microvilli). In contrast, pulmonary and many extrapulmonary adenocarcinomas display short microvilli. In addition, perinuclear condensation of keratin intermediate filaments (tonofilaments) may be seen in mesothelial cells. Although ultrastructural examination remains useful in the diagnosis of mesothelioma, it is generally more cumbersome than immunohistochemical evaluation.

CYTOGENETIC AND MOLECULAR GENETIC FINDINGS

There is experimental evidence that asbestos fibres are directly responsible for chromosomal damage in mesothelial cells. Also, there is evidence that the SV40 virus serves as an oncogenic cofactor in promoting the development of some mesotheliomas. Mesotheliomas differ from many other solid tumours, however, in that the early, preinvasive, stages of the disease are rarely observed (Henderson *et al*., 1998). In addition, it can be very difficult to distinguish *in situ* mesothelioma from non-neoplastic mesothelial hyperplasia. Therefore, the *in situ* stage of mesothelioma has not been readily available for comparison with invasive mesothelioma, and the molecular mechanisms responsible for mesothelioma initiation remain to be defined. On the other hand, various studies have shed light on cytogenetic, molecular and cell biology mechanisms in invasive mesotheliomas. Several of these mechanisms are discussed below.

Cytogenetic Alterations

Cytogenetic studies have revealed a characteristic profile of chromosomal deletions in most mesotheliomas. These cytogenetic findings implicate loss of several tumour-suppressor genes as a critical event in mesothelioma pathogenesis. Although mesothelioma karyotypes are often described as 'complex,' they are in fact less complex than those in many carcinomas. In particular, epithelial-type mesothelioma karyotypes often have fewer than five chromosomal abnormalities per cell (**Figure 6**), whereas bronchogenic adenocarcinomas typically have more than 20 chromosomal abnormalities per cell. Sarcomatoid mesotheliomas, on the other hand, have more complex karyotypes than do the epithelial-type cases. A rather detailed account of mesothelioma chromosomal aberrations has emerged over the past 10 years (Taguchi *et al*., 1993), and it is clear that the nonrandom chromosomal

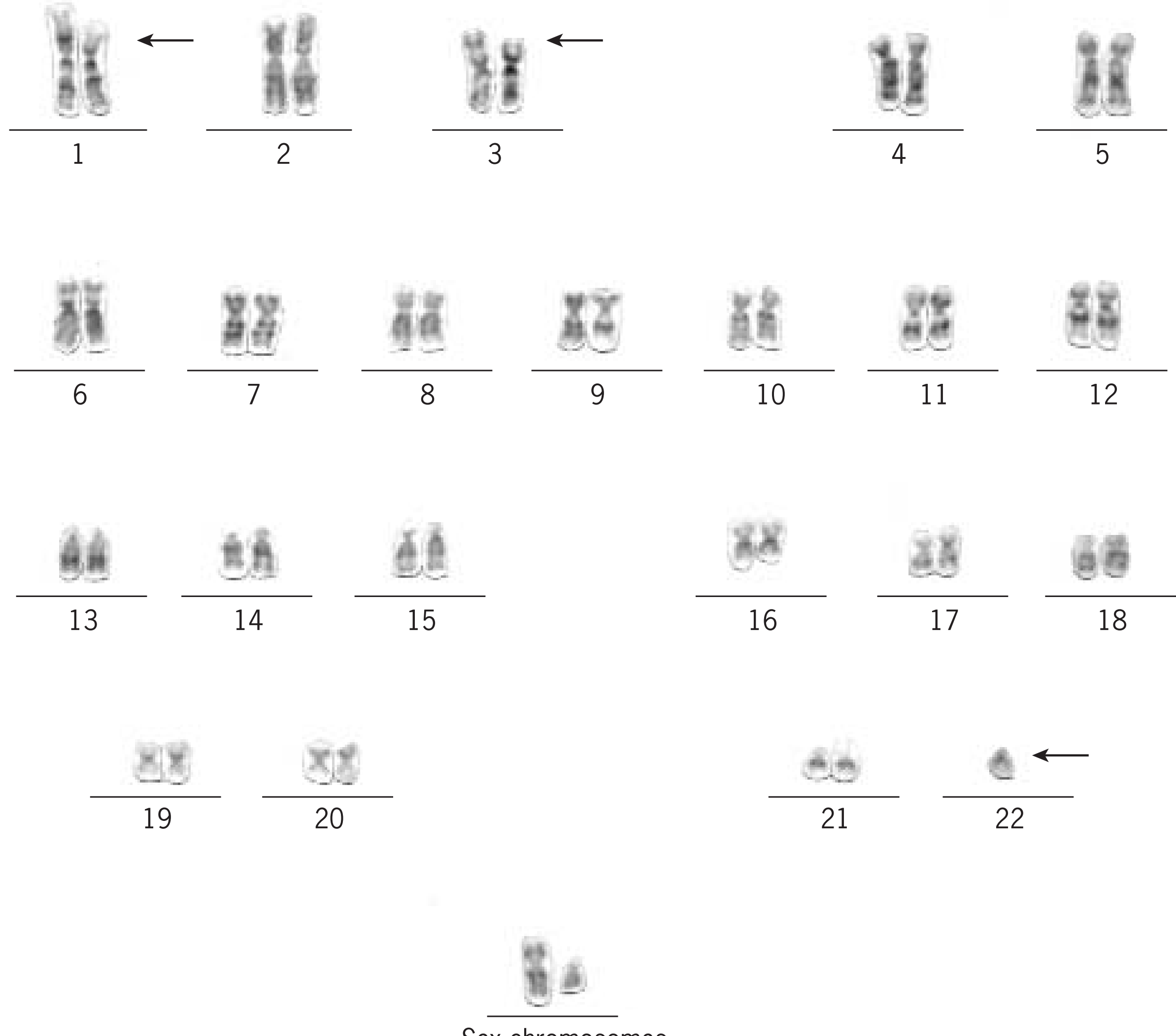


Figure 6 Noncomplex karyotype in an epithelial-type mesothelioma, showing characteristic losses of 1p, 3p and chromosome 22 (arrows).

deletions are found in mesotheliomas irrespective of histology (epithelial-type versus spindle-cell) or primary site (pleural versus peritoneal). The deletional hotspots are within the long arms of chromosomes 6 and 22 and within the short arms of chromosomes 1, 3 and 9 (**Figure 6**). Notably, the pathogenesis of some mesothelioma chromosome deletions may be directly attributable to asbestos-mediated chromosome damage (Lechner *et al.*, 1985; Ault *et al.*, 1995).

Molecular Alterations

Cytogenetic studies implicate loss of several tumour genes in the pathogenesis of virtually all mesotheliomas. However, the molecular targets for most of these chromosomal deletions are not yet known. Exceptions are the *CDKN2A* and *CDKN2B* genes, which are targetted by chromosome 9p deletions, and the *NF2* gene, which is targetted by chromosome 22q deletions (Cheng *et al.*, 1994). The *CDKN2A* and *CDKN2B* genes encode proteins which inhibit the CDK4 cell cycle checkpoint kinase and which thereby serve to block cell cycling. *CDKN2A* is targetted by chromosomal deletions in many human tumours, and *CDKN2A* genomic loss or functional inactivation, resulting in dysregulated cell cycling, is one of the more common tumour-suppressor mechanisms in human cancers. It is particularly interesting that *CDKN2A* mutations are often either/or in relationship to mutations of several 'downstream' effectors, including the CDK4 kinase and the retinoblastoma protein (Rb) (He *et al.*, 1995). CDKN2A inhibits the function of CDK4–cyclin D1 complexes, which, in turn, inactivate the Rb protein via phosphorylation. Hypophosphorylated Rb blocks cell entry into the S phase of the cell cycle, whereas phosphorylated Rb fails to block entry to S phase and thereby promotes cell proliferation. Hence CDKN2A, CDK and Rb

are control points along a single biological pathway, and disruption of any one of these points can disable normal constraints on cell cycling. These either/or relationships also appear to hold in mesothelioma, where CDKN2A inactivation is nearly universal, but Rb or CDK4 mutations are very uncommon (Cheng *et al*., 1994; Shimizu *et al*., 1994).

The NF2 protein is inactivated in virtually all mesotheliomas with chromosome 22 deletions (Sekido *et al*., 1995), and these represent approximately 75% of all mesotheliomas. NF2 was identified originally through analyses of its tumour-suppressor role in the genetic cancer syndrome, neurofibromatosis type 2, and through study of the various tumours, particularly benign schwannomas, which are characteristic of that syndrome. The normal functions of NF2 are poorly understood but it belongs to a family of proteins, each of which contains a so-called 'ezrin–radixin–moesin' domain, which communicate signals between cell surface receptors and the cytoskeleton. Recent studies show that NF2 interacts with substrates of the Met receptor tyrosine kinase and with scaffolding proteins interacting with the platelet-derived growth factor receptor tyrosine kinase. These findings suggest that NF2 modulates signal transduction pathways downstream of Met, PDGFR and other cell surface receptors. Such findings are relevant clinically because inhibition of receptor tyrosine kinases, whether by immunological or small molecule methods, has proved very effective in treatment of some cancers.

p53 is a seminal tumour-suppressor protein, whose inactivation contributes to the development of many human cancers. Initial reports suggested that p53 mutation was commonplace in mesotheliomas. This impression was not confirmed in other studies, and it is unlikely that p53 mutation plays a major role in mesothelioma pathogenesis (Mor *et al*., 1997). Nonetheless, it is possible that the p53 pathway is inhibited by other oncogenic mutations, and just such a role has been proposed for SV40 oncoproteins in some mesotheliomas.

Diagnostic Relevance

Further characterization of mesothelioma chromosomal and molecular aberrations may lead to improvements in diagnosis of mesothelioma (Granados *et al*., 1994). For example, virtually all mesotheliomas contain clonal chromosomal deletions whereas such alterations are not found in reactive mesothelial hyperplasia. Cytogenetic and molecular assays can be useful diagnostically because neoplastic and reactive mesothelial proliferations cannot be distinguished consistently using histological adjuncts such as immunohistochemistry and electron microscopy. There is also evidence that molecular assays are useful in distinguishing mesothelioma from adenocarcinoma (Sekido *et al*., 1995). Certain of the cytogenetic methods can now be performed routinely in cytological preparations (**Figure 7; see colour plate section**), and therefore should be useful in enabling a diagnosis of mesothelioma in pleural or peritoneal fluid specimens.

PROGNOSTIC FACTORS

Overall survival for patients with malignant mesothelioma is unquestionably poor. Nevertheless, intensive combined modality therapeutic strategies, as summarized below, have enabled a subgroup of patients to achieve 5-year survival. This accomplishment is striking, given that very few mesothelioma patients survive even 1 year when less intensive treatment methods are used. In the face of these therapeutic advances, it becomes more pressing to identify reliable prognostic factors. In particular, it is important to identify the subgroup of patients who will benefit from major surgical procedures, such as extrapleural pleuropneumonectomy.

Various potential prognostic variables have been evaluated in different mesothelioma series. Such studies have considered the prognostic implications of presenting symptoms (e.g. dyspnea, chest pain, cough), performance status, age, sex, histology, stage of disease and asbestos exposure. Tumour histology has been a major prognostic factor in all studies. Survival for patients who are untreated, or who receive single-modality therapy, is typically in the range of 4, 7 and 12 months depending on whether histology is sarcomatoid, mixed or epithelial-type, respectively (Fusco *et al*., 1993). The prognostic role of histology is also evident in patients receiving intensive treatment approaches, including the trimodality approach of extrapleural pleuropneumonectomy, chemotherapy and radiation therapy. In one such study, patients with epithelial-type mesothelioma had median survivals of 52 and 21% at 2 and 5 years, respectively, whereas only 16% of those with sarcomatoid mesothelioma survived for 2 years and none survived for 5 years (Sugarbaker *et al*., 1999). Primary site has also influenced prognosis in virtually all studies. Patients with peritoneal mesothelioma – with the exception of cases that are very well differentiated – have a median survival which is less than half of that in patients with pleural mesothelioma (Sridhar *et al*., 1992).

Although most mesotheliomas are disseminated locally at time of diagnosis, disease stage is nonetheless important prognostically (De Pangher Manzini *et al*., 1993). Not surprisingly, mediastinal lymph node involvement is an adverse prognostic factor in patients undergoing cytoreductive surgical approaches, such as extrapleural pleuropneumonectomy (Sugarbaker *et al*., 1993). Other prognosticators are undoubtedly influenced by the treatment approach. These include age less than 65 years, female gender, absence of pain, longer duration of symptoms, excellent performance status and absence of malignant cells in the associated effusion, all of which have been favourable prognostic factors in some studies but not in others (Fusco *et al*., 1993; Sugarbaker *et al*., 1993).

OVERVIEW OF PRESENT CLINICAL MANAGEMENT

Mesothelioma is relatively resistant to conventional chemotherapy and radiation therapy approaches. There are numerous reports, generally in small numbers of patients, which show impressive response rates – often in the range 25–50% for partial response – after single- or multiagent chemotherapy. Invariably, however, such findings have been refuted in follow-up studies, where the partial response rates prove to be well under 20%. Nonetheless, there are several chemotherapeutic agents, including doxorubicin and cisplatin, which unquestionably induce occasional partial responses, albeit typically lasting for no more than a few months (Taub and Antman, 1997). Multiagent chemotherapy protocols, e.g. cisplatin + doxorubicin, may achieve partial response rates in 20–30% of patients (Chahinian *et al*., 1993). Radiation therapy as a single modality is relatively ineffective, with very few partial responders, although some patients receive temporary benefit in the form of symptomatic improvement (Gordon *et al*., 1982). Likewise, mesothelioma is rarely eradicated by surgery alone, although meaningful palliation can be achieved (Butchart *et al*., 1981). Even the more radical surgical approaches invariably leave tumour cells behind, and are therefore cytoreductive rather than curative.

Although single modality therapies – whether with chemotherapy, radiation therapy or surgery – have been relatively ineffective, the combination of these methods has been highly successful in certain subgroups of patients. In particular, cytoreduction by extrapleural pleuropneumonectomy, when followed by multiagent chemotherapy and radiation therapy, leads to 5-year survival in many patients with epithelial-type mesothelioma (Boutin *et al*., 1993). Unfortunately, only one-third of pleural mesothelioma patients are candidates for this radical surgical approach, which involves *en bloc* removal of the lung, pleura, pericardium and diaphragm from the involved side of the chest (Boutin *et al*., 1993). In addition, combined modality methods have been less successful in patients with sarcomatoid mesothelioma, very few of whom survive for 2 years. Another constraint is that postoperative mortality is high, typically in excess of 10%, in centres which do not specialize in the radical surgical methods needed for effective cytoreduction (Rusch, 1999). Survival after the combined approach of extrapleural pleuropneumonectomy, chemotherapy and radiation therapy has been particularly prolonged in patients with epithelial-type histology whose resection margins and mediastinal lymph nodes were free of apparent tumour. This subgroup, represented by 31 of 176 extrapleural pleuropneumonectomy patients in one study, had 2- and 5-year survivals of 68 and 46%, respectively, whereas those for the overall group were 38 and 15%, respectively (Boutin *et al*., 1993). Similar trimodality approaches have also proved effective in pilot studies of peritoneal mesothelioma, where cytoreduction is accomplished by surgical debulking and then followed with chemotherapy and local irradiation to destroy residual tumour. Such studies often use intraperitoneal chemotherapy, which can be delivered readily using a Port-A-Cath set up. In one such study, with an intraperitoneal chemotherapy component of doxorubicin and cisplatin, three of five patients were disease-free at intervals of 46, 60 and 61 months after diagnosis (Lederman *et al*., 1988).

Despite the impressive advances provided by combined modality treatment, including a small but growing number of 10-year survivors, most mesothelioma patients do not survive past 2 years. This in part reflects the reality that many patients are fairly sick and/or inoperable at the time of diagnosis. Delays in diagnosis are common, particularly in those patients presenting with nonspecific symptoms, and many patients lose weight rapidly and develop significant respiratory compromise before therapy can be initiated. Therefore, improved survival for the overall group of mesothelioma patients probably depend on prompt diagnosis and development of more effective pharmacological interventions. Several novel therapies provide basis for optimism, although it is as yet unclear whether they will serve as effective solutions in their own right or as stepping stones in the quest for curative therapies. Examples include the intrapleural administration of immunomodulatory agents, such as interleukin-2 and γ-interferon, which have induced partial responses in up to 50% of patients with early-stage mesothelioma (Astoul *et al*., 1998). Beyond these tangible advances, it is certain that the growing understanding of mesothelioma biology, including oncogenes, tumour-suppressor genes and key growth factor signalling pathways, will translate in the near future to novel and biologically rational therapies.

REFERENCES

Antman, K. H. (1993). Natural history and epidemiology of malignant mesothelioma. *Chest*, **103**, 373S–376S.

Astoul, P., *et al*. (1998). Intrapleural administration of interleukin-2 for the treatment of patients with malignant pleural mesothelioma: a Phase II study. *Cancer*, **83**, 2099–2104.

Attanoos, R. L., *et al*. (2000). Anti-mesothelial markers in sarcomatoid mesothelioma and other spindle cell neoplasms. *Histopathology*, **37**, 224–231.

Ault, J. G., *et al*. (1995). Behavior of crocidolite asbestos during mitosis in living vertebrate lung epithelial cells. *Cancer Research*, **55**, 792–798.

Bocchetta, M., *et al*. (2000). From the cover: human mesothelial cells are unusually susceptible to simian virus 40-mediated transformation and asbestos cocarcinogenicity. *Proceedings*

of the National Academy of Sciences of the USA, **97**, 10214–10219.

Boutin, C., *et al*. (1993). Thoracoscopy in pleural malignant mesothelioma: a prospective study of 188 consecutive patients. Part 2: Prognosis and staging. *Cancer*, **72**, 394–404.

Brown, R. W., *et al*. (1993). Multiple-marker immunohistochemical phenotypes distinguishing malignant pleural mesothelioma from pulmonary adenocarcinoma. *Human Pathology*, **24**, 347–354.

Butchart, E. G., *et al*. (1981). The role of surgery in diffuse malignant mesothelioma of the pleura. *Seminars in Oncology*, **8**, 321–328.

Cagle, P. T., *et al*. (1989). Immunohistochemical differentiation of sarcomatoid mesotheliomas from other spindle cell neoplasms. *American Journal of Clinical Pathology*, **92**, 566–571.

Carbone, M., *et al*. (1994). Simian virus 40-like DNA sequences in human pleural mesothelioma. *Oncogene*, **9**, 1781–1790.

Chahinian, A. P., *et al*. (1993). Randomized phase II trial of cisplatin with mitomycin or doxorubicin for malignant mesothelioma by the Cancer and Leukemia Group B. *Journal of Clinical Oncology*, **11**, 1559–1565.

Cheng, J. Q., *et al*. (1994). p16 alterations and deletion mapping of 9p21–p22 in malignant mesothelioma. *Cancer Research*, **54**, 5547–5551.

Corson, J. M. (1997). Pathology of diffuse malignant pleural mesothelioma. *Seminars in Thoracic and Cardiovascular Surgery*, **9**, 347–355.

Craighead, J. E. (1987). Current pathogenetic concepts of diffuse malignant mesothelioma. *Human Pathology*, **18**, 544–557.

Craighead, J. E. and Mossman, B. T. (1982). The pathogenesis of asbestos-associated diseases. *New England Journal of Medicine*, **306**, 1446–1455.

Crotty, T. B., *et al*. (1992). Desmoplastic malignant mesothelioma masquerading as sclerosing mediastinitis: a diagnostic dilemma. *Human Pathology*, **23**, 79–82.

De Pangher Manzini, V., *et al*. (1993). Prognostic factors of malignant mesothelioma of the pleura. *Cancer*, **72**, 410–417.

Doglioni, C., *et al*. (1996). Calretinin: a novel immunocytochemical marker for mesothelioma. *American Journal of Surgical Pathology*, **20**, 1037–1046.

Fusco, V., *et al*. (1993). Malignant pleural mesothelioma. Multivariate analysis of prognostic factors on 113 patients. *Anticancer Research*, **13**, 683–689.

Gordon, W., *et al*. (1982). Radiation therapy in the management of patients with mesothelioma. *International Journal of Radiation Oncology and Biological Physics*, **8**, 19–25.

Granados, R., *et al*. (1994). Cytogenetic analysis of effusions from malignant mesothelioma. A diagnostic adjunct to cytology. *Acta Cytologia*, **38**, 711–717.

He, J., *et al*. (1995). Lack of p16INK4 or retinoblastoma protein (pRb), or amplification-associated overexpression of cdk4 is observed in distinct subsets of malignant glial tumors and cell lines. *Cancer Research*, **55**, 4833–4836.

Henderson, D. W., *et al*. (1998). Reactive mesothelial hyperplasia vs mesothelioma, including mesothelioma in situ: a brief review. *American Journal of Clinical Pathology*, **110**, 397–404.

Kannerstein, M., *et al*. (1977). Pathogenic effects of asbestos. *Archives of Pathology and Laboratory Medicine*, **101**, 623–628.

Lechner, J. F., *et al*. (1985) Asbestos-associated chromosomal changes in human mesothelial cells. *Proceedings of the National Academy of Sciences of the USA*, **82**, 3884–3888.

Lederman, G. S., *et al*. (1988). Combined modality treatment of peritoneal mesotheliomas. *National Cancer Institute Monographs*, 321–322.

Modi, S., *et al*. (2000). Protein expression of the RB-related gene family and SV40 large T antigen in mesothelioma and lung cancer. *Oncogene*, **19**, 4632–4639.

Mor, O., *et al*. (1997). Absence of p53 mutations in malignant mesotheliomas. *American Journal of Respiratory Cellular and Molecular Biology*, **16**, 9–13.

Ordonez, N. G. (1999). Role of immunohistochemistry in differentiating epithelial mesothelioma from adenocarcinoma. Review and update. *American Journal of Clinical Pathology*, **112**, 75–89.

Pache, J. C., *et al*. (1998). Increased epidermal growth factor-receptor protein in a human mesothelial cell line in response to long asbestos fibers. *American Journal of Pathology*, **152**, 333–340.

Peto, J., *et al*. (1995). Continuing increase in mesothelioma mortality in Britain. *Lancet*, **345**, 535–539.

Ribak, J., *et al*. (1988). Malignant mesothelioma in a cohort of asbestos insulation workers: clinical presentation, diagnosis, and causes of death. *British Journal of Industrial Medicine*, **45**, 182–187.

Rusch, V. W. (1999). Indications for pneumonectomy. Extrapleural pneumonectomy. *Chest Surgery Clinics of North America*, **9**, 327–338.

Schorge, J. O., *et al*. (1998). Molecular evidence for multifocal papillary serous carcinoma of the peritoneum in patients with germline BRCA1 mutations. *Journal of the National Cancer Institute*, **90**, 841–845.

Sekido, Y., *et al*. (1995). Neurofibromatosis type 2 (NF2) gene is somatically mutated in mesothelioma but not in lung cancer. *Cancer Research*, **55**, 1227–1231.

Selcuk, Z. T., *et al*. (1992). Malignant pleural mesothelioma due to environmental mineral fiber exposure in Turkey. Analysis of 135 cases. *Chest*, **102**, 790–796.

Sherman, M. E. and Mark, E. J. (1990). Effusion cytology in the diagnosis of malignant epithelioid and biphasic pleural mesothelioma. *Archives of Pathology and Laboratory Medicine*, **114**, 845–851.

Shimizu, E., *et al*. (1994). RB protein status and clinical correlation from 171 cell lines representing lung cancer, extrapulmonary small cell carcinoma, and mesothelioma. *Oncogene*, **9**, 2441–2448.

Sridhar, K. S., *et al*. (1992). New strategies are needed in diffuse malignant mesothelioma. *Cancer*, **70**, 2969–2979.

Strickler, H. D., *et al*. (1996). Simian virus 40 and pleural mesothelioma in humans. *Cancer Epidemiology Biomarkers and Prevention*, **5**, 473–475.

Sugarbaker, D. J., *et al*. (1993). Node status has prognostic significance in the multimodality therapy of diffuse, malignant mesothelioma. *Journal of Clinical Oncology*, **11**, 1172–1178.

Sugarbaker, D. J., *et al*. (1999). Resection margins, extrapleural nodal status, and cell type determine postoperative long-term survival in trimodality therapy of malignant pleural mesothelioma: results in 183 patients. *Journal of Thoracic and Cardiovascular Surgery*, **117**, 54–63.

Taguchi, T., *et al*. (1993). Recurrent deletions of specific chromosomal sites in 1p, 3p, 6q, and 9p in human malignant mesothelioma. *Cancer Research*, **53**, 4349–4355.

Taub, R. N. and Antman, K. H. (1997). Chemotherapy for malignant mesothelioma. *Seminars in Thoracic and Cardiovascular Surgery*, **9**, 361–366.

Timblin, C. R., *et al*. (1998). Patterns of c-fos and c-jun protooncogene expression, apoptosis, and proliferation in rat pleural mesothelial cells exposed to erionite or asbestos fibers. *Toxicology and Applied Pharmacology*, **151**, 88–97.

Wagner, J. C., *et al*. (1960). Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape Province. *British Journal of Industrial Medicine*, **17**, 26.

Walker, C., *et al*. (1992). Possible cellular and molecular mechanisms for asbestos carcinogenicity. *American Journal of Industrial Medicine*, **21**, 253–273.

Weiss, L. M. and Battifora, H. (1993). The search for the optimal immunohistochemical panel for the diagnosis of malignant mesothelioma. *Human Pathology*, **24**, 345–346.

FURTHER READING

Antman, K. H. (1993). Natural history and epidemiology of malignant mesothelioma. *Chest*, **103**, 373S–376S.

Ault, J. G., *et al*. (1995). Behavior of crocidolite asbestos during mitosis in living vertebrate lung epithelial cells. *Cancer Research*, **55**, 792–798.

Bocchetta, M., *et al*. (2000). From the cover: human mesothelial cells are unusually susceptible to simian virus 40-mediated transformation and asbestos cocarcinogenicity. *Proceedings of the National Academy of Sciences of the USA*, **97**, 10214–10219.

Corson, J. M. (1997). Pathology of diffuse malignant pleural mesothelioma. *Seminars in Thoracic and Cardiovascular Surgery*, **9**, 347–355.

Craighead, J. E. (1987). Current pathogenetic concepts of diffuse malignant mesothelioma. *Human Pathology*, **18**, 544–557.

Modi, S., *et al*. (2000). Protein expression of the RB-related gene family and SV40 large T antigen in mesothelioma and lung cancer. *Oncogene*, **19**, 4632–4639.

Ordonez, N. G. (1999). Role of immunohistochemistry in differentiating epithelial mesothelioma from adenocarcinoma. Review and update. *American Journal of Clinical Pathology*, **112**, 75–89.

Pache, J. C., *et al*. (1998). Increased epidermal growth factor-receptor protein in a human mesothelial cell line in response to long asbestos fibers. *American Journal of Pathology*, **152**, 333–340.

Sekido, Y., *et al*. (1995). Neurofibromatosis type 2 (NF2) gene is somatically mutated in mesothelioma but not in lung cancer. *Cancer Research*, **55**, 1227–1231.

Sugarbaker, D.J., *et al*. (1999). Resection margins, extrapleural nodal status, and cell type determine postoperative long-term survival in trimodality therapy of malignant pleural mesothelioma: results in 183 patients. *Journal of Thoracic and Cardiovascular Surgery*, **117**, 54–63.

Taguchi, T., *et al*. (1993). Recurrent deletions of specific chromosomal sites in 1p, 3p, 6q, and 9p in human malignant mesothelioma. *Cancer Research*, **53**, 4349–4355.

Taub, R. N. and Antman, K. H. (1997). Chemotherapy for malignant mesothelioma. *Seminars in Thoracic and Cardiovascular Surgery*, **9**, 361–366.

Chapter 50

Heart

Renu Virmani and Allen Burke
Armed Forces Institute of Pathology, Washington, DC, USA

CONTENTS

- Normal Development and Structure
- Tumour Pathology (Clinical, Gross and Histological Features)
- Epidemiology
- Aetiology
- Screening and Prevention
- Ultrastructure/and Immunohistochemistry
- Molecular Genetic Findings
- Prognostic Factors
- Overview of Present Clinical Management

NORMAL DEVELOPMENT AND STRUCTURE

The heart forms as a tube which loops between 24 and 26 days postovulation. Ventricular and atrial septation occur by 39 and 43 days postovulation, respectively; at the latter time, the heart structure is generally complete. There is little information relating cardiac development and neoplasia of the heart, as cardiac embryology is studied primarily in the context of congenital cardiac malformations (Burke and Virmani, 1996). Cardiac masses that develop *in utero* include rare lesions such as rhabdomyomas (frequently seen in conjunction with extracardiac hamartomas as part of the tuberous sclerosis syndrome), cardiac fibromas, benign rests of endodermal tissues (primarily benign AV nodal tumours) and potentially malignant tumours derived from developmentally misplaced germ cells (cardiac teratomas and related lesions). Rhabdomyomas and fibromas are generally considered to be hamartomas, that is, they are benign growths, usually developmental, that are composed of cells native to the organ of origin. An AV nodal tumour is not strictly a tumour, in that it represents a cluster of misplaced cells that migrated incorrectly within the foetus. These three lesions (rhabdomyoma, fibroma and AV nodal tumour) are not, then, strictly tumours, and result in potentially lethal symptoms only because of their delicate location, often near the conduction system of the heart, and not in their capability to grow and metastasize, which is absent. Germ cell tumours of the heart and pericardial space occur as a neoplastic growth of germ cell rests that remained in the mediastinum due to arrested migration from the yolk sac to the gonads. Mediastinal germ cell tumours in themselves are fairly rare; nevertheless, they outnumber intrapericardial germ cell tumours by a large margin.

The cell of origin of the commoner heart tumours in adults, namely myxomas and sarcomas, is controversial, as is any potential relationship between these two entities. Cardiac myxoma arises almost exclusively in the endocardium of the atrial septum. This rare neoplasm is made of undifferentiated mesenchymal cells (derived from connective tissue) that elaborate a matrix rich in proteoglycans and inflammatory cells which release various growth factors and thrombogenic substances. The embryological origin of cells that become cardiac myxoma is unknown, but because of the location and myxoid appearance, it has been suggested that myxomas are derived from cells that make up the endocardial cushion. Cardiac sarcomas are a heterogeneous groups of tumours with presumably various cells of origin. The most common group appear to arise from the endocardial lining of the atria, often elaborate myxoid ground substance similar to cardiac myxomas, and generally demonstrate myofibroblastic differentiation. There is no evidence that they derive from the same cell as myxomas, however, in that they do not arise in benign myxomas, tend to originate in sites in the atrium away from the septum and are histologically similar to sarcomas that arise in the intima of great arteries (a site in which myxomas do not arise). The second most common group of cardiac sarcomas are angiosarcomas, which are histologically identical with angiosarcomas that arise in extracardiac sites, and arise from endothelial cells usually in the right atrium or pericardium. Despite the fact that the heart is comprised mostly of muscle, only a small fraction of cardiac sarcomas show characteristics of striated muscle (i.e. rhabdomyosarcomatous differentiation). Because

working heart muscle cells are terminally differentiated and show little propensity to divide, it is perhaps not a surprise that neoplasms of the heart are more likely to derive from less committed stromal cells that line the endocardium.

TUMOUR PATHOLOGY (CLINICAL, GROSS AND HISTOLOGICAL FEATURES)

For the purposes of this review, only primary cardiac neoplasms will be addressed (**Table 1**). Neoplasms arising in the heart include cardiac myxoma (about 75%) and cardiac sarcomas (less than 25%); other neoplasms, such as paragangliomas and teratomas, are extremely rare and will not be discussed in detail. Cardiac sarcomas are heterogeneous, but will be grouped into endocardial-based sarcomas, or myofibroblastic intimal sarcomas, which usually arise in the left atrium; angiosarcomas, which generally arise in the right atrium and pericardium; and miscellaneous sarcomas.

Two additional malignancies that arise in or near the heart need brief mention here. Malignant mesotheliomas of the pericardium, or the serosal lining of the heart, are rare neoplasms that share many clinical and pathological features with the much more common mesothelioma of the pleura. Pericardial mesotheliomas are often extensions of pleural mesotheliomas, and share the same pathological and epidemiological features as pleural mesotheliomas. The second rare malignancy of the heart that needs mention is the malignant lymphoma. Lymphomas generally arise from haematopoietic tissues that reside in lymph nodes, the spleen or the bone marrow, but may occasionally arise in parts of the body that do not normally possess large numbers of lymphoid cells. These locations include the soft tissue and internal organs, including brain, liver, and heart. The incidence of extranodal lymphomas has recently increased, because of the large numbers of people with immune deficiency syndromes, either due to infection with human immunodeficiency virus or because of iatrogenic immunosuppression in patients with allografts. Malignant lymphomas and malignant mesotheliomas will be discussed briefly in this chapter, but the reader is directed to other portions of this volume that address these neoplasms in greater detail. (See the chapters *Lymph Nodes* and *Pleura and Peritoneum*.)

Myxoma

Cardiac myxoma is a benign endocardial-based neoplasm that usually arises in the left atrium (70%) at the region of the oval fossa (**Figure 1**). About 20% of myxomas arise in the right atrium, also usually in the region of the atrial septum or, less commonly, in ventricular sites. Myxomas that originate on the valve surfaces are exceptionally rare, and many reported cases actually represent a different lesion, usually a papillary fibroelastoma, not a true neoplasm.

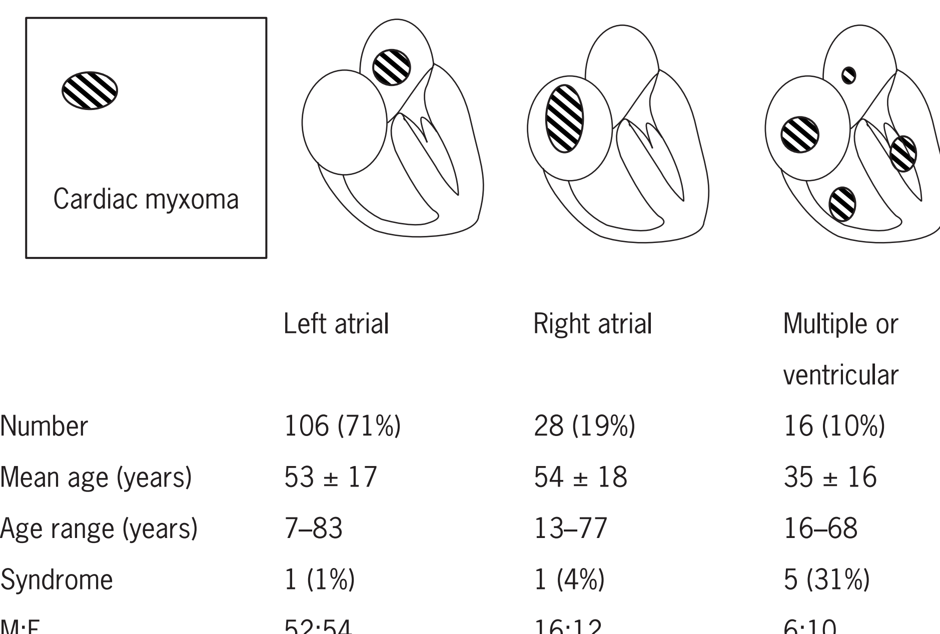
The mean age at presentation is 50 years in cases of sporadic myxoma (Burke and Virmani, 1996). Cardiac myxoma arises from the endocardial surface by a stalk, which may be either broad-based or pedunculated. There are two gross appearances which characterize the surface of cardiac myxoma. Approximately 70% of tumours have an intact, smooth, endothelialized surface, with no surface thrombus; these lesions appear to be quiescent, with little possibility for further growth. Such tumours cause symptoms generally by chronic obstruction of the mitral valve, often resulting in pulmonary hypertension and symptoms of congestive heart failure (**Figure 2**). Occasional tumours cause symptoms as a result of hormone-like substance (cytokines) that stimulate the immune system, causing fever, weight loss and other constitutional symptoms. Because cardiac imaging is becoming more sophisticated and commonplace, an increasing percentage of cardiac myxomas are being detected as incidental findings in asymptomatic patients. Although a routine chest X-ray is unlikely to detect a cardiac myxoma in the majority of cases, about 5% of left-sided myxomas and up to 50% of right-sided tumours, will be detected on the basis of tumour calcification. More non-specific findings, such as cardiomegaly or evidence of mitral valve obstruction, are common in left-sided myxomas.

The majority of cardiac myxomas have a smooth surface (**Figure 3; see colour plate section**). A small

Table 1 Classification of primary cardiac neoplasms[a]

Tumour type	Site in heart	Biological behaviour	Associations
Myxoma	Left atrium, right atrium	Benign	Myxoma syndrome (2%)
Sarcoma		Highly malignant	None
Myofibroblastic	Atria (left > right)		
Angiosarcoma	Right atrium, pericardium		
Other	Various		
Lymphoma	Atria, epicardium, ventricles	Malignant	Immunosuppression (50%)
Mesothelioma	Pericardium	Highly malignant	Asbestos (20–30%)

[a]Lesions that are variably considered hamartomas, such as rhabdomyoma, fibroma, haemangioma and papillary fibroelastoma, are not included. Likewise, primary neoplasms that are extremely rare, such as paraganglioma, are also not included.



	Left atrial	Right atrial	Multiple or ventricular
Number	106 (71%)	28 (19%)	16 (10%)
Mean age (years)	53 ± 17	54 ± 18	35 ± 16
Age range (years)	7–83	13–77	16–68
Syndrome	1 (1%)	1 (4%)	5 (31%)
M:F	52:54	16:12	6:10

Figure 1 Sites of occurrence, cardiac myxoma (150 tumours seen by the authors). Multiple and ventricular cardiac myxomas demonstrate a clinically distinct clinical profile. They arise in younger patients and are more likely to be associated with extracardiac lesions (myxoma syndrome).

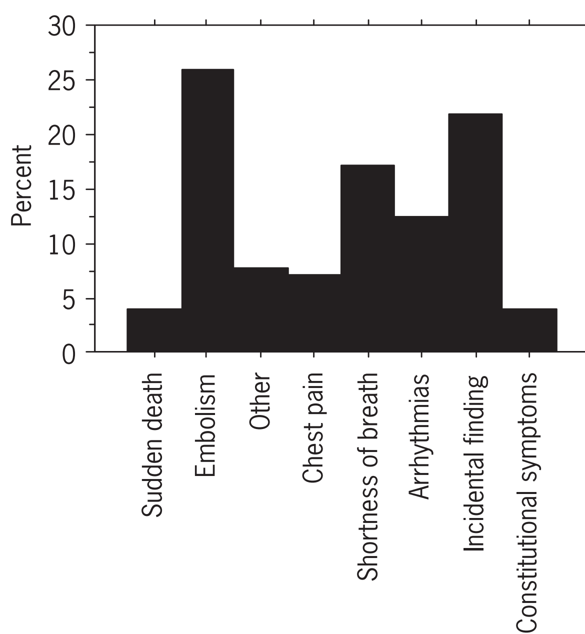


Figure 2 Cardiac myxoma, clinical symptoms at presentation. The most common single presentation relates to embolization. However, these tumours account for only 25% of cases. Sudden death may occur from cardiac myxoma by embolization to coronary vessels.

but significant proportion of cardiac myxomas (about 30%) have a rough, papillated surface partly covered by thrombus (**Figure 4; see colour plate section**). These tumours often cause symptoms because of embolization; that is, a portion of tumour or attached clot will break off and lodge somewhere in the patient's circulation. In most cases, the embolic site will be in the systemic circulation, as most cardiac myxomas are left-sided. Typical embolic sites are the brain, kidneys and arteries of the extremities. Emboli to the brain may be asymptomatic, or cause minor or even major strokes; if the retinal artery is involved, temporary blindness (amaurosis fugax) may ensue. Dislodged tumour is not synonymous with metastasis, as the myxoma does not seed organs and lymph nodes, forming new growths (metastatic deposits). However, embolic myxoma may occasionally cause weakness and aneurysm of the involved artery, especially in the cerebral circulation (Furuya *et al*., 1995). These aneurysms may be detected by angiography, but their relationship to a patient's symptoms is often unclear and, after removal of the primary tumour, they tend to regress. If a cardiac myxoma dislodges into the arteries of the arm or legs, ischaemic symptoms may result, including muscle cramping and pain, and rarely even gangrene.

Other symptoms that can be precipitated by cardiac myxoma include palpitations, syncope and chest pain. Right-sided tumours may cause pulmonary embolism if they possess an irregular surface. Approximately 2% of cardiac myxomas occur in families, and are associated with noncardiac diseases, including skin lesions and tumours and endocrine abnormalities (Singh and Lansing, 1996). In these patients, tumours tend to cause symptoms at an early age (childhood to young adulthood), by embolization.

The microscopic features of cardiac myxoma are heterogeneous, often within a single tumour, explaining the diversity of imaging findings, especially with contrast media. The name 'myxoma' derives from the abundant myxoid ground substance which is produced by the tumour cells (**Figure 5; see colour plate section**). Unlike most neoplasms, the tumour cells themselves are often a minority of the total cells of the tumour, which include reactive vessels, inflammatory cells and reactive stromal cells. Unlike sarcomas, myxomas never infiltrate the cardiac muscle, and respect the borders of the endocardium. The tumour cells are primitive, undifferentiated cells that form syncytia shaped as cords and rings, and have a variety of cytological features. In general, the nuclei are round to ovoid without prominent nucleoli, although in some tumours, large nucleoli may be present. The cytoplasm is abundant and eosinophilic. In many areas, especially away from the attachment to the myocardium, the tumour cells are intimately related to endothelial cells and capillaries, and appear to give rise to vascular structures. In 2% of cardiac myxomas, the tumour cells exhibit glandular differentiation, similar to the lining of the gastrointestinal tract. The interstitium of the tumour often contains dendritic cells, macrophages, lymphocytes, haemorrhage and breakdown blood products in the form of haemosiderin. In about 10% of myxomas, large deposits of haemosiderin may precipitate in elastic tissue of the interstitium, forming calcified structures that may be seen on radiographs (gamna-gandy bodies). In general, however, the majority of the haemosiderin is found engulfed within foamy macrophages. Ossification is also not infrequently present in cardiac myxoma, often those arising on the right side; approximately 10% of myxomas demonstrate radiographically detectable calcification (Basso *et al*., 1997).

Myofibroblastic (Intimal) Sarcoma

Myofibroblastic sarcomas are among the most common cardiac sarcomas, followed by angiosarcomas and undifferentiated sarcomas (**Figure 6**). Malignant neoplasms that occur in the atrial cavities are generally highly malignant myofibroblastic sarcomas grossly and clinically closely mimic the benign myxoma (Burke *et al*., 1992; Burke and Virmani, 1996). The initial symptoms may be similar to those of myxoma, including shortness of breath, chest pain, signs and symptoms of mitral valve or tricuspid valve obstruction and symptoms resulting from tumour embolism. Cardiac sarcomas tend to embolize less frequently than myxoma, however. Like myxoma, atrial sarcomas are often pedunculated, but are more likely to arise in the wall of the atria, and not the atrial septum, and often invade the wall of the atrium, epicardium or mitral valve. Those sarcomas that demonstrate infiltrative characteristics are easily distinguished from myxomas, which always respect the endocardial boundaries of the heart. Imaging studies may not be able to differentiated myxoma from sarcoma, although certain features suggest a malignancy, i.e. attachment site in the wall of the atrium and infiltration of the atrial wall. Contrast techniques that take advantage of the heterogeneous vascularity of cardiac myxomas and imaging studies that detect the large amount of haemorrhage within myxomas may also prove to be helpful adjuncts in the differential diagnosis of atrial myxoma and sarcoma. A small proportion (less than 20%) of intimal sarcomas of the heart arise in the cardiac ventricles, where they occur as endocardial-based tumours that obstruct blood flow.

The histological features of intimal sarcoma are heterogeneous, and include sarcoma subtypes that would by roughly synonymous with fibrosarcoma, malignant fibrous histiocytoma, osteosarcoma and chondrosarcoma. The

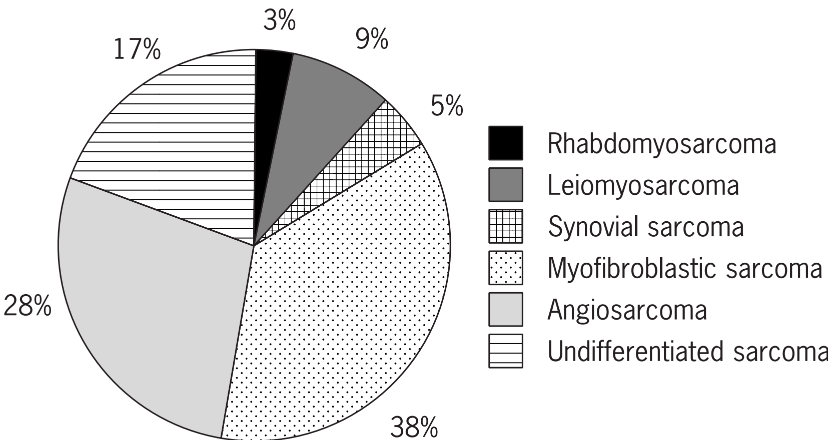


Figure 6 Histological types of cardiac sarcoma. In the authors' files, the majority of cardiac sarcomas include myofibroblastic, leiomyosarcomatous and undifferentiated tumours. Leiomyosarcomas are similar to myofibroblastic sarcomas, but are more purely smooth muscle cell in origin.

degree of osteo- or chondrosarcomatous differentiation is usually small, and it is rare to have a diffusely calcified sarcoma arising in the atrium. It must be kept in mind that ossifying tumours of the left atrium are usually malignant; small foci of ossification are not rare, however, in right atrial myxomas. The histological appearance of cardiac sarcomas is similar to that of sarcomas that arise in soft tissue and, overall, the full histological spectrum of soft tissue sarcomas may be encountered in the heart. However, those that arise within the atrium are primarily intracavitary. Because of their predominantly endoluminal growth, and because they are histologically very similar to sarcomas that arise within the great arteries (pulmonary artery and aorta), it is convenient for descriptive and taxonomic purposes to assume that they derive from the endocardium. Like the intima of arteries, the endocardium is composed of a layer of endothelial cells overlying an elastic layer, with a variable number of smooth muscle cells (or myofibroblastic cells) that occur sandwiched within. It is presumably these myofibroblastic cells, which are prone to proliferate, that undergo malignant transformation and result in the formation of sarcomas of the atria.

Angiosarcoma

Myofibroblastic sarcomas and angiosarcomas are the most common subtypes of cardiac sarcomas. The mean age at presentation for these sarcomas is 40 years (**Figure 7**). Angiosarcomas of the heart are one of the more common

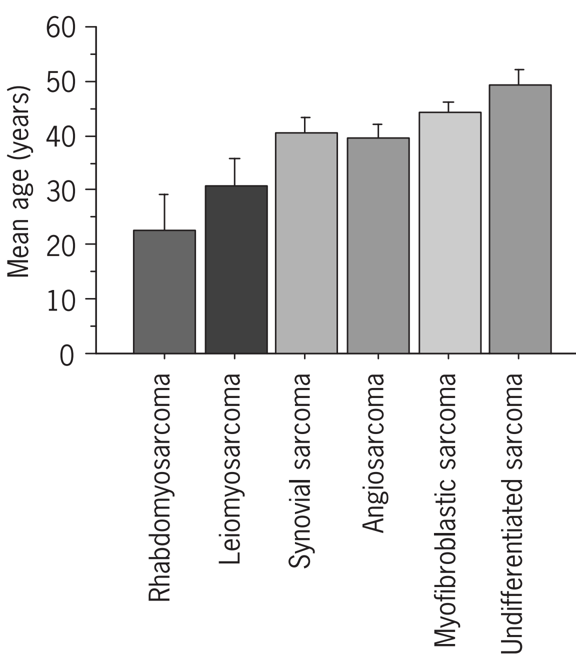


Figure 7 Mean age at presentation (years), by type of cardiac sarcoma. Rhabdomyosarcomas are tumours of adolescents and young adults. The overall mean age at presentation of cardiac sarcoma is 40 years.

cardiac sarcomas and, for unknown reasons, usually arise in the right atrium (Putnam *et al*., 1991; Burke *et al*., 1992; Tazelaar *et al*., 1992). There may be a slight bias for males over females (**Figure 8**). Unlike myofibroblastic sarcomas, they do not form pedunculated intracavitary masses, but invade early into the atrial wall and pericardium (**Figure 9; see colour plate section**). A few cardiac angiosarcomas are entirely confined to the pericardial space. The presenting symptoms of cardiac angiosarcomas, therefore, are typically related to pericardial disease, especially recurrent pericardial effusions or haemopericardium. A minority of cardiac angiosarcomas present with symptoms of obstruction of cardiac blood flow. Because cardiac angiosarcomas typically are right-sided, pulmonary metastases occur very early in the course of disease. In some patients, it is the presence of these lung lesions which cause the initial symptoms. The histological appearance of cardiac angiosarcoma is similar to that of soft tissue sarcomas. Rare examples of low-grade angiosarcomas (epithelioid haemangioendothelioma) have been described in the heart. However, epithelioid angiosarcomas, which are as lethal as histologically typical angiosarcomas, have not yet been reported to occur within the pericardium.

Miscellaneous Sarcomas

One would think initially that sarcomas with muscular differentiation (i.e. tumours composed of cells demonstrating features of striated muscle, or rhabdomyosarcomas) would predominate in the myocardium, as the bulk of the heart is composed of working muscle cells. However, cardiac muscle is terminally differentiated, with limited propensity for cellular division. Therefore, it makes sense that rhabdomyosarcomas are rare tumours, and make up less than 10% of all heart malignancies. As one would expect, the majority of these tumours occur in the walls of the heart, and they do not usually project into the lumen like intimal sarcomas, and rarely infiltrate the pericardium like angiosarcomas. Rhabdomyosarcomas of the heart are almost always of the embryonal type, and tend to occur in children and adolescents (Burke *et al*., 1992). Histologically, they are similar to embryonal rhabdomyosarcomas of soft tissue, although the so-called ‘alveolar’ subtype of rhabdomyosarcoma is infrequently seen in the myocardium.

Another variety of sarcoma that makes up about 5% of heart sarcomas is the synovial sarcoma, which, in the soft tissue, generally arises at or near joint spaces (Burke *et al*., 1992; Karn *et al*., 1994; Iyengar *et al*., 1995). The cell of origin of synovial sarcomas is unknown. These tumours typically demonstrate a biphasic histological growth pattern of spindle and glandular cells. In the heart, synovial sarcomas have a predilection for the pericardial space, but may occur in any cardiac site.

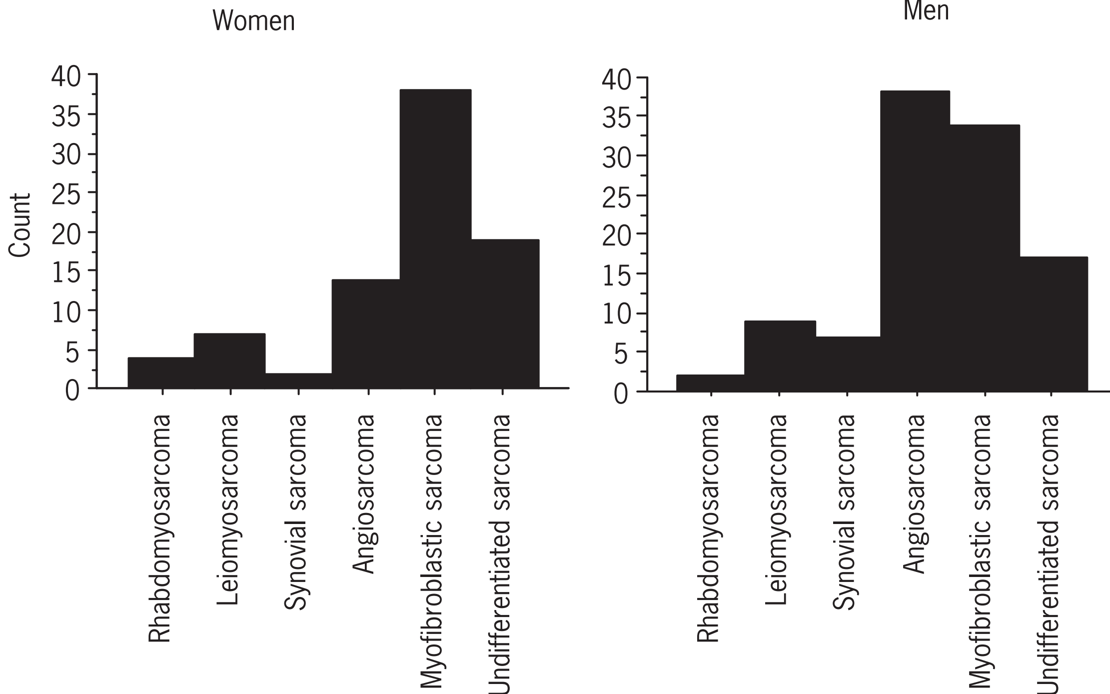


Figure 8 Gender distribution, cardiac sarcomas. There is no particular gender bias in cardiac sarcomas. In our experience, however, myofibroblastic sarcomas are slightly more common in women and angiosarcomas in men.

Sarcomas that are extremely rare in the cardiac location include liposarcomas, malignant localized fibrous tumours, malignant peripheral nerve sheath tumours and Ewing sarcoma (primitive neuroectodermal tumour). A fairly large proportion (up to 25%) of cardiac sarcomas defy classification, and represent a variety of undifferentiated spindle cell or small cell sarcomas.

Malignant Mesothelioma

Mesotheliomas are malignancies of the serosal lining tissues and, as such, may occur in the pleura, pericardium, peritoneum and inguinal sac. For reasons probably related to causation, they are most common in the pleura, which is at higher risk of exposure to inhaled asbestos than the other serosal sites. Pericardial mesotheliomas cause symptoms of chronic pericarditis, including pericardial pain, jugular venous distension, shortness of breath and eventually heart failure. The typical patient with pericardial mesothelioma may experience several episodes of pericarditis and pericardial effusion before a definitive diagnosis is made. The diagnosis may be confirmed only by biopsy sampling of pericardial tissue, and may be suggested by cytological analysis of pericardial fluid. Pericardial mesotheliomas, like angiosarcomas, often result in chronic effusion, but are much more likely to encase the heart in an unresectable constricting mass (**Figure 10; see colour plate section**). In early stages of disease, the primary differential diagnosis is chronic constrictive pericarditis, which is usually idiopathic, but may be related to autoimmune and other disorders (Burke and Virmani, 1996).

Pathologically, malignant mesotheliomas of the pericardium are identical with those of other sites. There is a bimodal appearance to these tumours microscopically, as there is for synovial sarcoma. However, the antigenic profile of mesothelioma is somewhat different from that of synovial sarcoma, and the latter tumour produces mucins not found in mesothelioma.

Malignant Lymphoma

Lymphomas arising in the heart are extremely uncommon; conversely, secondary involvement of the myocardium in patients with well advanced lymphomas arising in lymph nodes is not uncommon. Primary cardiac lymphomas are virtually all of B cell lineage, and may represent one of many histological subtypes, including low-, intermediate- and high-grade tumours (Burke and Virmani, 1996). For reasons that are not clearly understood, T cell lymphomas and Hodgkin disease are exceedingly rare as primary lesions within the heart or pericardium. The symptoms that cause lymphomas of the heart are varied, and depend on the site of origin within the myocardium. Typically, cardiac lymphomas arise within the walls of the organ, often in the atria with prominent epicardial involvement. There may be depressed cardiac function, obstruction to blood flow or pericardial effusions that produce the initial symptoms in patients with cardiac lymphoma.

Miscellaneous Neoplasms

Besides myxomas and sarcomas, primary neoplasms of the myocardium are vanishingly rare (Burke and Virmani, 1996). These include paragangliomas (phaeochromocytomas), neurofibromas and granular cell tumours. Cardiac paragangliomas generally occur in the atrial and atrial septum, where there are paraganglial tissues as a result of autonomic innervation of the heart. Cardiac paragangliomas may cause symptoms of hypertension. Rarely, cardiac paragangliomas may be malignant, and demonstrate the capability of developing distant metastases. However, in most patients, excision of the lesion is curative. The pathological and biological features of cardiac paraganglioma are similar to those that arise in paraganglial tissues elsewhere in the body, such as the adrenal gland and carotid bodies.

EPIDEMIOLOGY

Primary cardiac tumours are rare, so there are few accurate data to suggest their true prevalence. In general, it has been estimated that the incidence of cardiac neoplasms is approximately 0.01%, or one in 10 000 persons (Lam *et al*., 1993; Burke and Virmani, 1996). Of these, the majority of adult tumours are myxomas and sarcomas, at a ratio of between 2 : 1 and 8 : 1. The incidence of surgically excised cardiac tumours is estimated at more than three per million population per year (Reyen *et al*., 1998). There is no strong sex or racial predilection among primary heart tumours, although in some studies, more women than men suffer from cardiac myxomas, sometimes by a factor of 3 : 1 (Endo *et al*., 1997). In our experience, women are more likely to develop myofibroblastic sarcomas and men angiosarcomas (**Figure 8**). Epidemiological studies demonstrate that there are a variety of environmental factors and hereditary conditions that predispose to extracardiac soft tissue sarcomas. In the case of mesothelioma, it has become clear that there is a strong association with pericardial mesothelioma and asbestos exposure, as had been demonstrated for pleural and abdominal mesotheliomas. Pericardial mesotheliomas are rare, account for 1% of all mesotheliomas, and it has been estimated that there were 140 cases reported in the world literature in 1994 (Kaul *et al*., 1994), and fewer than 30 in the English literature (Thomason *et al*., 1994).

AETIOLOGY

Cardiac myxoma

There is little known about the aetiology of primary cardiac tumours. The aetiology of the various types of primary cardiac tumours is presumably the same as that for histologically similar neoplasms that arise in extracardiac sites. In the case of cardiac myxoma, which is unique to the heart, the only clues to aetiology come from those cases that are familial and may have a genetic basis. In 1980, Atherton *et al*. described a patient with skin pigmentation, neurofibromas and cardiac myxomas (Atherton *et al*., 1980). Later, the associations were refined to include lentigines, mucocutaneous myxomas and blue naevi, when it was recognized that the myxoid skin lesions were better classified as myxomas than myxoid neurofibromas. The list of associated conditions continues to grow, and includes Sertoli cell tumours of the testis, myxoid fibroadenoma of the breast, pituitary hyperactivity including growth hormone-producing adenoma, pigmented nodular adrenocortical disease and psammomatous melanotic schwannoma (Carney, 1995; Stratakis *et al*., 1996). The genetic mode of transmission of myxoma syndrome has not been determined, but autosomal dominant transmission is favoured in most families. It has been shown by several investigators that familial cardiac myxomas are more often multiple, recurrent and right-sided than sporadic myxomas.

Cardiac Sarcoma

The aetiology of cardiac sarcoma is unknown, and is probably similar to that of soft tissue sarcomas. Angiosarcoma may be induced by ionizing radiation, especially those of the liver in patients having received injections of radiation contrast material. There is a case report of a primary cardiac (pericardial) angiosarcoma in a patient with prior radiation therapy for mediastinal seminoma (Killion *et al*., 1996). In the majority of patients with cardiac angiosarcoma, however, there is no such prior history. Chromosomal and genetic defects have been described in soft tissue sarcomas but an extensive study of such defects in cardiac sarcomas has not been undertaken. It has been estimated that 7–33% of soft tissue sarcomas may have a genetic component (Hartley *et al*., 1993). A variety of exposures have been associated with extracardiac sarcomas, including arsenical pesticides and medications, phenoxy herbicides, dioxin, vinyl chloride, immunosuppressive drugs, alkylating agents, androgen-anabolic steroids, human immunodeficiency virus and human herpes virus type 8 (Zahm and Fraumeni, 1997).

Pericardial Mesothelioma

The most common site of malignant mesothelioma is the pleura. A majority of patients with pleural mesothelioma have a history of asbestos exposure. There is increasing evidence that at least a large minority of patients with pericardial mesotheliomas also have a history of asbestos exposure, and pericardial mesothelioma may be

experimentally induced by topical asbestos application (Burke and Virmani, 1996). The mechanism by which asbestos causes mesothelioma is uncertain, but it is known that certain forms of the inhaled element are more likely to result in malignancies than others. About 1% of asbestos-related mesotheliomas are pericardial and the rest are pleural and other sites. There are undoubtedly other stimuli for the development of mesothelioma, as some animals develop malignant mesothelioma of pericardium without any apparent asbestos exposure (Maltoni *et al.*, 1991; Chandra and Mansfield, 1999; Closa *et al.*, 1999). Simian virus induces pericardial mesotheliomas in hamsters, but a viral aetiology for human mesothelioma has not been demonstrated (Cicala *et al.*, 1993).

Cardiac Lymphoma

Approximately 50% of primary lymphomas of the heart arise in patients with acquired immunodeficiency, either iatrogenic (in cases of allograft recipients) or viral (in cases of AIDS) (Burke and Virmani, 1996). However, the heart is not a common site for the development of lymphomas in patients with AIDS or allografts (Tirelli *et al.*, 1994). The aetiology of immunosuppression-related lymphomas is believed related to infection by Epstein–Barr virus, which has the ability to transform B cells *in vitro*. In immunocompetent patients, Epstein–Barr virus is also associated with nasopharyngeal carcinomas and Burkitt lymphoma. Another virus which may have an aetiological role in immunosuppression-related lymphomas is another herpes virus, HHV-8, also known has Kaposi sarcoma-associated herpes virus. HHV-8-related lymphomas are often confined to serosal spaces without an identifiable contiguous tumour mass. The usual location for these lesions, however, is the peritoneum or pleura, with only rare examples occurring in the pericardium (Nador *et al.*, 1996).

The remaining 50% of primary cardiac lymphomas are not related to immunosuppression or herpes viruses (Bogaert *et al.*, 1995; Versluis *et al.*, 1995; Ito *et al.*, 1996). These tumours are fairly rare and aetiological agents have not been identified. Presumably, the pathoaetiology of cardiac lymphoma is similar to that of nodal lymphomas.

SCREENING AND PREVENTION

Because of the rarity of primary cardiac neoplasms, there is little if any role for screening or attempts at prevention. The only possible exception to this rule would be in easily identifying individuals at greatly increased risk for the development of a primary cardiac myxoma. In cases of myxoma syndrome, family relatives should be aggressively evaluated for the presence of cardiac myxoma, as these can present with embolization and even sudden death. We are aware of a case of a young boy with cutaneous features of the myxoma syndrome, whose father had a cardiac myxoma removed at a young age. The boy was not evaluated medically, and died suddenly due to coronary embolism from cardiac myxoma at a young age.

ULTRASTRUCTURE/AND IMMUNOHISTOCHEMISTRY

Cardiac Myxoma

The electron microscopic features of cardiac myxoma are of limited value in diagnosis, owing to the heterogeneous nature of cardiac myxoma and limited sampling afforded by ultrastructural techniques. Studies of myxoma have demonstrated myxoma cells to be primitive mesenchymal cells with intermediate filaments and cytoplasmic organelles. The cells are present either as single stellate cells, or groups, and have primitive tight intracellular junctions lacking desmosomes. Myxoma cells are best described ultrastructurally as embryonic cells that occasionally show features of myofibroblasts. The ground substance contains fine electron-dense granules which have been described as identical with the proteoglycan granules seen in the ground substance of cartilage (Feldman *et al.*, 1977; Goldman *et al.*, 1987).

The immunohistochemical features of cardiac myxoma indicate a pluripotent myxoma cell capable of divergent differentiation (Burke and Virmani, 1993, 1996; Berrutti and Silverman, 1996; Deshpande *et al.*, 1996). The most consistent antigenic type of cellular antigen expressed in myxomas is that of endothelial differentiation (Silverman and Berrutti, 1996). Other antigens found on myxoma cells have included smooth muscle cell antigens, neural antigens and epithelial antigens. The use of immunohistochemical techniques in the diagnosis of cardiac myxoma is limited. Recently, immunohistochemical data suggest that the myxoma is a weakly proliferative lesion without modulation of oncogene/tumour suppressor gene products (Suvarna and Royds, 1996).

Cardiac Sarcoma

Immunohistochemical studies are useful for distinguishing cardiac sarcomas from other processes, but are not particularly useful in subtyping the lesions (Burke *et al.*, 1992; Tazelaar *et al.*, 1992). An exception to this rule includes the diagnosis of rare types of sarcoma, such as malignant localized fibrous tumour (CD 34 positivity) and synovial sarcoma. Immunohistochemistry is currently of little use in providing prognostic data for cardiac sarcomas, which are uniformly lethal. In general, the immunohistochemical findings of cardiac sarcomas are similar to those of soft tissue sarcomas, and the reader is directed to relevant chapters in this volume. (See the chapter *Soft Tissues*.)

Cardiac Lymphoma and Pericardial Mesothelioma

The immunohistochemical and ultrastructural features of these neoplasms vary by cardiac or noncardiac origin. The reader is referred to other sections of this volume for immunohistochemical and ultrastructural characterization of lymphomas and mesotheliomas. (See the chapters *Lymph Nodes* and *Pleura and Peritoneum*.)

MOLECULAR GENETIC FINDINGS

Myxoma

Tissue cultures of cardiac myxoma cells have demonstrated chromosomal abnormalities both in familial and non-familial cases. A locus on chromosome 2p16 has been identified that may be involved in the regulation of genomic stability of dividing myxoma cells, in particular the structure of telomeres in replicating chromosomes (Stratakis *et al*., 1996). Other studies have shown clonal and non-clonal abnormalities including dicentric chromosomes and telomeric associations, clonal telomeric associations between chromosomes 13 and 15 and nonclonal telomeric associations between chromosomes 12 and 17 (Richkind *et al*., 1994; Dijkhuizen *et al*., 1995).

Cardiac Sarcoma

Molecular genetic and chromosomal abnormalities are helpful in the classification of soft tissue sarcomas (Pollock, 1994; Mohamed *et al*., 1997; Singer, 1999). In the heart, molecular diagnosis has been helpful in the classification of cardiac synovial sarcomas (Karn *et al*., 1994; Iyengar *et al*., 1995), specifically the X;18 translocation. The molecular aetiology of extracardiac soft tissue sarcoma involves growth factors and their receptors, nuclear and cytoplasmic oncogenes and tumour-suppressor genes (Pollock, 1994).

Cardiac Lymphoma

Most cardiac lymphomas are of B cell origin (Bogaert *et al*., 1995; Chao *et al*., 1995; Versluis *et al*., 1995; Ito *et al*., 1996; Chim *et al*., 1997). Those that arise in immunocompromised patients are often associated with Epstein–Barr virus. The molecular genetics of nodal lymphomas have been extensively studied, and are covered in other areas of this volume.

Pericardial Mesothelioma

The karyotypes of malignant mesotheliomas are complex, and often preclude the identification of primary chromosome abnormalities. Recently, peritoneal mesothelioma with a balanced t(3;3)(p14;q29) translocation has been described. Molecular genetic abnormalities of pericardial mesotheliomas have not been described.

PROGNOSTIC FACTORS

Cardiac Myxoma

The prognosis for cardiac myxoma is excellent. With adequate surgical excision, the recurrence rate is vanishingly small. Patients with recurrent myxoma should be considered possible myxoma syndrome victims, especially if they are younger than 40 years of age. Alternatively, an incorrect diagnosis should be considered, as many recurrent cardiac sarcomas are initially misdiagnosed as myxomas. In the majority of a series of cardiac myxoma, surgical treatment is curative in virtually all cases (Lazzara *et al*., 1991; Actis Dato *et al*., 1993; Bastos *et al*., 1995; Gawdzinski and Sypula, 1996; Bjessmo and Ivert, 1997). Reports of malignant myxoma should be regarded with suspicion. Rare well-documented examples of myxomas with embolization (‘metastasis’) to the cerebral vessels and skin have been reported (Reed *et al*., 1989; Scarpelli *et al*., 1997).

Cardiac Sarcoma

In stark contrast to cardiac myxoma, the prognosis of cardiac sarcoma is dismal, and generally measured in months. Although surgery is indicated for accurate diagnosis and palliation, and incurs a low mortality risk, the outcome is death within 2–3 years (Murphy *et al*., 1990; Putnam *et al*., 1991; Endo *et al*., 1997; Centofanti *et al*., 1999). The most important histopathological prognostic indicators are increased mitotic rate and the presence of necrosis (Burke *et al*., 1992). In addition, patients with right-sided tumours and those presenting with metastatic disease have an especially poor prognosis. Even in patients with completely resectable tumours, most recur at a median time of 10 months. The median time to progression, in one series, was shorter in patients presenting a cardiac angiosarcoma than other histological types (3 vs 14 months, $P < 0.01$), with an overall 2-year survival rate of 26% (Llombart-Cussac *et al*., 1998).

Pericardial Mesothelioma

The prognosis of pericardial mesothelioma is poor: 50% of patients are dead at 6 months, and an exceptional patient may live as long as 48 months. The mean survival of patients with pericardial mesothelioma is shorter than that of patients with pleural and peritoneal mesotheliomas (Hillerdal, 1983).

Cardiac Lymphoma

The prognosis of primary cardiac lymphoma is difficult to ascertain, because of the extreme rarity of the lesion,

and concomitant diseases in patients who are immunosuppressed. As with nodal lymphomas, the prognosis, and also the treatment, would vary by the grading of the lymphoma. As low-, intermediate- and high-grade lymphomas have all been described to arise in the heart, the prognosis would probably vary with the histological subtype.

OVERVIEW OF PRESENT CLINICAL MANAGEMENT

The clinical management of cardiac myxoma is that of surgical excision, which is curative (Mishra *et al*., 1991; Actis Dado *et al*., 1993; Castells *et al*., 1993; Bjessmo and Ivert, 1997). Occasionally, a patch repair of the atrial septum is required and, rarely, valve reconstruction or repair.

The initial treatment of cardiac sarcoma is surgical excision. In many cases, the preoperative diagnosis is cardiac myxoma. However, complete excision is often difficult, unlike myxoma, and surgical margins for adequacy of excision should be ascertained by intraoperative consultation with the pathologist. The long-term management of cardiac sarcomas is difficult, as the use of radiation and chemotherapy is of limited benefit (Noirclerc *et al*., 1997; Movsas *et al*., 1998). However, many centres offer these modalities of treatment, which are similar to those administered for soft tissue sarcomas. The only chance at a cure in a patient with heart sarcoma is orthotopic heart transplantation (Noirclerc *et al*., 1997); long-term survival has also been attained with fractionated radiation therapy (Movsas *et al*., 1998). Adjuvant chemotherapy has been advised in high-grade cardiac sarcomas (Ceresoli *et al*., 1999), although it has not been demonstrated to be effective (Llombart-Cussac *et al*., 1998).

REFERENCES

Actis Dato, G. M. (1993). Long-term follow-up of cardiac myxomas (7–31 years). *Journal of Cardiovascular Surgery*, **34**, 141–143.

Atherton, D., *et al*. (1980). A syndrome of various cutaneous pigmented lesions, myxoid neurofibromata and atrial myxoma: the NAME syndrome. *British Journal of Dermatology*, **103**, 421–429.

Basso, C., *et al*. (1997). Cardiac lithomyxoma. *American Journal of Cardiology*, **80**, 1249–1251.

Bastos, P., *et al*. (1995). Cardiac myxomas: surgical treatment and long-term results. *Cardiovascular Surgery*, **3**, 595–597.

Berrutti, L. and Silverman, J. S. (1996). Cardiac myxoma is rich in factor XIIIa positive dendrophages: immunohistochemical study of four cases. *Histopathology*, **28**, 529–535.

Bjessmo, S. and Ivert, T. (1997). Cardiac myxoma: 40 years' experience in 63 patients. *Annals of Thoracic Surgery*, **63**, 697–700.

Bogaert, J., *et al*. (1995). High-grade immunoblastic sarcoma: an unusual type of a primary cardiac non-Hodgkin lymphoma. *RoFo. Fortschritte aut dem Gebiete der Röntgenstranlen under neuen bildgebenden Verfahren*, **162**, 186–188.

Burke, A. P., *et al*. (1992). Primary sarcomas of the heart. *Cancer*, **69**, 387–395.

Burke, A. P. and Virmani, R. (1993). Cardiac myxoma. A clinicopathologic study. *American Journal of Clinical Pathology*, **100**, 671–680.

Burke, A. P. and Virmani, R. (1996). Classification and incidence of cardiac tumours. In: *Atlas of Tumor Pathology. Tumors of the Cardiovascular System*. 1–12, 21–46, 127–170, 181–194. (Armed Forces Institute of Pathology, Washington, DC).

Carney, J. A. (1995). Carney complex: the complex of myxomas, spotty pigmentation, endocrine overactivity, and schwannomas. *Seminars in Dermatology*, **14**, 90–98.

Castells, E., *et al*. (1993). Cardiac myxomas: surgical treatment, long-term results and recurrence. *Journal of Cardiovascular Surgery*, **34**, 49–53.

Centofanti, P., *et al*. (1999). Primary cardiac tumors: early and late results of surgical treatment in 91 patients. *Annals of Thoracic Surgery*, **68**, 1236–1241.

Chandra, M. and Mansfield, K. G. (1999). Spontaneous pericardial mesothelioma in a rhesus monkey. *Journal of Medical Primatology*, **28**, 142–144.

Cicala, C., *et al*. (1993). SV40 induces mesotheliomas in hamsters. *American Journal of Pathology*, **142**, 1524–1533.

Closa, J. M., *et al*. (1999). Pericardial mesothelioma in a dog: long-term survival after pericardiectomy in combination with chemotherapy. *Journal of Small Animal Practice*, **40**, 383–386.

Deshpande, A., *et al*. (1996). Phenotypic characterization of cellular components of cardiac myxoma: a light microscopy and immunohistochemistry study. *Human Pathology*, **27**, 1056–1059.

Dijkhuizen, T., *et al*. (1995). Rearrangements involving 12p12 in two cases of cardiac myxoma. *Cancer Genetics and Cytogenetics*, **82**, 161–162.

Endo, A., *et al*. (1997). Characteristics of 161 patients with cardiac tumors diagnosed during 1993 and 1994 in Japan. *American Journal of Cardiology*, **79**, 1708–1711.

Feldman, P. S., *et al*. (1977). An ultrastructural study of seven cardiac myxomas. *Cancer*, **40**, 2216–2232.

Furuya, K., *et al*. (1995). Histologically verified cerebral aneurysm formation secondary to embolism from cardiac myxoma. Case report. *Journal of Neurosurgery*, **83**, 170–173.

Gawdzinski, M. P. and Sypula, S. (1996). The long term results of treatment of heart myxomas with special attention to very rare myxoma of the right ventricle. *Journal of Cardiovascular Surgery*, **37**, 121–129.

Goldman, B., *et al*. (1987). Glandular cardiac myxomas. Histologic, immunohistochemical, and ultrastructural evidence of epithelial differentiation. *Cancer*, **15**, 1767–1775.

Hartley, A. L., *et al*. (1993). Patterns of cancer in the families of children with soft tissue sarcoma. *Cancer*, **72**, 923–930.

Hillerdal, G. (1983). Malignant mesothelioma 1982: review of 4710 published cases. *British Journal of Diseases of the Chest*, **77**, 321–343.

Iyengar, V., *et al*. (1995). Synovial sarcoma of the heart. Correlation with cytogenetic findings. *Archives of Pathology and Laboratory Medicine*, **119**, 1080–1082.

Karn, C. M., *et al*. (1994). Cardiac synovial sarcoma with translocation (X;18) associated with asbestos exposure. *Cancer*, **73**, 74–78.

Kaul, T. K., *et al*. (1994). Primary malignant pericardial mesothelioma: a case report and review. *Journal of Cardiovascular Surgery*, **35**, 261–267.

Killion, M. J., *et al*. (1996). Pericardial angiosarcoma after mediastinal irradiation for seminoma. A case report and a review of the literature. *Cancer*, **78**, 912–917.

Lam, K. Y., *et al*. (1993). Tumors of the heart. A 20-year experience with a review of 12,485 consecutive autopsies. *Archives of Pathology and Laboratory Medicine*, **117**, 1027–1031.

Llombart-Cussac, A., *et al*. (1998). Adjuvant chemotherapy for primary cardiac sarcomas: the IGR experience. *British Journal of Cancer*, **78**, 1624–1628.

Lazzara, R. R., *et al*. (1991). Cardiac myxomas: results of surgical treatment. *Journal of Cardiovascular Surgery*, **32**, 824–827.

Mohamed, A. N., *et al*. (1997). Cytogenetic aberrations and DNA ploidy in soft tissue sarcoma. A Southwest Oncology Group Study. *Cancer Genetics and Cytogenetics*, **99**, 45–53.

Movsas, B., *et al*. (1998). Primary cardiac sarcoma: a novel treatment approach. *Chest*, **114**, 648–652.

Murphy, M., *et al*. (1990). Surgical treatment of cardiac tumors: a 25 year experience. *Annals of Thoracic Surgery*, **49**, 612–617.

Nador, R. G., *et al*. (1996). Primary effusion lymphoma: a distinct clinicopathologic entity associated with the Kaposi's sarcoma-associated herpes virus. *Blood*, **88**, 645–656.

Noirclerc, M., *et al*. (1997). Primary cardiac sarcoma treated by orthoptic cardiac transplantation. Apropos of a case. *Archives des Maladies du Coeur et des Vaisseaux*, **90**, 1539–1543.

Putnam, J. B., Jr, *et al*. (1991). Primary cardiac sarcomas. *Annals of Thoracic Surgery*, **51**, 906–910.

Reed, R. J., *et al*. (1989). Embolic and metastatic cardiac myxoma. *American Journal of Dermatopathology*, **11**, 157–165.

Reyen, K., *et al*. (1998). Heart operations for heart tumors in Germany – results of 1996 survey. *Zeitschrift für Kardiologie*, **87**, 331–335.

Richkind, K. E., *et al*. (1994). Cardiac myxoma characterized by clonal telomeric association. *Genes, Chromosomes and Cancer*, **9**, 68–71.

Scarpelli, M., *et al*. (1997). Cardiac myxoma with glandular elements metastatic to the brain 12 years after the removal of the original tumor. *Clinical Neuropathology*, **16**, 190–194.

Silverman, J. S. and Berrutti, L. (1996). Cardiac myxoma immunohistochemistry: value of CD34, CD31, and factor XIIIa staining. *Diagnostic Cytopathology*, **15**, 455–456.

Singer, S. (1999). New diagnostic modalities in soft tissue sarcoma. *Seminars in Surgical Oncology*, **17**, 11–22.

Stratakis, C. A., *et al*. (1996). Cytogenetic and microsatellite alterations in tumors from patients with the syndrome of myxomas, spotty skin pigmentation, and endocrine overactivity (Carney complex). *Journal of Clinical Endocrinology and Metabolism*, **81**, 3607–3614.

Suvarna, S. K. and Royds, J. A. (1996). The nature of the cardiac myxoma. *International Journal of Cardiology*, **57**, 211–216.

Tazelaar, H. D., *et al*. (1992). Pathology of surgically excised primary cardiac tumors. *Mayo Clinic Proceedings*, **67**, 957–965.

Tirelli, U., *et al*. (1994). Malignant tumours in patients with HIV infection. *British Medical Journal*, **308**, 1148–1153.

Thomason, R., *et al*. (1994). Primary malignant mesothelioma of the pericardium. Case report and literature review. *Texas Heart Institute Journal*, **21**, 170–174.

Versluis, P. J., *et al*. (1995). Primary malignant lymphoma of the heart: CT and MRI features. *RoFo. Fortschritte auf dem Gebiete der Röntgenstrahlen und der neuen bildgebenden Verfahren*, **162**, 533–534.

FURTHER READING

Actis Dato, G. M., *et al*. (1993). Long-term follow-up of cardiac myxomas (7–31 years). *Journal of Cardiovascular Surgery*, **34**, 141–143.

Bear, P. and Moodie, D. (1987). Malignant primary cardiac tumors. The Cleveland Clinic experience, 1956 to 1986. *Chest*, **92**, 860–862.

Bjessmo, S. and Ivert, T. (1997). Cardiac myxoma: 40 years' experience in 63 patients. *Annals of Thoracic Surgery*, **63**, 697–700.

Burke, A. P. and Virmani, R. (1993). Cardiac myxoma. A clinicopathologic study. *American Journal of Clinical Pathology*, **100**, 671–680.

Burke, A. P. and Virmani, R. (1996). In: *Atlas of Tumor Pathology. Tumors of the Cardiovascular System.* 1–12 (Armed Forces Institute of Pathology, Washington, DC).

Dijkhuizen, T., *et al*. (1995). Rearrangements involving 12p12 in two cases of cardiac myxoma. *Cancer Genetics and Cytogenetics*, **82**, 161–162.

Llombart-Cussac, A., *et al*. (1998). Adjuvant chemotherapy for primary cardiac sarcomas: the IGR experience. *British Journal of Cancer*, **78**, 1624–1628.

Movsas, B., *et al*. (1998). Primary cardiac sarcoma: a novel treatment approach. *Chest*, **114**, 648–652.

Murphy, M., *et al*. (1990). Surgical treatment of cardiac tumors: a 25 year experience. *Annals of Thoracic Surgery*, **49**, 612–617.

Putnam, J. B., Jr, *et al*. (1991). Primary cardiac sarcomas. *Annals of Thoracic Surgery*, **51**, 906–910.

Chapter 51

Neuromuscular System

Kar-Ming Fung
University of Texas, MD Anderson Cancer Center, Houston, TX, USA

John Q. Trojanowski
University of Pennsylvania, Philadelphia, PA, USA

CONTENTS

- Normal Development and Structure
- Tumour Pathology
- Epidemiology
- Aetiology
- Screening and Prevention
- Gross/Histopathology/Preinvasive lesions/Ultrastructure/Immunohistochemistry
- Molecular Genetics
- Prognostic Factors
- Overview of Present Clinical Management

NORMAL DEVELOPMENT AND STRUCTURE

The nervous system and skeletal muscular system are both electrically active and they have an intimate functional relationship. However, tumours that arise in these two systems are very different. Whereas primary tumours of neuroepithelial differentiation arising from the central nervous system (CNS) and peripheral nervous system (PNS) are common, primary tumours of skeletal muscular differentiation that arise from skeletal muscle are extremely rare. In contrast, most primary tumours in skeletal muscles have mesenchymal differentiation of other lineages and they presumably arise from nonmuscular tissue adjacent to or within the muscle. Interestingly, primary tumours with muscular differentiation (e.g. rhabdomyosarcoma) usually arise in soft tissue rather than skeletal muscle. Therefore, it is most appropriate to consider these tumours with other soft tissue tumours. The reader should refer to the chapter on soft tissue tumours in this section and the Further Reading section for further details.

Tumours of the nervous system display morphological, immunohistochemical and molecular features that recapitulate their normal developing and mature counterparts. Knowledge on normal development and anatomy of the nervous system is fundamental for the study of these tumours.

The mature nervous system is composed of neurons and their supporting cells. In the CNS, neurons and axons are supported by glial cells, namely astrocytes, oligodendrocytes, ependymal cells and microglia. Choroid plexus is neuroepithelial tissue specialized in the production of cerebral spinal fluid. Other supporting cells of the nervous system include the arachnoid cells and a variety of mesenchymal cells. Neurons and axons in the PNS are supported by Schwann cells that are of neuroepithelial (neurocrest) origin and other mesenchymal cells. Tumours of the nervous system can have features reminiscent of any or a combination of these components but they do not necessarily arise from these components.

Formation of the CNS and PNS is a highly organized process that involves induction, cell division, differentiation, migration and cell death. During early development, the ectoderm undergoes neural induction by the underlying chordamesoderm and becomes the neuroectoderm. The neuroectoderm becomes a thickened layer of cells and forms the neural plate that in turn folds into a tube that will eventually develop into the brain and spinal cord. Part of the neural plate will give rise to the neural crest cells that give rise to Schwann cells, adrenal medulla and cell types such as melanocytes, odontoblasts and many mesenchymal and musculoskeletal components of the head.

At the dawn of development, the neural tube is composed of a thick, pseudostratified, columnar neuroepithelium that is composed of primitive neuroepithelial cells. Then, these

cells segregate into a highly cellular ventricular zone and the paucicellular marginal zone. The ventricular zone will give rise to all neurons and glial cells with the exception of microglial cells that belong to the mononuclear phagocytic cell lineage. Through the guidance of radial glia, developing neurons and glial cells migrate out from the ventricular zone and segregate into several foetal layers between the ventricular zone and marginal zone. Such foetal laminations will evolve into the six-layer adult cortex during maturation.

The maturation process involves interaction between many temporally and developmentally regulated genes. Young neurons will lose their ability to divide as they develop into mature neurons and over 50% of the young neurons will undergo apoptosis during maturation. Glial cells will develop into several specialized forms including astrocytes, oligodendrogliocytes, ependymal cells and epithelium of choroid plexus. In contrast to neurons, glial cells maintain their ability to divide in the mature nervous system.

In-depth discussion on development is beyond the scope of this chapter and the reader is referred to the Further Reading section.

TUMOUR PATHOLOGY

The central goal of tumour classification is to correlate pathological and biological properties with prognostic data. Classification of tumours of the nervous system, similar to tumours of other organ systems, is essentially based on a comparison of similarities between the neoplastic tissue and normal mature or developing counterparts. In the past 100 years or so, such comparison has been limited primarily to morphological features. In the recent past, comparison of the molecular phenotypes, mainly achieved by utilization of immunohistochemical techniques, *in situ* hybridization and cytogenetics have been used. The World Health Organisation (WHO) Classification of Tumours of the Nervous System (Kleihues and Cavenee, 2000) was formulated on the basis of such information (**Table 1**). In general, tumours are named according to the cell type that they most resemble. For example, an astrocytoma is a tumour that has features resembling astrocytes. Recent advances in molecular biology and cell biology and recognition of molecular signatures specific to particular types of tumour will presumably play a more important role in future classification systems.

Tumours of the Central Nervous System

The pathology of CNS tumours is a vast topic and it is not the intention of the authors to cover every aspect in this

Table 1 2000 WHO Classification of Tumours of the Nervous System

Tumours of neuroepithelial tissue
Astrocytic tumours
Diffuse astrocytoma
Fibrillary astrocytoma
Protoplasmic astrocytoma
Gemistocytic astrocytoma
Anaplastic astrocytoma
Glioblastoma
Giant cell glioblastoma
Gliosarcoma
Pilocytic astrocytoma
Pleomorphic xanthoastrocytoma
Subependymal giant cell astrocytoma
Oligodendroglial tumours
Oligodendroglioma
Anaplastic oligodendroglioma
Mixed gliomas
Oligoastrocytoma
Anaplastic oligoastrocytoma
Ependymal tumours
Ependymoma
Cellular
Papillary
Clear cell
Tanycytic
Anaplastic ependymoma
Myxopapillary ependymoma
Subependymoma
Choroid plexus tumours
Choroid plexus papilloma
Choroid plexus carcinoma
Glial tumours of uncertain origin
Astroblastoma
Gliomatosis cerebri
Chordoid glioma of the 3rd ventricle
Neuronal and mixed neuronal–glial tumours
Gangliocytoma
Dysplastic gangliocytoma of cerebellum (Lhermitte–Duclos)
Desmoplastic infantile astrocytoma/ganglioglioma
Dysembryoplastic neuroepithelial tumour
Ganglioma
Anaplastic ganglioglioma
Central neurocytoma
Cerebellar liponeurocytoma
Paraganglioma of the filum terminale
Neuroblastic tumours
Olfactory neuroblastoma (aesthesioneuroblastoma)
Olfactory neuroepithelioma
Neuroblastomas of the adrenal gland and sympathetic nervous system
Pineal parenchymal tumours
Pineocytoma
Pineoblastoma
Pineal parenchymal tumour of intermediate differentiation
Embryonal tumours
Medulloepithelioma

Table 1 (*Continued*)

Ependymoblastoma
Medulloblastoma
Desmoplastic medulloblastoma
Large cell medulloblastoma
Medullomyoblastoma
Melanotic medulloblastoma
Supratentorial primitive neuroectodermal tumour (PNET)
Neuroblastoma
Ganglioneuroblastoma
Atypical teratoid/rhabdoid tumour
Tumours of peripheral nerves
Schwannoma (neurilemmoma, neurinoma)
Cellular
Plexiform
Melanotic
Neurofibroma
Plexiform
Perineurioma
Intraneural perineurioma
Soft tissue perineurioma
Malignant peripheral nerve sheath tumour (MPSNST)
Epithelioid
MPSNST with divergent mesenchymal and/or epithelial differentiation
Melanotic
Melanotic psammomatous
Tumours of the meninges
Tumours of meningothelial cells
Meningioma
Meningothelial
Fibrous (fibroblastic)
Transitional (mixed)
Psammomatous
Angiomatous
Microcystic
Secretory
Lymphoplasmacyte-rich
Metaplastic
Clear cell
Chordoid
Atypical
Papillary
Rhabdoid
Anaplastic meningioma
Mesenchyma, nonmeningothelial tumours
Lipoma
Angiolipoma
Hibernoma
Liposarcoma (intracranial)
Solitary fibrous tumour
Fibrosarcoma
Malignant fibrous histiocytoma
Leiomyoma
Leiomyosarcoma
Rhabdomyoma
Rhabdomyosarcoma
Chondroma
Chondrosarcoma
Osteoma
Osteochondroma
Osteosarcoma
Haemanagioma
Epithelioid haemangioendothelioma
Haemangiopericytoma
Angiosarcoma
Kaposi sarcoma
Primary melanocytic lesions
Diffuse melanocytosis
Melanocytoma
Malignant melanoma
Meningeal melanomatosis
Tumours of uncertain histiogenesis
Haemangioblastoma
Lymphomas and haematopoietic neoplasms
Malignant lymphomas
Plasmacytoma
Granulocytic sarcoma
Germ cell tumors
Germinoma
Embryonal carcinoma
Yolk sac tumour
Choriocarcinoma
Teratoma
Mature
Immature
Teratoma with malignant transformation
Mixed germ cell tumours
Tumours of the sellar region
Craniopharyngioma
Adamantinomatous
Papillary
Granular cell tumour
Metastatic tumours

Source: Kleihues and Cavenee, 2000, *World Health Organisation Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System* (International Agency for Research on Cancer, Lyon).

chapter. The reader should consult the Further Reading section for further details.

Tumours of the CNS can be separated into several major categories (**Table 2**). First and foremost, they can be separated into primary and secondary tumours. Primary tumours are those that arise within the central nervous system and can be categorized into several major groups.

The first category consists of tumours that have features of primitive neuroepithelium. In the WHO Classification (Kleihues and Cavenee, 2000), this category includes the embryonal tumours and pineoblastomas. Some tumours such as olfactory esthesioneuroblastoma also belong to this category, although they do not occur within the cranial cavity. Essentially, these tumours resemble the primitive neuroectoderm or primitive neuroepithelium.

Table 2 Major biological categories of tumours of the nervous system

Central nervous system (CNS)
Category 1: Primary tumour with features of primitive neuroepithelial cell
Category 2: Primary tumour with features of mature neuroepithelial cell
Category 3: Primary tumour with features of meninges and mesenchyme
Category 4: Primary tumour with features of tissue normally found outside the central nervous system
Category 5: Secondary (metastatic) tumour
Peripheral nervous system (PNS)
Category 1: Primary tumours with features of neuroblasts/mature neurons
Category 2: Primary tumours with features of mature supporting elements of the peripheral nervous system
Category 3: Primary tumours of paraganglia
Category 4: Secondary tumour

The second category consists of tumours which display features of mature neuroepithelial cells such as glial cells and neurons. These tumours are further separated into three major subcategories. The first subcategory has features of glial cells, the supporting elements of the CNS. These are among the most common primary tumours of the CNS and they are formed by cells that resemble astrocytes and/or oligodendrocytes and/or ependymal cells. They are, therefore termed astrocytoma, oligodendroglioma and ependymoma, respectively. Not infrequently, they may have features of more than one type of glial cell in which case they are termed mixed gliomas. The second subcategory contains a variable neuronal component. Tumours composed exclusively of neoplastic neurons are uncommon. Very often, the neoplastic neurons or ganglion cells are mixed with neoplastic glial cells, most often neoplastic astrocytes. This category constitutes the neuronal and mixed neuronal–glial tumours of the WHO Classification. The third subcategory consists of choroid plexus tumours. Although the choroid plexus is regarded as modified ependyma, it has features of epithelial cells rather than glial cells. Choroid plexus tumours also behave like epithelial tumours rather than glial tumours. For this reason, choroid plexus tumours are regarded as a separate category.

The third category consists of tumours that have features resembling meninges and other supporting mesenchymal components of the CNS. Meningiomas have features of arachnoid cells that cover leptomeninges. Tumours composed of other mesenchymal elements are named after the mature tissue they most resemble. Melanocytic tumours of the CNS arise most commonly from the leptomeninges.

The fourth category consists of primary tumours that resemble mature and developing tissue that may or may not be normally present in the cranial cavity, including germ cell tumours, primary haematopoietic tumours such as lymphoma and craniopharyngioma.

Secondary (metastatic) tumours are extremely common, primarily in adults. In fact, they are the most common tumour of the CNS. This should not be surprising since the brain consumes about 20% of the cardiac output but comprises only about 2% of the body mass. Secondary tumours of the CNS include metastatic carcinomas, malignant melanomas and sarcomas as well as secondary involvement of the CNS by haematopoietic tumours such as lymphomas and leukaemias.

Tumours of the Peripheral Nervous System

Tumours of the PNS can also be separated into several major categories (**Table 2**). Similar to CNS tumours, there are PNS tumours that recapitulate features of primitive neuroepithelial cells or neuroblasts and some of them contain mature neuronal elements. These tumours are classified under neuroblastic tumours in the WHO Classification.

The second major category consists of tumours that have features of the mature supporting elements of the PNS such as Schwann cells. These tumours are named after their normal counterparts and are termed schwannoma, neurofibroma and perineuroma. Some have aggressive biological properties and cannot be clearly identified phenotypically clearly as schwannoma or neurofibroma. Because of their malignant behaviour and mixed histological features, they are collectively termed malignant peripheral nerve sheath tumours (MPNSTs) and stand by themselves in a distinct subcategory.

The third major category has features of paraganglia. Paraganglia are widely dispersed collections of specialized neural crest cells that arise in association with the segmental or collateral autonomic ganglia throughout the body. Histologically, they have features of neuroendocrine cells.

Owing to the small size of the PNS, metastatic tumours limited strictly to the PNS are extremely rare, although they may occur theoretically. Direct invasion of peripheral nerve by malignant tumour is common and peripheral nerves are preferentially invaded by some primary tumours such as prostate carcinoma and adenoid cystic carcinoma of the salivary gland.

EPIDEMIOLOGY

The overall incidence of primary brain tumours is around 11.8 per 100 000 persons per year. Congenital brain tumours are rare. With the exception of meningioma that affects women twice as often as men, they are more prevalent in man. In the United States, the incidence of

Table 3 Incidence of CNS tumours

Tumour	Incidence/age
Primitive neuroectodermal tumours of posterior fossa (medulloblastoma)	*Incidence:* About 0.5 per 100 000 persons per year *Age:* Peak incidence during the 1st decade of life, a small number of them are seen in adults
Diffuse astrocytoma	*Incidence:* About 0.2 per 100 000 persons per year *Age:* Mean age of onset is 34 years. Peak incidence between 20 and 45 years of age. About 15% of them are seen in children under 10 years old
Anaplastic astrocytoma	*Incidence:* About 0.5 per 100 000 persons per year *Age:* Peak incidence in the 4th to 5th decade. Uncommon in children
Glioblastoma	*Incidence:* Most common malignant astrocytic tumour and the incidence is about 2.0–3.2 per 100 000 persons per year *Age:* Peak incidence in the 5th to 7th decades. They also comprise about 8.8% of all brain tumours in children
Pilocytic astrocytoma	*Incidence:* Most common glioma in children and 85% of them occur in the cerebellum. The incidence is about 0.3 per 100 000 persons per year *Age:* Peak incidence in the first two decades
Pleomorphic xanthoastrocytoma	*Incidence:* Exact incidence is not known. They account for less than 1% of all astrocytic tumours *Age:* Usually seen in young adults with peak incidence in the 2nd decade
Oligodendroglioma	*Incidence:* Usually seen in adults, although children are also affected. Peak incidence in the 5th to 6th decade. The incidence is about 0.3 per 100 000 persons per year *Age:* The peak incidence is in the 4th to 6th decades
Ependymoma	*Incidence:* They account for about 3–9% of all neuroepithelial tumours and are frequently seen in the spinal cord. The incidence is about 0.2–0.3 per 100 000 persons per year *Age:* The first peak of incidence is in the 1st decade and second peak of incidence is in the 4th decade
Choroid plexus tumour	*Incidence:* They account for about 0.4–0.6% of all brain tumours, about 2–4% of all brain tumours in children *Age:* Peak incidence in the 1st decade
Ganglioglioma and gangliocytoma	*Incidence:* They represent about 0.4% of all CNS tumours and 1.3% of all brain tumours *Age:* They can be seen in all ages but are most common in children and young adults
Meningioma	*Incidence:* The annual incidence is about 2.8–6.0 per 100 000 persons per year. They account for 13–26% of primary intracranial tumours *Age:* Most common in middle-aged and elderly patients, peak incidence in the 6th decade
Malignant lymphoma	*Incidence:* Relatively uncommon in the general population, marked increase in incidence in immunocompromised patients *Age:* For immunocompetent patients, the peak incidence is in the 7th decade. Occur in young patients who are immunocompromised
Germ cell tumours	*Incidence:* The incidence is less than 0.1 per 100 000 population per year. They account for about 0.3–2% of all primary intracranial tumours. The incidence is lower in the West and higher in Asia *Age:* Most common in the first two decades
Craniopharyngioma	*Incidence:* 0.05–0.25 per 100 000 persons per year *Age:* Most common in children and a second peak in the 6th decade
Schwannoma	*Incidence:* They are most commonly seen in peripheral nerves. About 8% of all intracranial tumours and 29% of all spinal tumours are schwannomas *Age:* Peak incidence in the 4th to 6th decades
Neuroblastic tumours of adrenal gland and sympathetic system	*Incidence:* 0.7–1.0 per 100 000 persons per year *Age:* 85% occur before the age of 5 years

Source: Kleihues and Cavenee, 2000, *World Health Organisation Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System* (International Agency for Research on Cancer, Lyon).

brain tumours, with the exception of meningioma and craniopharyngioma, has been consistently higher in whites than in blacks. The overall incidence of brain tumours has a bimodal pattern. The first peak occurs in infants and children and the second peak occurs at the sixth to seventh decade. Each individual tumour, however, has its own pattern of age distribution (**Table 3**). Patients affected by hereditary cancer syndromes may develop neoplasms at an earlier age than sporadic cases. There is also a general tendency for high-grade glial tumours to occur in older

patients. There is an apparent increase in the incidence of brain tumours both in the United States and in other countries. However, the increase in the early 1980s may partly reflect improved detection techniques, mainly the use of magnetic resonance image (MRI) and computed tomography (CT) scans. There also appears to be an increase in incidence of childhood brain tumours, most apparent in children under 5 years of age. Increase in the rate of primary CNS lymphomas is dramatic and correlates with the large population of patients with compromised immune system secondary to immune suppression, most often after solid organ transplantation, and human immunodeficiency virus (HIV). Some countries and geographic areas have higher incidences than others. The incidence of a particular tumour may be unusually high in some country. For example, germ cell tumour and craniopharyngioma are far more common in Japan than elsewhere. The overall survival of primary brain tumours varies significantly among the different histological types.

AETIOLOGY

Numerous epidemiological studies have been conducted in the search for risk factors and aetiological agents for brain tumours. Mutations of specific genes have been strongly associated with the development of brain tumours. Some of them are associated with hereditary cancer syndromes that involve the nervous system, others are found with increased frequency in some types of tumours. These genes will be discussed later.

It has been reported repeatedly that virus or virus-like particles have been observed in human cerebral tumours or tumour cell lines. Between 1955 and 1962, poliomyelitis vaccines were contaminated with Simian virus 40 (SV40), a member of the polyoma virus family. SV40 has been used to generate experimental choroid plexus tumour and primitive neuroectodermal tumours in transgenic mice. In addition, a partial genome of SV40 has been identified in a few ependymomas, meningiomas and primitive neuroectodermal tumours (also known as medulloblastomas). Despite these findings, there is no epidemiological evidence that the contaminated poliomyelitis vaccine is associated with any increase in incidence of primary brain tumours. JC and BK viruses, also members of the polyoma virus family, have been suspected for being tumorigenic in human brain tumours. To this date, however, there is no solid proof. The association of HIV infection and primary CNS lymphoma primarily rests on the destruction of immune tumour surveillance by HIV.

A wide variety of other physical and chemical agents have also been studied but ionizing radiation is the only proven risk factor for glial tumours and meningiomas. Children who receive therapeutic doses of X-rays (1.5 Gy) for treatment of tinea capitis have an overall increased relative risk of 8.4 for brain tumours. A correlation at higher dosages (1.0–6.9 Gy) is even stronger (Ron *et al.*, 1988). Correlation between risk of brain tumours and low-dose X-rays has not been well established, although association between dental X-rays and possible increase in meningiomas has been reported. The tumorigenic effects of electromagnetic fields on brain tumours have been shown in industrial settings but the relationship with weaker electromagnetic fields has not been well established and the results of different studies have been inconsistent.

Numerous other aetiological agents and risk factors, including *N*-nitroso compounds, diet, tobacco, alcohol and occupational factors, have also been studied. No conclusive evidence of their association with brain tumours has been identified.

SCREENING AND PREVENTION

Many organs, such as the vagina, uterine cervix, peripheral blood, skin, oral cavity and gastrointestinal tract, are easily accessible. Cancer screening involves simple procedures and is highly effective. The situation is different for the CNS. Except for germ cell tumours that may secret markers detectable in the serum, there is no reliable marker in the blood for most primary brain tumours. The anatomical complexity and bony encasement of the central nervous system make biopsy-based screening impossible. The only reliable method for detecting an asymptomatic or presymptomatic brain or spinal cord tumour is a high-resolution imaging technique such as MRI. This method, however, is too expensive, time consuming and complicated to provide a practical screening test for the general population.

Early Recognition of Brain Tumours

Early recognition of brain tumours relies on astute clinical evaluation since there are no specific symptoms or signs of brain tumours. However, headache, new-onset seizures, hemiparesis and mental-status abnormalities are among the most common symptoms. Diagnosis is largely dependent on imaging techniques such as MRI or CT scan. These tools have revolutionized the diagnosis of CNS tumours and other intracranial and intraspinal conditions.

CNS tumours, like any intracranial and intraspinal space-occupying lesions, generate two major problems. The first is increased intracranial pressure because of oedema and mass effect caused by the tumour or obstruction of the flow of cerebrospinal fluid, a typical situation when the tumour is located in or around the ventricles. Increased intracranial pressure produces headache, vomiting and papilloedema; occasionally, it may cause paralysis of the sixth cranial nerve (abducens nerve) due to excessive stretching. Recognition of papilloedema,

often an asymptomatic clinical sign, is important for presymptomatic recognition of brain tumours. Brain herniation is a life-threatening condition often associated with a severe increase in intracranial pressure.

Local effects of CNS tumours produce symptoms depending upon locations. For example, cortical lesions are prone to cause seizures. Tumours adjacent to the optic nerve often produce visual symptoms due to compression of the optic nerves and chiasma. Tumour masses compressing the cervical cord often cause hemiplegia or quadriplegia. The interested reader should refer to the Further Reading section for further details on clinical aspects of brain tumours.

High-risk Patients

Two major groups of patients have increased risk of developing brain tumours. The first includes those who have underlying nonhereditary conditions that predispose them to develop brain tumours, specifically patients with a primary carcinoma or melanoma in some other part of the body. Immune compromised patients also have an increased risk of developing primary CNS lymphoma.

Patients with hereditary syndromes that predispose them to develop tumours of the CNS comprise the second group. Typically, a variety of tumours from several organs occur in a syndromic manner. Recent advances in genetics and molecular biology have identified the genetic abnormalities in many of these syndromes (**Table 4**), and their molecular mechanisms will be discussed later. Well-structured periodic screening programmes have been invaluable in detecting and treating tumours in early stages. With these screening and treatment programmes, the expectancy and quality of life have been improved in these patients.

GROSS/HISTOPATHOLOGY/PREINVASIVE LESIONS/ULTRASTRUCTURE/ IMMUNOHISTOCHEMISTRY

Tumours of the CNS

Pathological features of brain tumours are diverse and fascinating. Pathological examination for diagnostic purposes must include a thorough macroscopic and microscopic evaluation that often includes utilization of immunohistochemical techniques. Occasionally, ultrastructural examination may be necessary. Molecular biology and/or cytogenetic studies may also provide crucial information for diagnostic and prognostic purposes in some cases. Although a full discussion of immunohistochemistry in CNS tumour is beyond the scope of this chapter, the reader should be familiar with a few commonly used diagnostic molecular markers such as Ki-67, synaptophysin, neurofilament proteins (NFPs), glial fibrillary acidic proteins (GFAPs), vimentin, S-antigen and epithelial membrane antigen (EMA) (**Table 5**).

Primary Tumours of Primitive Neuroepithelial Tissue

This is a family of CNS tumours, two of which recapitulate the phenotypic features of the developing nervous system, i.e. medulloepithelioma, primitive neuroectodermal tumours (PNETs) and related tumours. The third tumour in this group, atypical teratoid/rhabdoid tumour, displays unique features and a remarkably consistent cytogenetic abnormality.

Medulloepithelioma

Medulloepithelioma is a tumour that recapitulates the neural tube stage of the developing CNS. These tumours are rare and predominantly seen in infancy and early childhood. Over half arise in the lateral ventricles, occasionally in the posterior fossa and other locations. Macroscopically, they are well-defined, often massive tumours with extensive haemorrhage and necrosis, and are typically associated with extensive cerebrospinal dissemination. The microscopic appearance is highly distinctive and composed of tubular or papillary structures that are morphologically and immunohistochemically similar to the primitive epithelium (**Figure 1**) of the neural tube. Cytologically, they have hyperchromatic nuclei and a high nuclear-to-cytoplasmic ratio. Sheets of patternless undifferentiated tumour cells are invariably present. Ultrastructural examination of these areas also discloses extensive primitive lateral cell junctions and basal lamina, both features of the neural tube. The tumour cells may also differentiate along neuronal, astrocytric, ependymal, oligodendroglial and, on rare occasions, mesenchymal lineage.

Primitive Neuroectodermal Tumours (PNETs) and Related Tumours

This is a family of tumours that phenotypically resembles the primitive neuroectoderm of developing CNS. PNET of the posterior fossa, also known as medulloblastoma, is the prototype and such tumours can also occur in any CNS locations. PNET is the most common malignant childhood brain tumour but sometimes they are seen in adults, particularly young adults. With adequate treatment, the 5-year survival rate is about 60–80% (Packer *et al.*, 1999). PNETs arising in the posterior fossa typically present in midline and involve the vermis. They project into the fourth ventricle and often extend into the leptomeninges (**Figure 2**). The strategic location in the posterior fossa often leads to hydrocephalus (i.e. pathological expansion of the ventricles). Macroscopically they often appear as soft to firm masses with well-demarcated margins. The cut surface has a pale pink to tan colour and there may be a variable amount of haemorrhage. PNETs are typically composed of

Table 4 Hereditary tumour syndromes

Syndrome	Pathology	Genetics	Reference
Tuberous sclerosis	*Nervous system:* Cortical hamartomas (tuber), subependymal hamartoma and subependymal giant cell tumour *Extraneural:* Adenoma sebaceum and other manifestations of skin, retinal astrocytoma, renal angiolipoma, cardiac rhabdomyoma and other systemic manifestations	*Inheritance:* Autosomal dominant *Prevalence:* Between 1 in 5 000 and 1 00 000 Gene: *TSC1* gene (tuberin) on chromosome 9q34 and *TSC2* gene on chromosome 16p13.3.	Kandt *et al*. (1992); van Slegtenhorst *et al*. (1997)
Neurofibromatosis type 1	*Nervous system:* Neurofibromas and malignant peripheral nerve sheath tumour of the peripheral nerve, gliomas of the brain *Extraneural:* Multiple café-au-lait spots, rhabdomyosarcoma, phaeochromocytoma, carcinoid tumour, juvenile chronic myeloid leukaemia, bone lesions and other manifestations	*Inheritance:* Autosomal dominant *Prevalence:* 1 in 3000–4000 of the general population *Gene: NF1* gene (neurofibromin) on chromosome 17q2	Seizinger *et al*. (1987); Pollack and Mulvihill (1997); Von Deimling, *et al*. (2000)
Neurofibromatosis type 2	*Nervous system:* Bilateral vestibular schwannomas, peripheral schwannomas, meningiomas and meningioangiomatosis, ependymomas, astrocytomas, glial hamartoma, and cerebral calcifications *Extraneural:* Posterior lens opacity	*Inheritance:* Autosomal dominant *Prevalence:* 1 in 50 000 of the general population *Gene: NF2* gene (merlin) on chromosome 22q12	Trofatter *et al*. (1993); Pollack and Mulvihill (1997)
Von Hippel–Lindau disease	*Nervous system:* Haemangioblastoma of the retina and CNS *Extraneural:* Renal cysts and renal cell carcinoma, pancreatic cysts, islet cell tumours, phaeochromocytoma, and other manifestations	*Inheritance:* Autosomal dominant *Prevalence:* 1 in 36 000 to 1 in 45 500 of the general population *Gene: VHL* gene is located on chromosome 3p25.3	Latif *et al*. (1993); Maddock *et al*. (1996)
Naevoid basal cell carcinoma syndrome (Gorlin syndrome)	*Nervous system:* PNET in the posterior fossa *Extraneural:* Multiple basal cell carcinoma and keratocyst of the jaw. Abnormal ribs and other skeletal abnormalities, epidermal cyst, ovarian cysts and other features	*Inheritance:* Autosomal dominant *Incidence:* 1 in 57 000 of the general population *Gene:* Human homologue of the *Drosophilia* segment polarity gene patched (*PTCH*) on chromosome 9q22.3	Hahn *et al*. (1996); Vorechovsky *et al*. (1999)

Cowden disease	*Nervous system:* Dysplastic gangliocytoma of the cerebellum (Lhermitte–Duclos disease). Other pathologic changes include megalencephaly and heterotopic grey matter. Meningiomas and medulloblastomas have also been described *Extraneural:* Verrucous skin changes, papules and fibromas of oral mucosa, multiple facial trichilemmomas, hamartomas polyps of the colon, thyroid tumour and breast cancer	*Inheritance:* Autosomal dominant *Gene: PTEN/MMAC1* gene on chromosome 10q23.	Sutphen *et al.* (1999); Robinson and Cohen (2000)
Turcot syndrome (type 1)	*Nervous system:* Usually glioblastoma *Extraneural:* Café-au-lait spots. Small number of large colorectal polyps and high incidence of colorectal carcinoma. Some patients are associated with hereditary non-polyposis colorectal carcinoma syndrome (HNPCC)	*Inheritance:* Autosomal dominant *Gene:* Several genes involved in mismatch repair including *hMLH1* at chromosome 3p21, *hMLH2* at 2p16, *hMSH3* at 5q11-q13, *hMSH6/GTBP* at 2p16, *hPMS1* at 2q32 and *hPMS2* at 7p22	Paraf *et al.* (1997); Cavenee *et al.* (2000)
Turcot syndrome (type 2)	*Nervous system:* Usually PNET in the posterior fossa *Extraneural:* Associated with familial adenomatous polyposis syndrome (FAP). Patient has innumerable adenomatous colorectal polyposis and high incidence of colorectal carcinoma	*Inheritance:* Autosomal dominant *Gene: APC* gene on chromosome 5q21 that is associated with familial adenomatous polyposis syndrome (FAP)	Hamilton *et al.* (1995); Paraf *et al.* (1997); Cavenee *et al.* (2000)
Retinoblastoma (RB) gene deletion syndrome	*Nervous system:* Retinoblastoma in the retina with or without PNET in the pineal gland (pineoblastoma) *Extraneural:* Increased incidence of second malignancy, multiple congenital abnormalities and mental retardation	*Inheritance:* Autosomal dominant *Gene: RB1* gene on chromosome 13q14.2	Sopta *et al.* (1992); Pratt *et al.* (1994)
Li–Fraumeni syndrome	*Nervous system:* Astrocytic tumour, oligodendroglioma, PNET within and outside the posterior fossa, and choroid plexus tumour *Extraneural:* Tumour in various organs including the breast, lung, stomach, colon, pancreas, skin and others	*Inheritance:* Autosomal dominant *Gene: p53* on chromosome 17p13	Malkin *et al.* (1990); Tachibana *et al.* (2000)

Source: Kleihues and Cavenee, 2000, *World Health Organisation Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System* (International Agency for Research on Cancer, Lyon).

Table 5 Diagnostic immunohistochemical markers for brain tumours

Protein	Properties	Application
Ki-67	Non-histone proteins that are present in all phases of the cell cycle, except G0 and the early G0–G1 transitional phase	Immunohistochemical recognition of Ki-67 is used to estimate the proportion of proliferating cells in brain tumours
Glial fibrillary acidic protein (GFAP)	This is a class III intermediate filament of 55 kDa. GFAP is expressed in mature astrocytes. It may co-express with other intermediate filaments in developing CNS	Expression of GFAP in a brain tumour is highly suggestive but not an absolute indicator of astrocytic differentiation
Synaptophysin	Synaptophysin is an acidic, *N*-glycosylated integral membrane glycoprotein of 38–42 kDa that is expressed in the CNS	Detection of neuroendocrine differentiation in CNS tumours
Chromogranin A	This is a 70–80-kDa protein that is found in neurosecretory granules. It is expressed in neuroendocrine cells outside the CNS	Detection of neuroendocrine differentiation in PNS tumours
Neurofilament proteins (NEPs)	Neurofilament proteins are class IV intermediate filaments that exist in three different isoforms (low, medium and high molecular mass) and different phosphorylation state	Expression of neurofilament proteins is highly suggestive but not an absolute indicator of astrocytic differentiation
Vimentin	This is class III intermediate filament of 56 kDa. It is widely expressed in the developing nervous system. It is not expressed in most mature neurons. It is co-expressed with GFAP in mature and developing astrocytes	Vimentin is strongly and extensively expressed by atypical teratoid/rhabdoid tumours. It is also variably expressed by many neuroepithelial tumours
S-antigen and rhodopsin	S-antigen and rhodopsin are proteins expressed by photoreceptor cells in the retina	Detection of photoreceptor differentiation in PNETs, retinoblastomas and pineal tumours
S-100	S-100 is an acidic protein that is widely expressed in the CNS, PNS and a variety of cells outside the nervous system	It is used in the diagnosis of tumours of the PNS and soft tissue
Cytokeratins	Cytokeratins are proteins of class I and II intermediate filaments. They are predominantly expressed by epithelial cells	Used in diagnosis of tumours with epithelial differentiation such as metastatic carcinomas and craniopharyngioma
Epithelial membrane antigen (EMA)	This is a group of high molecular mass molecules with a high carbohydrate content. They are widely expressed by many epithelial cells	Used in diagnosis of metastatic carcinomas. Some CNS tumours such as meningiomas, ependymomas and atypical teratoid/rhabdoid tumours also express EMA

solid sheets of small cells with an extremely high nuclear-to-cytoplasmic ratio and small hyperchromatic nuclei. The tumour cells inconstantly form Homer Wright rosettes (neuroblastic rosettes), but may arrange themselves in other patterns (**Figure 3**). Individual cell and small areas of necrosis are common but extensive necrosis is uncommon. The most common genetic aberration in PNETs arising from the posterior fossa is isochromosome 17 but other chromosomal abnormalities may also be found. Spinal leptomeningeal dissemination is common.

Desmoplastic medulloblastoma is a variant of PNET that tends to occur in cerebellar hemispheres rather than midline. They are also found in older patients. Histologically these tumours contain nodular, reticulin-free islands ('pale-islands') of tumour cells surrounded by densely packed PNET cells with dense intercellular reticulin deposition. The pale islands are usually less cellular and more differentiated than the surrounding cells. Medulloblastoma with extensive nodularity and advanced neuronal differentiation is a variant that is associated with favourable prognosis and contains intranodular cells that resemble mature neurons. Large cell medulloblastoma is a variant in which the tumour cells have large, round and pleomorphic nuclei, prominent nucleoli and more obvious cytoplasm. They usually carry a grave prognosis.

PNETs typically contain numerous mitoses and apoptotic bodies. Depending on the studies, the proliferation rate based on Ki-67 labelling may range from 5 to 80% (**Figure 4**). Such a high proliferation rate has been confirmed by other labelling method such as iododeoxyuridine that discloses a range of 3.9–38.2% (Onda *et al*., 1996).

Immunohistochemically, PNETs frequently express synaptophysin, a marker indicative of neuroendocrine differentiation. Many PNETs also express nestin, a marker for primitive neuroepithelial cells. Although PNET cells appear histologically similar, they may express different developmentally regulated antigens that recapitulate different developmental stages of the CNS. While some PNETs may

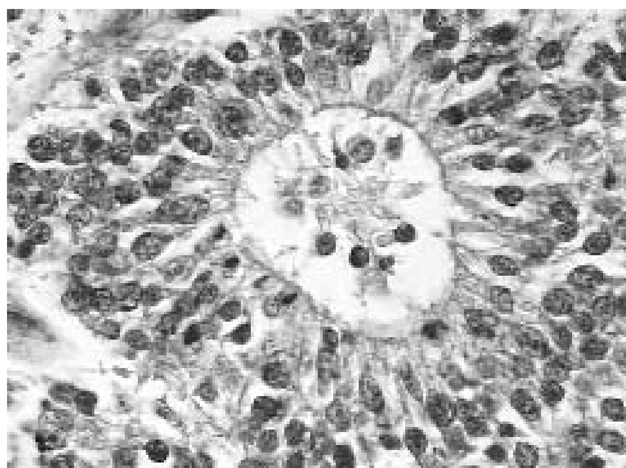


Figure 1 Medulloepitheliomas are characterized by small blue cells that arrange in glandular or tubular structures reminiscent of the primitive neuroepithelium. (Case courtesy of Dr Lucy B. Rorke, Children's Hospital of Philadelphia.)

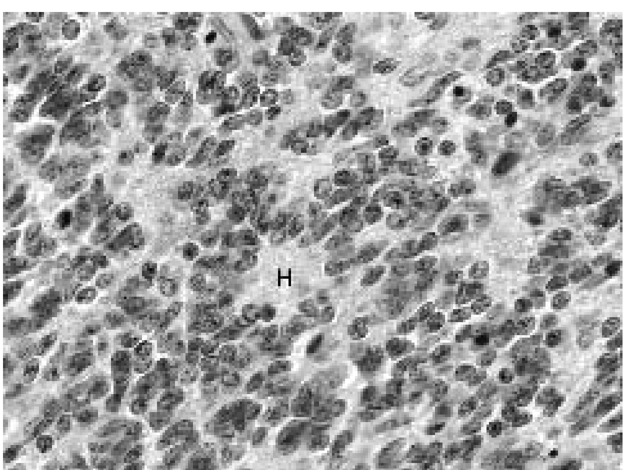


Figure 3 Homer Wright rosettes are formed by small islands of neoplastic neuropil surrounded by tumour cells. Same case as in **Figure 4**.

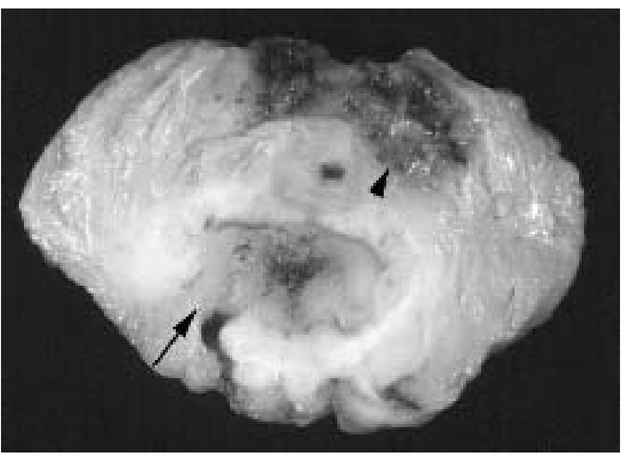


Figure 2 The fourth ventricle (arrow) is filled and expanded by this primitive neuroectodermal tumour. The haemorrhagic region (arrowhead) is in the leptomeningeal extension of the lesion.

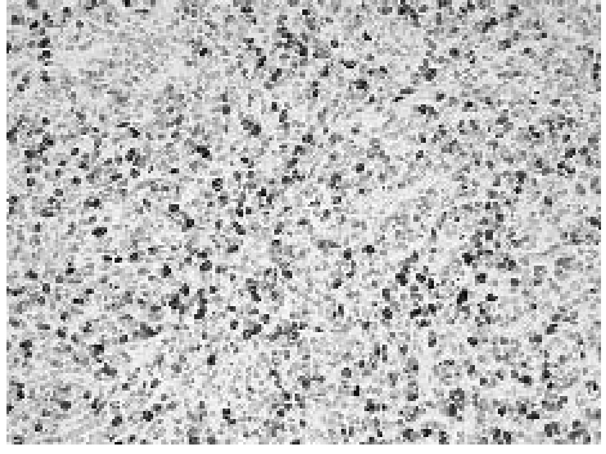


Figure 4 Proliferating cells express Ki67 strongly and comprise about 25% of the population in this primitive neuroectodermal tumour. Immunostaining for Ki-67, ABC peroxidase method.

express NFP indicative of neuronal differentiation, others may express GFAP indicative of glial differentiation and many PNETs express both. A small proportion of PNETs also express cytokeratins, intermediate filament proteins that are expressed in the very early stage of neural development. Ependymal differentiation can also be seen, in which case some investigators prefer to call them ependymoblastomas. Photoreceptor-specific proteins such as retinal S-antigen and rhodopsin are also expressed by a subpopulation of PNETs. Rarely, they express desmin or actually contain skeletal muscle cells in which case they are termed medullomyoblastoma. Melanotic medulloblastoma is the variant that contains melanin pigment. It is obvious that PNETs have the potential to differentiate along different lineages.

Ultrastructually, tumour cells may contain few, if any, specific ultrastructural features. Neurosecretary granules and synapses, an indication of differentiation towards neuronal lineage, may be seen in some tumours. Intermediate filaments are also common.

Less commonly, PNETs occur in supratentorial locations. In the WHO Classification, neuroblastoma and ganglioneuroblastoma are grouped under the heading of supratentorial PNET. These tumours are seen in early childhood. They occur most frequently in the frontal and frontoparietal regions but rare examples are seen in the spinal cord. Macroscopically, they are well-defined tumours. Histologically and immunohistochemically, they are very similar to if not identical to PNETs arising

in the posterior fossa (medulloblastoma). Although infratentorial and supratentorial PNETs are given different names in the WHO Classification (i.e. medulloblastoma vs neuroblastoma), they are best regarded as part of the spectrum of PNETs.

Atypical Teratoid/Rhabdoid Tumour (ATRT)

This is a highly malignant embryonal CNS tumour that contains rhabdoid cells (**Figure 5**). They are predominantly seen in infants under 2 years of age and the mean post-operative survival is less than 1 year. ATRTs are equally frequent in both infratentorial and supratentorial locations. Dissemination throughout the CNS is seen in about one-third of the cases at initial presentation. Macroscopically, they are well-demarcated tumours. Rhabdoid cells are round to oval cells with medium to large nuclei and distinct nucleoli. About two-thirds of ATRTs also have a PNET component. In addition, they may contain epithelial (carcinomatous) and mesenchymal (sarcomatous) components. The tumour cells typically express vimentin, smooth muscle actin and epithelial membrane antigen, but may also express GFAP and NFP. They do not express germ cell markers such as α-fetoprotein, placental alkaline phosphatase or β-human chorionic gonadotrophin, nor do they express desmin. Ultrastructually, they are characterized by massive whorls of intermediate filaments. Characteristically, they show homozygous or heterozygous, complete or partial deletion of chromosome 22q and presumably are related to germline and somatic mutation of the *hSNF5/INI1* gene (Biegel *et al.*, 2000). Although ATRTs often contain a PNET component, their distinctive histology, characteristic chromosomal abnormality and aggressive clinical behaviour distinguish them from PNETs. The true nature of ATRT remains to be determined.

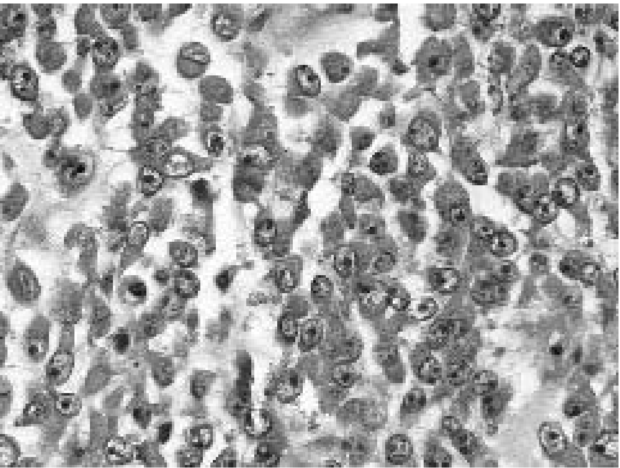


Figure 5 Rhabdoid cells in atypical teratoid/rhabdoid tumours are loosely arranged large egg-shaped cells with eosinophilic cytoplasm, large eccentrically located nuclei containing prominent nucleoli. (Case courtesy of Dr Lucy B. Rorke, Children's Hospital of Philadelphia.)

Primary Tumours of Mature Neuroepithelium

Tumours considered in this section are the most common primary tumours in the CNS. They display features of mature neuroepithelial cells although they do not necessarily arise from such cells. Rather, they may arise from primitive neuroepithelial cells that show differentiation during tumorigenesis. Three major categories will be considered here: glial tumours, glial–neuronal tumours and choroid plexus tumours.

Glial Tumours

Glial tumours, otherwise known as gliomas, are tumours that display features of mature glial cells specifically astrocytes, oligodendrogliocytes and ependymal cells. It is not unusual to have tumours composed of a mixture of these cells. In addition, they may also display a spectrum of biological aggressiveness, often reflected by their histology. In general, low-grade gliomas occur more commonly in younger patients whereas high-grade gliomas are more common in older adults.

Astrocytoma, Anaplastic Astrocytoma, Glioblastoma. Astrocytic tumours are composed of cells that look like astrocytes and are by far the most common of all gliomas. They comprise about one-third of the gliomas in the cerebral hemispheres. Astrocytic tumours are separated into different histological grades that reflect their biological behaviour. Astrocytomas are histologically benign/low-grade tumours, anaplastic astrocytoma has features of a malignant tumour, and glioblastoma is the most malignant version.

Astrocytoma occurs in young adults with a peak incidence in the fourth decade and 15% are seen in children. They are most commonly located in the cerebral hemispheres with frontal and temporal regions as the most common site. They are also common in the spinal cord. In children, the optic nerve, diencephalon, brainstem and cerebellum are the most favoured locations.

Astrocytomas in adults and children behave differently. In adults, the survival rate of astrocytoma is variable and they are notorious for recurring as anaplastic astrocytoma or glioblastoma with a mean time interval of 4–5 years. This occurs more commonly in older patients.

MRI often discloses a poorly defined nonenhancing lesion with oedema. Macroscopically, they are poorly demarcated, soft, grey, expanding lesions that involve mainly the white matter (**Figure 6**). Tumours that arise from the brainstem often diffusely enlarge the latter. Hemispherical astrocytomas may involve the basal ganglia and thalamus. Degenerative changes and cyst formation are common. Histologically, they are composed of neoplastic astrocytes with variable morphology. The architecture ranges from streaming bundles of fibrillary astrocytes to a meshwork-like arrangement. Most often, the tumour cells have poorly defined cytoplasm and small and mature-looking nuclei (**Figure 7**). Some astrocytomas are composed of

fibrillary astrocytes or polygonal to stellate neoplastic cells that arrange themselves into a fine cobweb resembling protoplasmic astrocytes. Gemistocystic forms which tend to be large oval with eosinophilic homogeneous cytoplasm and large nuclei, often binucleated, may also be seen. Often, all morphological subtypes of neoplastic astrocytes can be seen in the same tumour. Degenerative changes and formation of microcysts are very common. Features indicative of aggressive behaviour such as necrosis and mitoses are not present. The low proliferative rate is also reflected by a low Ki-67 labelling index of less than 4% in most cases. Astrocytes strongly express GFAP. These tumours also express other antigens such as S-100 proteins and vimentin, but these markers are not diagnostically useful. Onset at a young age and location at a site that allows gross total resection are two favourable prognostic indicators.

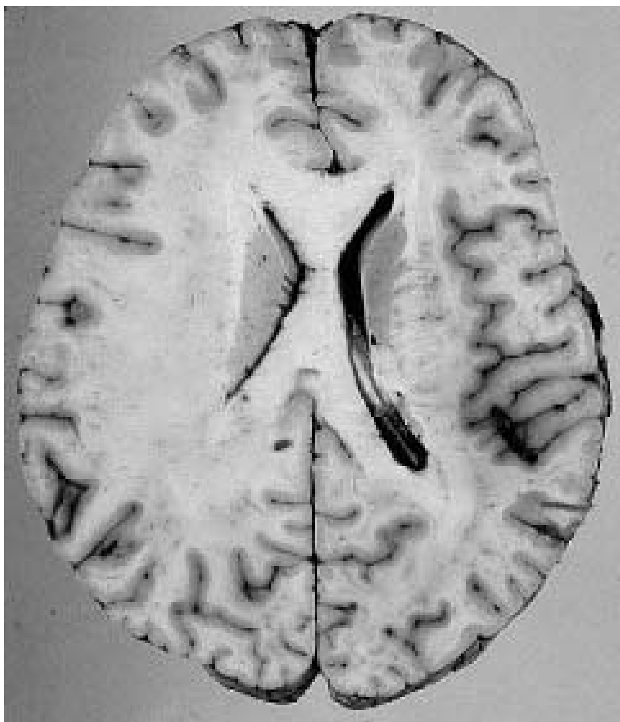


Figure 6 The left side of the brain is expanded by a diffusely infiltrating astrocytoma that also compresses the ventricle. The cut surface of the tumour is similar to normal white matter and no delineation between tumour and normal tissue in appearance. Attenuation of the gyral architecture, however, is obvious.

In contrast to astrocytomas occurring in adults, astrocytoma arising from cerebellum and midbrain in children have excellent prognosis. Childhood astrocytomas arising from pons often carry a grave prognosis.

Anaplastic astrocytomas are aggressive tumours with a peak age of occurrence about 10 years later than that of astrocytomas. MRI usually discloses an enhancing mass with its epicentre in the white matter. These tumours usually enhance with gadolinium. Macroscopically, the tumours have poorly defined borders and are friable, greyish and occasionally haemorrhagic. They are densely cellular, exhibit nuclear pleomorphism, are mitotically active (**Figure 8**) and show endothelial proliferation (**Figure 9**). In some cases, the bulk of the tumour may appear to be a low-grade astrocytoma but microscopic foci of anaplasia may be scattered around. GFAP is usually widely expressed by tumour cells. The brisk proliferative activity is well reflected by the presence of mitotic figures and a Ki-67 labelling index that is usually >5%.

Glioblastoma is the most aggressive astrocytic tumour and unfortunately the most common glial tumour in adults. They also occur in children. The peak incidence is around the sixth and seventh decade and they are 1.5 times more common in men than women. Glioblastoma may be present initially as such or evolve from low-grade glial

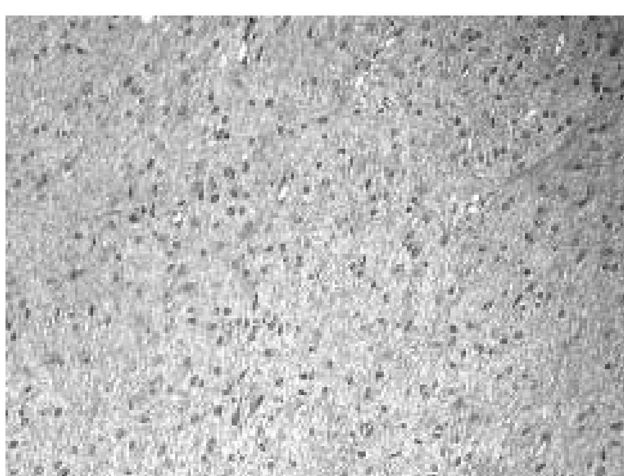


Figure 7 Astrocytoma cells have small hyperchromatic nuclei and a substantial amount of cytoplasm.

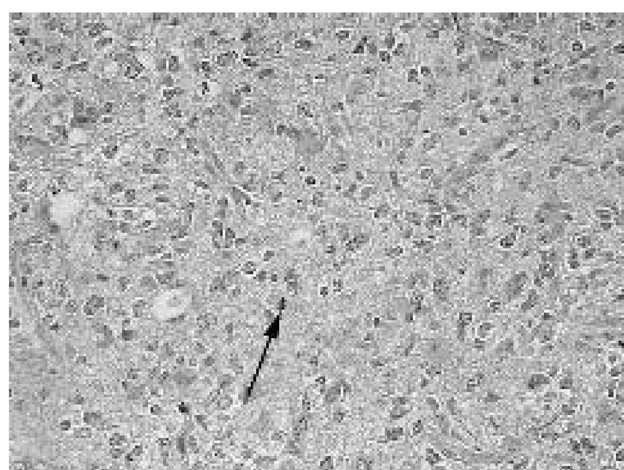


Figure 8 When compared with astrocytomas, cells in anaplastic astrocytocytomas are larger, more pleomorphic and mitotically active (arrow).

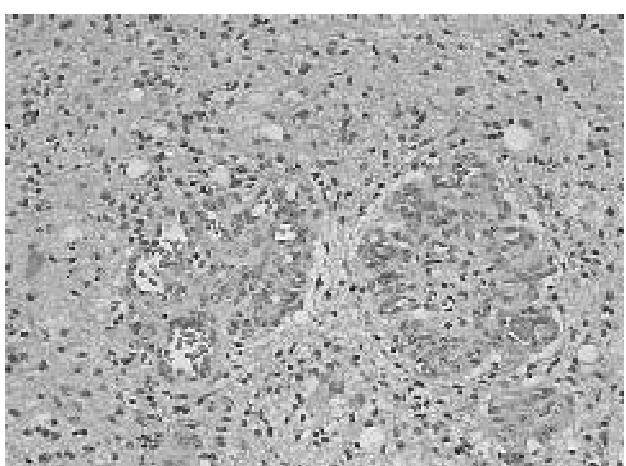


Figure 9 Endothelial proliferation is a hallmark of anaplastic astrocytoma and glioblastoma multiforme. Note glomeruloid architecture of the blood vessels with large and tombstone-like endothelial cells projecting into the lumen.

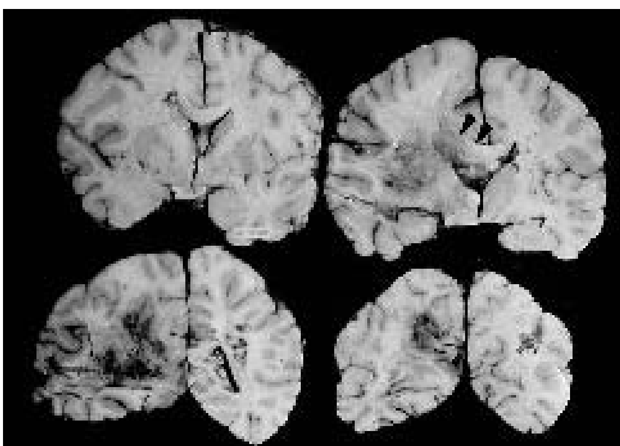


Figure 10 Four coronal sections of cerebrum showing large, poorly demarcated partially haemorrhagic glioblastoma multiforme. Note thickening of the corpus callosum due to tumour infiltration (arrow head).

tumours such as astrocytoma, oligodendroglioma or mixed astrocytoma–oligodendroglioma. The frontal half of the brain is more affected. The tumour may appear as an irregular to multinodular mass with heterogeneous signal and ring enhancement on MRI. They are often associated with a substantial amount of oedema.

Macroscopically, they are typically large hemispherical tumours that extend across the corpus callosum to involve the contralateral hemispheres, the so-called 'butterfly tumour.' They may be poorly demarcated or sharply demarcated; necrosis and haemorrhages are common (**Figure 10**). Histologically, the cellular morphology is highly diversified, as reflected by old term 'glioblastoma multiforme.' The salient histological features of glioblastoma include pseudopalisading necrosis (**Figure 11**),

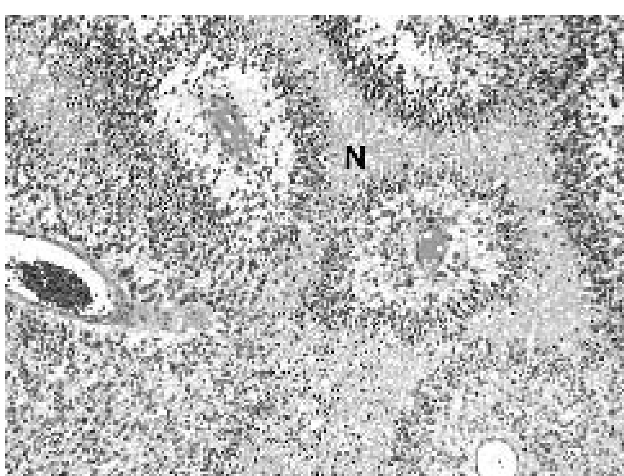


Figure 11 Pseudopalisading necrosis in GBM is composed of necrotic tissue (N) rimmed by viable tumour cells that arrange themselves in a palisading fashion.

striking endothelial proliferation, markedly anaplastic cells and prominent mitotic activity. Bizarre tumour cells and multinucleated giant cells are often seen adjacent to fields of densely packed small anaplastic cells. In general, expression of GFAP is inversely proportional to the degree of anaplasia. Although the majority of glioblastomas express GFAP, it is often expressed in a patchy manner. Ki-67 labelling may vary significantly in different areas of the same tumour and mean values of 15–20% have been reported.

Giant cell glioblastoma is a variant that contains a substantial number of giant cells; it has a slightly better prognosis than regular glioblastoma. Gliosarcoma is a variant that contains a glioblastoma component and sarcomatous component. In rare examples, the latter consist of osteosarcoma or chondrosarcoma.

Astrocytoma with Special Features. Pilocytic astrocytoma, pleomorphic xanthoastrocytoma, desmoplastic infantile astrocytoma and subependymal giant cell astrocytomas are tumours that exhibit distinct clinical and histological features.

Pilocytic astrocytomas are most commonly seen in the first two decades of life. They tend to remain histologically benign, i.e. they do not progress to high-grade gliomas. They are most common in the cerebellum as discrete, slow-growing lesions with or without cystic change. Other preferred sites include optic nerve, optic chiasma/hypothalamus and brainstem. Those that are located in a surgically accessible position can be totally resected and have an excellent prognosis, whereas locations in some sites may preclude such an approach. Hence prognosis in the latter instance is not so sanguine. In contrast to astrocytomas in adults that infiltrate the surrounding brain tissue in a sinister fashion, pilocytic astrocytomas especially in the cerebellum are often well demarcated. Histologically they are characterized by alternating densely packed and loosely packed areas (**Figure 12**). The densely packed

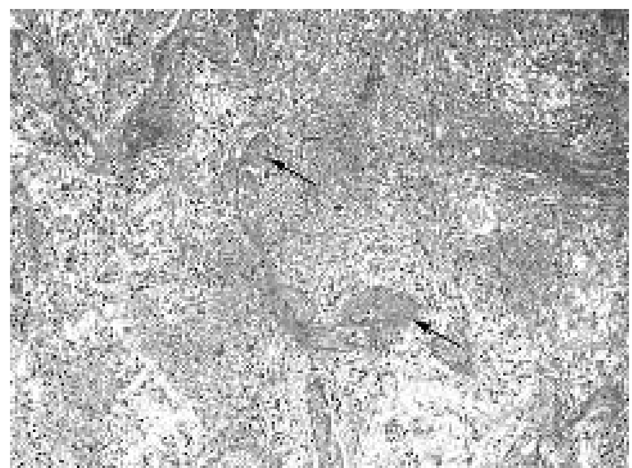


Figure 12 Pilocytic astrocytomas have alternating densely packed and loosely packed areas. Tumour cells in the densely packed areas are reminiscent of bundles of hair. Many thin-walled and dilated blood vessels are typically present (arrow).

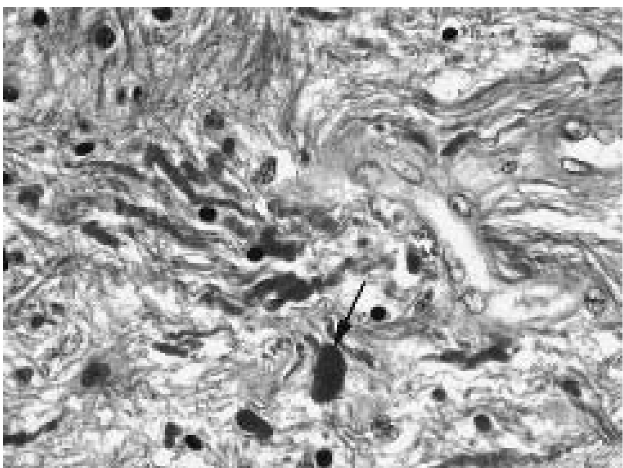


Figure 13 Rosenthal fibres are amorphous to fibrillary hyalinized eosinophilc structures that are intracytoplasmic components of glial processes (arrow). They are typical but not diagnostic for pilocytic astrocytoma.

areas are composed of elongated hair-like (piloid) astrocytes whereas the loosely packed areas are composed of astrocytes with small nuclei and stellate processes that branch and anastomose producing a 'chicken-wire' appearance. Pilocytic astrocytomas are usually vascular and contain peculiar glomeruloids of thin-walled blood vessels reminiscent of wickerwork. They must be distinguished from endothelial proliferation found in anaplastic astrocytoma and glioblastoma. Rosenthal material (**Figure 13**) and eosinophilic granular bodies are common and are most often seen in the densely packed areas. Mucoid and cystic degeneration are common. As expected, the tumour cells strongly express GFAP.

Pleomorphic xanthoastrocytoma and subependymal giant cell astrocytomas are two tumours that are classified as astrocytoma but should probably be in the category of neuronal–glial tumours as they often express both GFAP and NFP that indicate simultaneous glial and neuronal differentiation.

Pleomorphic xanthoastrocytoma is usually seen in children and young adults. They typically occur as supertentorial and superficial tumours that are attached to the leptomeninges and have cystic components. Histologically, they are composed of cells that show extreme variation in size and shape of both nuclei and cytoplasm, hence they are called 'pleomorphic.' Some cells may have foamy or xanthomatous cytoplasm. They contain a dense network of reticulin fibre and multinucleated giant cells are common. Admixed with these cells are also spindly cells. Collections of lymphocytes are often prominent. The large xanthomatous spindle cells express GFAP. Some also express NFPs and synaptophysin. The anaplastic variant exhibits significant mitotic activity and/or necrosis. Most pleomorphic xanthoastrocytomas carry a favourable prognosis but a small proportion of them may progress to high-grade tumours.

Desmoplastic infantile astrocytoma will be discussed later together with desmoplastic infantile ganglioglioma. Subependymal giant cell tumour is almost always associated with tuberous sclerosis and will be discussed together with the latter.

Oligodendroglioma and Anaplastic Oligodendroglioma. Oligodendrogliomas are neoplasms composed of oligodendrocytes and uncommon. Most arise in hemispherical white matter of individuals in the fourth to sixth decades of life. Macroscopically, they display a relatively well-demarcated, soft, greyish pink appearance. Often, mucoid changes in the cells produce a translucent mucoid consistency. Microscopically they are moderately cellular tumours. The cells grow in solid sheets with delicate vascular networks. The cells are round to polygonal cells with a well-defined cytoplasmic membrane, clear cytoplasm and isomorphic round, small nuclei producing the so-called 'fried-egg' appearance (**Figure 14**). Calcified deposits are more common than other gliomas. The proliferation rate is low and mitotic figures are rare, if present at all; the Ki-67 labelling index is <5% (Coons *et al.*, 1997). Oligodendrogliomas may express small amounts of GFAP but no immunohistochemical marker is specific for these tumours. The mean survival time for oligodendrogliomas varies from 3.3 to 7.1 years and the 5-year survival rate varies from 38 to 54% (Dehghani *et al.*, 1998). The biological behaviour is not always predicted by histology and oligodendrogliomas without any anaplastic features can also behave aggressively.

Anaplastic oligodendrogliomas occur in a slightly older age group than do the histologically 'benign' type but the

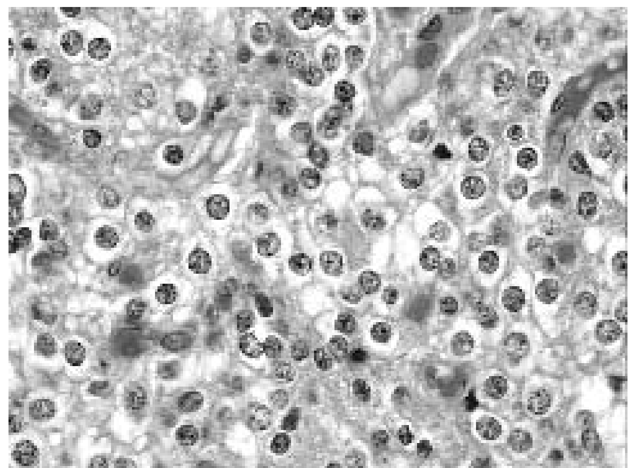


Figure 14 Oligodendrogliomas are characteristically composed of round to polygonal cells with clear cytoplasm and centrally located round nuclei. (Courtesy of Dr Lucy B. Rorke, Children's Hospital of Philadelphia.)

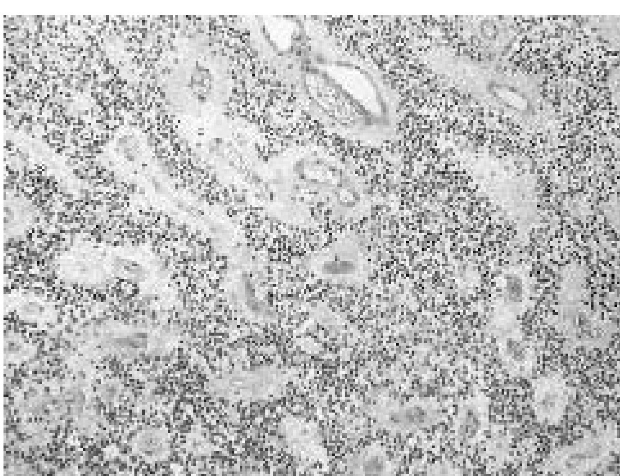


Figure 15 Ependymomas often appear as isomorphic cellular tumours with a rich vascular supply and pseudorosette formation.

preferred locations are similar. The histological features are similar to those of oligodendrogliomas but these tumour cells tend to contain amphophilic rather than clear cytoplasm. There is increased cellularity, a high nuclear-to-cytoplasmic ratio and variation in size and shape of nuclei. Endothelial proliferation and necrosis are often present and mitotic activity is usually brisk. It is not uncommon to find focal anaplastic changes in an otherwise oligodendroglioma, in which case the tumour behaves like an anaplastic oligodendroglioma. The survival rate of patients with anaplastic oligodendroglioma (median survival time 0.87–3.9 years) is overall worse than for its 'benign' counterparts (Shaw *et al*., 1992; Dehghani *et al*., 1998). Recent association of chromosome 1p and 19q deletion with improved responsiveness to chemotherapy in anaplastic oligodendrogliomas may improve survival of these patients (Cairncross *et al*., 1998).

Ependymal Tumours. Ependymal cells are modified glial cells that line the ventricles. Although they have partial features of epithelial cells, they do not have a basement membrane. Ependymal tumours can be segregated into several clinically and pathologically distinct groups on the basis of pathology and clinical behaviour.

Ependymomas are seen in all age groups but most commonly in children, amongst whom they comprise 6–12% of all intracranial tumours in children and about 30% of them occur before the age of 3 years. Ependymomas may occur at any site along the ventricular system and in the spinal canal. In adults they are most common in the spinal cord; in children they are most common in the fourth ventricle. Hydrocephalus and increased intracranial pressure are common manifestations, particularly in tumours that arise in the fourth ventricles.

Ependymomas are grossly well demarcated. Tumours that arise within the fourth ventricle often expand through the foramina of Luschka and Magendie and extend caudally through the foramen magnum, making complete excision difficult. Ependymoma may also have spinal dissemination. They have diverse histological features. Classical ependymomas are composed of densely packed isomorphic cells that have a small amount of cytoplasm (**Figure 15**). The typical architecture is a perivascular distribution of cells forming pseudorosettes (**Figure 16**). Tumour cells have cytoplasmic processes that taper towards the blood vessel wall. Less frequently canals formed by ependymal cells (also known as ependymal

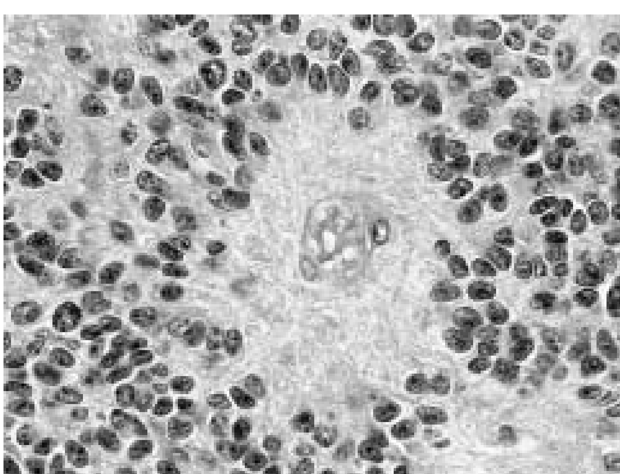


Figure 16 Pseudorosettes, the hallmark of ependymal tumours, are composed of small blood vessels rimmed by a nucleus free mantle of fibrillary cytoplasmic processes that are surrounded by tumour cells.

rosettes) of variable size may predominate in some tumours. A subtype called cellular ependymomas has paucity of pseudorosettes or ependymal canals. Clear cell ependymomas resemble oligodendrogliomas and central neurocytomas. Care must be taken not to confuse them. Papillary ependymoma is a rare variant that is composed of papillae and must be distinguished from choroid plexus papilloma. Tanycytic ependymomas histologically resemble the spindly bipolar tanycytic ependymal cells.

Myxopapillary ependymomas have a distinctive histology and carry an excellent prognosis if completely resected. They can be seen in all age groups but are most common in the fourth decade. These tumours occur almost exclusively in the conus medullaris, cauda equina and filum terminale. Rarely, they can be seen at other levels of the spinal cord or even in sacrococcygeal subcutaneous tissue. These tumours typically have a lobulated, soft, greyish appearance. Histologically, they consist of hyalinized or oedematous fibrovascular cores covered by variable amounts of ependymal cells (**Figure 17**). A mucoid matrix material between tumour cells and blood vessels is typically present. The tumour cells often express GFAP but not cytokeratin, a useful feature to differentiate them from metastatic papillary carcinomas.

Subependymoma is an uncommon, slow-growing, histologically benign periventricular tumour that is composed of cells resembling both ependymal cells and astrocytes. Such tumours are found predominantly in the lateral and fourth ventricles and all age groups are affected. Subendymomas often but not always remain asymptomatic. Clinical manifestations, when present, often result from ventricular obstruction and increased intracranial pressure. Macroscopically, they are sharply demarcated exophytic nodular masses arising from the ventricular wall. Histologically, they display features of both astrocytoma and ependymoma. Sparsely cellular fibrillary astrocytes form a matrix throughout which ependymal cells are unevenly scattered. These cells may form hollow rosettes or perivascular pseudorosettes or may merely consist of small clumps of polygonal cells with no special orientation. Mitotic figures are not usually seen.

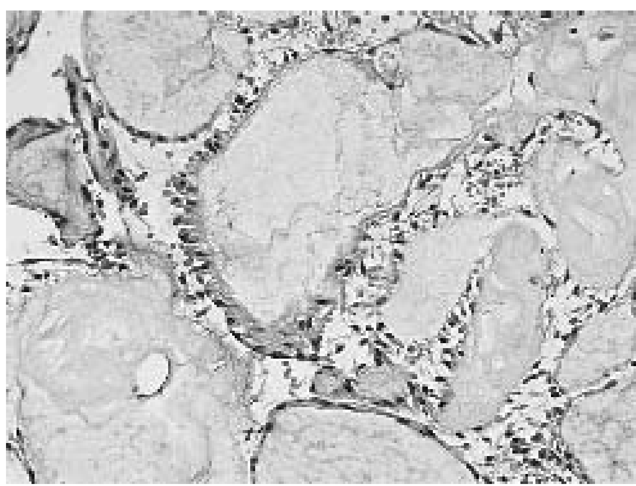


Figure 17 Myxopapillary ependymomas typically contain lakes of myxoid material bounded by a thin layer of cells.

The majority of ependymal tumours express GFAP. EMA is also expressed and most often at the luminal border of the ependymal canal, the internal border of the perivascular pseudorosettes and luminal surface of papillary ependymoma. Cytokeratin is occasionally expressed. The ultrastructural characteristics of normal ependymal cells such as a $9+2$ arrangement of cilia, blepharoblasts and microvilli are well maintained by the tumour cells.

Anaplastic ependymoma is a tumour with histological features suggestive of aggressive behaviour and malignancy. These features included nuclear pleomorphism, high mitotic activity and necrosis.

However, the clinical behaviour of ependymomas is not well correlated with their histological features. In a series of 298 cases, survival of patients could not be correlated with the histopathological features of malignancy (Schiffer *et al*., 1991). On the other hand, Ritter *et al*. (1998) noted that onset at young age, a Ki-67 labelling index of >20% and anaplastic features are associated with aggressive behaviour.

Mixed Gliomas. Some gliomas contain more than one neoplastic component. Oligoastrocytoma, the prototype of mixed glioma, is composed of a substantial mixture of distinct components of astrocytoma and oligodendroglioma. When features indicative of malignant behaviour are present, they are termed anaplastic oligoastrocytoma. The median survival for low-grade mixed oligoastrocytoma is 6.3 years and for high-grade oligoastrocytoma it is 2.8 years (Shaw *et al*., 1994). Other mixed gliomas may contain ependymal and oligodendroglial components, less commonly astrocytic and ependymal components. Rarely, neoplastic components of all three lineages can be found.

Neuronal and Mixed Glial–Neuronal Tumours

Primary neuroepithelial tumours that contain mature neuronal elements are most common in children and young adults. Many are histologically and biologically benign tumours and often have features of hamartomas. While paraganglioma is included in this category of CNS tumour in the WHO Classification, it is best regarded as a neuroendocrine tumour that arises from the nerve roots and hence is discussed in the section 'Tumours of the PNS' (see later).

Ganglioglioma and Gangliocytoma. Ganglioglioma is composed of neoplastic ganglion (neuronal) cells and neoplastic glial cells. Symptoms vary according to location, but intractable epilepsy is the most common clinical manifestation. Indeed, total resolution or significant

improvement for control of epilepsy is often noted following removal of the tumour. They can occur anywhere in CNS but the temporal lobe is the preferred site.

Macroscopically, they consist of a well-circumscribed solid mass or cyst with a mural nodule; they are not usually associated with significant mass effect. Calcification is often present but haemorrhage and necrosis are rare. Neoplastic ganglion cells, the salient feature of these tumours, are moderate to large dysmorphic cells that resemble neurons. They have eccentrically placed, round nuclei containing small but distinct nucleoli; binucleated cells are common. Their cytoplasm is usually basophilic. The ganglion cells may be arranged in clusters in a neuropil-like or glial background.

The amount of glial component is variable. In some tumours there is no significant glial component, in which case the tumour is gangliocytoma (**Figure 18**). Gangliocytomas are almost always benign and can be cured by resection if in an accessible location. Rare anaplastic gangliocytomas have been reported.

Ganglioglioma, a far more common entity than gangliocytoma, is composed of neoplastic ganglion cells admixed with neoplastic astrocytes and enlarged or bizarre-shaped neurons (**Figure 19**). In unusual cases, the glial component may be oligodendroglial rather than astrocytic. Both gangliocytoma and ganglioglioma may contain reticulin around ganglion cells and/or perivascular lymphocytic infiltrations. Neoplastic ganglion cells express NFPs, synaptophysin and other neuronal markers in the neoplastic ganglionic cells, whereas the glial component expresses GFAP. At the ultrastructural level, neurosecretory granules are often seen in the neoplastic neurons; synaptic junctions may or may not be present.

Most gangliogliomas can be cured by resection but a minority of them recur and undergo malignant transformation. Only in exceptional conditions is the product of malignant transformation a high-grade glial neoplasm. Gangliogliomas in the brainstem and spinal cord are 3.5–5 times more likely to recur than cerebral gangliogliomas (Lang *et al*., 1993).

Dysplastic Gangliocytomas of the Cerebellum (Lhermitte–Duclos Disease). This interesting benign cerebellar lesion combines features of hypertrophy, congenital malformation and neoplasm. It is benign and often associated with Cowden disease (see below). Neurological manifestations include ataxia, mild mental retardation and a self-limited seizure disorder. They are most common in young adults and have a marked male preponderance. The lesion is characterized by diffuse hypertrophy of cerebellar folia. Histologically, the normal architecture of cerebellum is effaced. Instead, folia are composed of an outer layer of well-developed radial and more superficially distributed parallel myelinated nerve fibres, and an inner layer of abnormal neurons that contains both small hyperchromatic neurons and large polygonal neurons with prominent nucleoli. The large cells resemble Purkinje cells. The white matter is either absent or greatly reduced. It has been suggested that this is a hamartomatous lesion, a hypothesis supported by its frequent association with other malformations such as megalencephaly and heterotopic grey matter.

Dysembryoplastic Neuroepithelial Tumour (DNET). DNETs are benign glial–neuronal neoplasms that are most commonly seen in children and young adults, often with a history of prolonged drug-resistant seizures. They are superficially located supratentorial tumours that frequently involve the temporal lobe. Macroscopically, they contain a central viscous mass associated with firmer nodules in the periphery. The affected cortex is expanded and foci

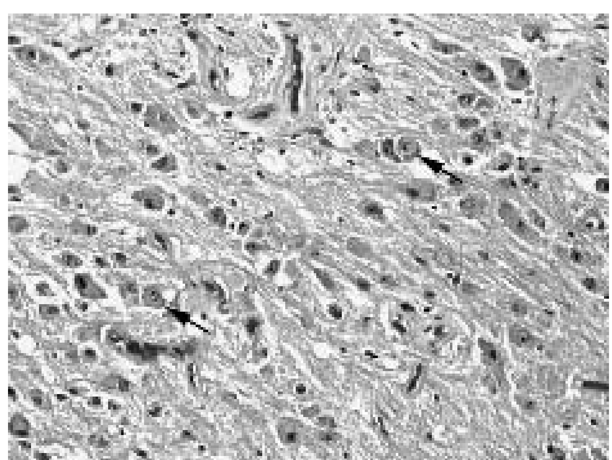


Figure 18 Photomicrograph of gangliocytoma showing tumour composed of abnormal ganglion (arrow) cells scattered through the neuropil.

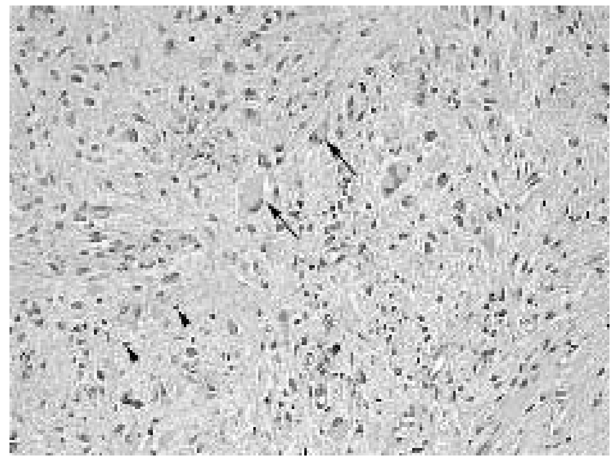


Figure 19 Gangliogliomas have a neuronal (ganglionic) component and an astrocytic component. Ganglionic cells (arrow) are usually much larger than astrocytic cells (arrow head).

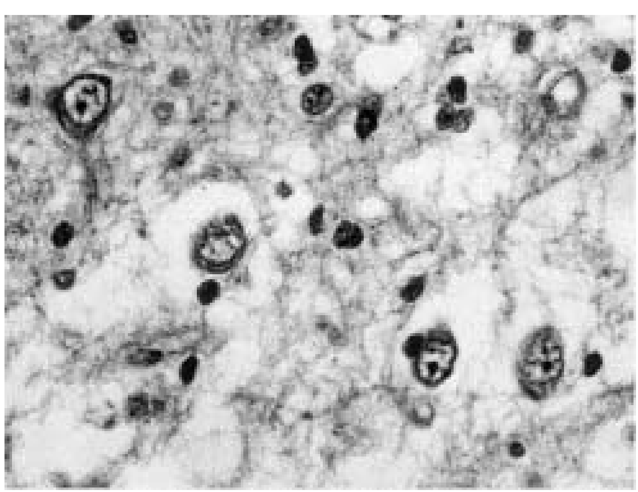


Figure 20 Differences between the neuronal component (large cells) and glial component (small cells) in DNETs are well illustrated here.

of cortical dysplasia (focally disorganized cortex) are invariably seen in the surrounding cortex. Such cortical topography is an important criterion for differentiating between DNETs and other gangliogliomas. Therefore, clinical imaging findings must be seriously examined in conjunction with the histological features. Classical DNETs contain a mucoid component composed of columns of axons perpendicularly oriented to pial surface and lined by small S-100-positive and GFAP-negative oligodendroglia-like cells. Suspended in between these bundles is mucoid material (**Figure 20**) that contains neurons with normal cytology and some GFAP-positive stellate astrocytes. This combination is termed 'specific glioneuronal elements' and is considered the histological hallmark of DNETs (Daumas-Duport, 1993).

Desmoplastic Infantile Ganglioglioma/Astrocytoma (DIG/DIA). Desmoplastic infantile ganglioglioma (DIG) and desmoplastic infantile astrocytoma (DIA) (also known as superficial cerebral astrocytoma or desmoplastic cerebral astrocytoma of infancy) are rare tumours that invariably arise in the supratentorial region and are almost exclusively seen in infants under 2 years of age, although noninfantile cases have been reported. They tend to be large and superficially placed cystic tumours that involve the leptomeninges and are often attached to the dura. The solid part of the tumour is firm or rubbery in consistency and grey or white in colour. Necrosis and haemorrhage are uncommon. They may casue cranial enlargement, tense and bulging fontanelles, downward ocular deviation ('sunset sign') and palsy of the sixth and seventh cranial nerve in infants. Hydrocephalus is common. Despite their large size and some suggestive malignant histological features, prognosis is favourable after surgical resection.

Histologically, DIGs are reticulin-rich astrocytic tumours with a variable neuronal component (DIG). There is an often a sharp demarcation between the cortical surface and tumour. Neoplastic ganglion cells and astrocytes are found within a desmoplastic background. Reticulin stain characteristically discloses deposition of reticulin substance around individual tumour cells. Aggregates of poorly differentiated neuroepithelial cells are often found. The glial cell express GFAP and vimentin; the neuronal component expresses synaptophysin, NFPs and class III β-tubulin. The identity of the poorly differentiated cells is not well defined but is most likely primitive neuroepithelial cells. Mitotic activity in the differentiated component is usually rare, however, the poorly differentiated cases may display mitotic activity. DIA is a reticulin-rich desmoplastic astrocytic tumour with no neuronal component.

Central Neurocytoma. Central neurocytomas are typically seen in young adults with peak incidence in the third decade. They are uncommon tumours that are usually found in the lateral ventricles and/or third ventricle, a feature that often leads to increased intracranial pressure and hydrocephalus. The tumours are greyish and friable and may be calcified and occasionally haemorrhagic. Histological architecture may vary from place to place. Some tumour cells are round to polygonal, contain clear cytoplasm and have well-defined cytoplasmic margins. In fact, they were previously regarded as clear cell ependymoma of the foramen of Monro. They often have a honeycomb appearance that simulates oligodendrogliomas. In other areas, tumour cells may have a substantial amount of neuropils. Ganglioid cells and Homer Wright rosettes are rare findings. Although histologically similar, neurocytoma are separated from oligodendroglioma by immunohistochemical and ultrastructural features. The most useful diagnostic marker is synaptophysin, which is strongly expressed in both the cytoplasm and neuropil of these tumours but not in oligodendrogliomas. If synaptophysin is not detected, electron microscopy may be used to confirm the presence of neurosecretary granules. Synapses are also common but are not required for diagnosis. NFPs are usually not expressed except in ganglioid cells and GFAP is expressed by a small number of tumours. Neurocytomas usually behave in a benign fashion but a small number of them may exhibit increased proliferative activity, necrosis and other features indicative of aggressive behaviour.

Choroid Plexus Tumour

Choroid plexus has features of epithelial cells including tight junctions and basement membrane. Choroid plexus tumours, namely papilloma and carcinoma, are intraventricular tumours. Over 90% arise from the lateral and fourth ventricles while about 5% are found in the third ventricles. Although 80% of these tumours occur before the age of 20 years, most of them occur before the age of 2 years; congenital tumours have been described. They

typically cause hydrocephalus by two mechanisms: obstruction and overproduction of spinal fluid.

Choroid plexus papilloma appears as a cauliflower-like mass that is adherent to the ventricular wall but well demarcated from the brain parenchyma. Histologically, there is recapitulation of normal choroid plexus structure. The tumours are composed of delicate, often oedematous, fibrovascular connective tissue fronds that are covered by single-layered columnar to cuboidal epithelium with basally located nuclei (**Figure 21**). A basement membrane can easily be demonstrated by a periodic acid–Schiff (PAS) stain. Pseudostratification may occur but mulilayered epithelial cells are not seen in papilloma. A few mitotic figures may be present but, unless excessive and atypical, do not indicate malignancy. Brain invasion and necrosis are absent. Like their normal counterparts, the tumours express cytokeratins. Most express vimentin and S-100. Although GFAP is absent in normal choroid plexus, it is focally expressed in about 25–55% of choroid plexus papillomas. Transthyretin (prealbumin), a protein involved in the transport of thyroxine, is expressed in normal choroid plexus and papillomas. Expression of synaptophysin occurs in normal choroid plexus, choroid plexus papilloma and carcinoma but the significance is unclear.

Choroid plexus carcinomas show partial lost of papillary pattern with tumour cells arranged in solid sheets. There are numerous mitoses, nuclear pleomorphism, increased cellularity, necrosis and often extensive brain invasion. The pattern of antigen expression is similar to choroid plexus papillomas. Expression of transthyretin, however, may be reduced. Proliferative activity of the carcinoma (mean 13.8%, range 7.3–60%) is significantly higher than found in papillomas (mean 1.9%, range 0.2–6.0%) (Vajtai *et al*., 1996). Choroid plexus carcinoma is rare in adults. Such a diagnosis should only be made when the possibility of metastatic papillary carcinoma is eliminated through thorough morphological and immunohistochemical study.

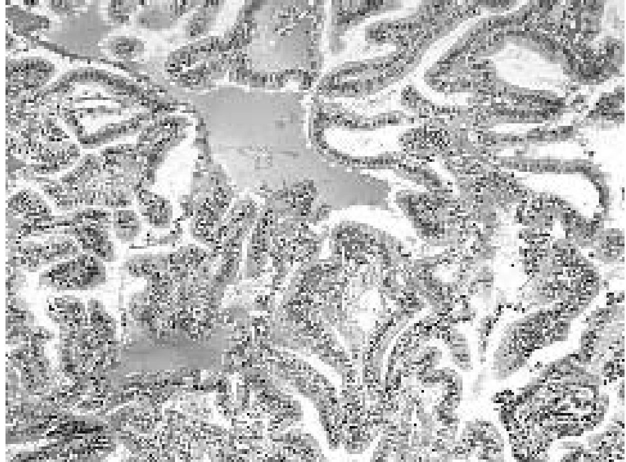


Figure 21 Choroid plexus papilloma show branching papillary structures covered by a single layer of epithelium.

Choroid plexus tumours tend to spread throughout the cerebrospinal fluid. Whereas the carcinomas may produce frank metastases, papilloma with spinal dissemination may remain asymptomatic. The overall 5-year survival rate for choroid plexus papilloma is 100% and for choroid plexus carcinoma 40%. Choroid plexus carcinoma is one of the rare brain tumours that can metastasize to visceral organs.

Pineal Parenchymal Tumour

Pineocytoma is a rare neoplasm that can occur at all ages but older children and young adults are predominantly affected. They tend to be well-defined masses which are composed of solid sheets of small, uniform, mature cells that resemble pineocytes. Some have neuropil or are arranged in a circular pattern that looks like an enlarged Homer Wright rosette. These large rosettes are termed ‘pineocytoma rosettes’ as they are only seen in these tumours. A papillary subtype of pineocytoma has also been described. These tumours strongly express synaptophysin and variably express NFP and other neuronal markers such as class III β-tubulin. Expression of retinal S-antigen and rhodopsin indicating their photoreceptor lineage can often be demonstrated. The 5-year survival of patients with pineocytoma is about 86% (Schild *et al*., 1996).

Pineoblastomas, in contrast, are highly malignant tumours and almost half of them occur within the first 10 years of life. They share many features of PNETs and are better regarded as PNETs with photoreceptor differentiation. Macroscopically, they are soft, friable, haemorrhagic and poorly demarcated tumours. Histologically, they are composed of small primitive cells with hyperchromatic nuclei and scant cytoplasm. Mitotic activity is brisk. They may contain rosettes that resemble those in PNETs in other locations (Homer Wright rosette) or in retinoblastoma (Flexner–Wintersteiner rosette). Dissemination through CSF is common. They are variably immunoreactive for synaptophysin, neurofilament proteins, class III β-tubulin, chromogranin A and retinal S-antigen.

The distinction between pineocytoma and pineoblastoma is blurred in a significant number of cases, in which event they are regarded as pineal parenchymal tumours with intermediate differentiation.

Primary Tumours of Meninges and Mesenchyme

This family of tumours arise within intracranial and intraspinal locations and recapitulate the normal supporting tissue of the CNS such as the leptomeninges and the surrounding mesenchyme.

Meningiomas

Meningiomas are tumours that consist of arachnoid (meningothelial) cells and are most commonly seen in

the areas where arachnoid villi are most abundant, such as the parasagittal region and olfactory groove, and display features of mature meningothelial cells; sometimes they arise from intraventricular locations and are rare in tissue outside the cranium and spinal cord. These are common tumours and comprise 13–26% of all primary intracranial mass lesions. Meningiomas are essentially tumours of middle-aged and elderly individuals, and are only rarely seen in children and infants. They are twice as common in women than in men. These tumours are slow growing and many remain asymptomatic for years. However, local recurrence after surgery is not uncommon. The neurological manifestations are closely related to compression of adjacent structures and the specific neurologic deficits depend on the location of the tumour.

As noted, meningiomas are preferentially found over the cerebral convexity, olfactory groove, sphenoid ridge, parasellar regions, petrous ridge and tentorium cerebelli. Multiple tumours are not infrequent, particularly in patients with neurofibromatosis 2 (see below), and are not necessarily an indication of malignancy. On MRI, meningiomas typically appear as well demarcated masses that enhance homogeneously and may be associated with a variable amount of oedema in the surrounding brain. Grossly, they are well-demarcated, round to lobulated, rubbery to firm masses (**Figure 22**). The cut surface is often homogeneous and tan and can be gritty when substantial calcification is present. Brain tissue is often compressed by the tumour. Intracranial meningiomas are typically attached to the dura whereas those in the vertebral canal are adhered to leptomeninges. These tumours may invade the skull bone and adjacent sinuses or may arise entirely within bone and be associated with osteoblastic activity.

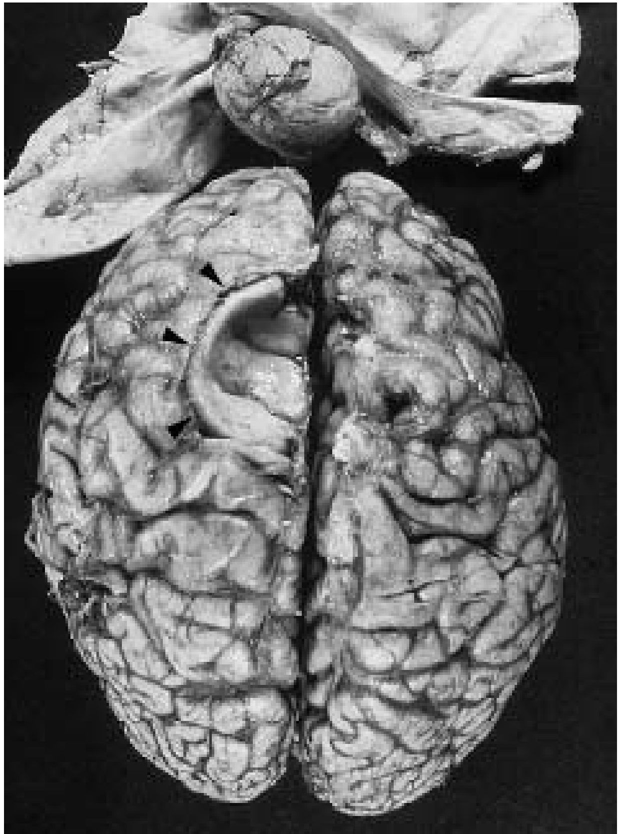


Figure 22 Parasagittal meningioma which has compressed but not invaded the surrounding gyri producing atrophy (arrow head).

Meningiomas are basically histologically benign tumours. They have the most diversified histopathological features and the histology is often variable within the same tumour. Nuclei of the tumour cells are small, round to oval, and display minimal pleomorphism; they may contain nuclear pseudoinclusions. Sporadic large and hyperchromatic nuclei can sometimes be seen. Proliferative activity, as measured by cell cycle labelling, is low; high proliferative activity may be correlated with an increased risk of recurrence. There is no significant pleomorphism, mitotic activity or necrosis.

Histological features of meningiomas are diverse. There are several common histological subtypes of meningiomas. The meningothelial type is composed of tumour cells that are arranged into sheets and lobules that are separated by thin collagenous septae (**Figure 23**). Cytologically they are benign looking and resemble arachnoid cells (**Figure 24**) and very often they arrange themselves into small whorls of cells resembling arachnoid granulations. The fibrous (fibroblastic) meningioma is composed of spindle-shaped cells with elongated cytoplasm that resemble fibroblasts (**Figure 25**). These cells are arranged in parallel and interlacing bundles intermingled with a variable amount of dense collagen. The transitional (mixed) type is essentially a hybrid that has features of both the meningothelial and fibroblastic types. Tumour cells of the psammomatous meningioma tend to arrange

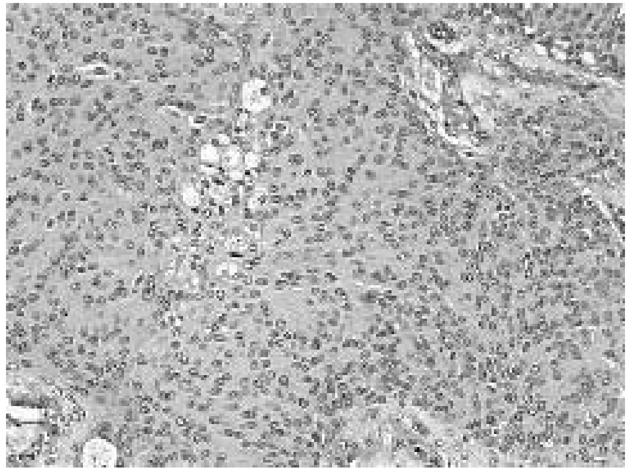


Figure 23 Typical field of meningotheliomatous meningioma showing sheets of arachnoid cells and a small focus of foamy macrophages.

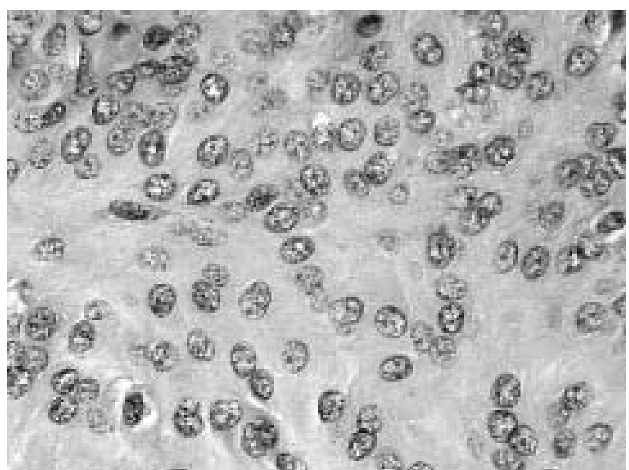


Figure 24 Nuclei of meningioma resemble those of the arachnoid cells and are round to oval and have a speckled chromatin pattern.

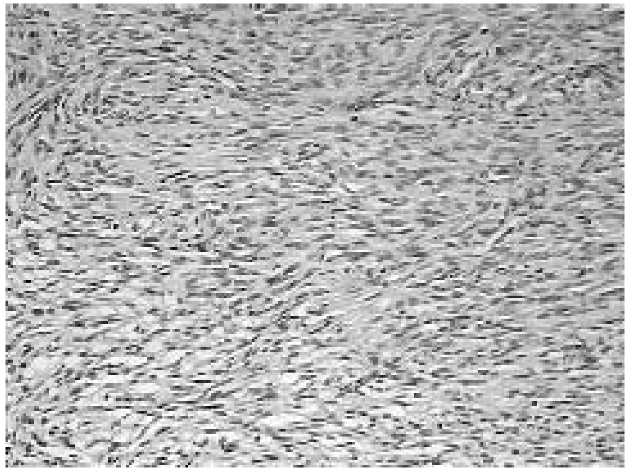


Figure 25 Photomicrograph of fibroblastic meningioma composed of enlarged spindle cells reminiscent of fibroblasts.

themselves into whorls that often contain psammoma bodies, a round to irregular, laminated calcified structure.

Less common types include angiomatous meningioma, microcystic meningioma, secretory meningioma, lymphoplasmacyte-rich meningioma and metaplastic meningioma.

While the aforementioned variants do not exhibit differences in their biological behaviour, papillary meningioma is a variant that behaves aggressively. These tumours are characterized by a perivascular pseudopapillary arrangement of tumour cells. They often have mitotic activity and pleomorphism and tend to recur locally and metastasize to other body regions. Clear cell and chordoid meningioma may also be associated with increased risk of recurrence and aggressiveness.

Immunohistochemically, most meningiomas express EMA in a patchy manner. Expression is less consistent in atypical and anaplastic meningiomas. Vimentin is seen in virtually all meningiomas. These tumours do not express GFAP, a useful feature for differentiating them from desmoplastic gliomas.

Characteristic ultrastructural features of meningiomas include interdigitations of the cytoplasmic membrane, hemidesmosones and desmosomes. They contain abundant amounts of intermediate filaments, sometimes with a pronounced whirling pattern. These filaments are also firmly anchored to the desmosomes, a feature that is also seen in normal arachnoid cells.

Atypical and Anaplastic Meningiomas

Atypical meningiomas display features that suggest aggressive biological behaviour including four mitotic figures per 10 high-power fields, uninterrupted patternless or sheet-like cell growth, increased pleomorphism and necrosis. Such atypical changes can occur with any of the histological variants.

Anaplastic (malignant) meninigiomas display more sinister histological features that far exceed those displayed by atypical meningiomas. Histologically, they are frankly malignant. Both atypical meningiomas and anaplastic meningiomas have a significantly higher proliferation rate than meningioma.

Brain invasion by meningioma is an indication of aggressive behaviour but not an equivalent to frank malignancy as histologically benign meningiomas may invade the brain. On the other hand, atypical and anaplastic meningiomas frequently invade the brain.

Mesenchymal Non-meningothelial Tumours

This is a family of rare tumours that arise within the intracranial and/or intraspinal cavity. They exhibit mesenchymal differentiation of various types including fat, fibrous tissue, muscle, bone, endothelial cells and others. This heterogeneous family includes benign tumours such as lipomas and highly malignant lesions such as mesenchymal chondrosarcomas. Some tumours, such as rhabdomyosarcoma, are preferentially seen in children, whereas other tumours such as chondrosarcoma are often seen in adults; no age group is spared. Only a few will be discussed here.

Vascular Tumours. Vascular lesions of the brain fall into three major groups. The most frequent type is malformative in nature and includes arteriovenous malformations, cavernous angiomas, venous angiomas and capillary telangiectases. These must be distinguished from true neoplastic vascular lesions. Interestingly, these vascular malformations are often associated with familial syndromes that involve malformations of the CNS and blood vessels.

The second type is a true neoplasm with the neoplastic cells displaying histological and immunohistochemical features of endothelial cells. The tumour cells express factor-VIII and CD31 and can also be identified with

Ulex europeus lectin. Haemangiomas are benign tumours. Epithelioid haemangioendotheliomas are rare vascular tumours of intermediate malignancy which arise from the skull base, dura or brain parenchyma. Angiosarcomas are rare but highly malignant vascular tumours. Kaposi sarcomas are only exceptionally encountered and almost always associated with acquired immune deficiency syndrome (AIDS).

In the third type, the tumours have such a rich vasculature to suggest a vascular tumour but the neoplastic cells do not exhibit features of endothelial cells. Examples include haemangiopericytoma and haemangioblastoma.

Haemangiopericytoma. Haemangiopericytoma is an uncommon tumour of intermediate-grade malignancy. The tumours are almost always dural-based and tend to occur in young adults. Macroscopically, they are dural-based solid masses well demarcated from the adjacent brain tissue, almost invariably solitary at initial presentation. Large, thin-walled and dilated blood vessels often with the appearance of a stag-horn are present. Haemangiopericytomas are highly cellular tumours composed of compactly arranged, monotonous polygonal cells with ill-defined cytoplasm and usually mitotically active. There are numerous small vascular channels between the tumour cells. The tumour cells arrange themselves in lobules separated by thin vascular spaces. The neoplastic cells are not immunohistochemically reactive for CD31, factor-VIII or *Ulex europeus*. Characteristically, reticulin material surrounds the individual cells. This is confirmed by ultrastructural demonstration of basal lamina-like amorphous material around the cells. Local recurrence is a rule.

Haemangioblastoma. Haemangioblastoma is a highly vascular tumour but the neoplastic cells do not express features of endothelium. In fact, their nature is poorly understood and, therefore, they are classified as tumours of uncertain histogenesis in the WHO Classification. For practical purposes, they are discussed with other vascular tumours. Haemangioblastomas are relatively common and about 25% of them are associated with von Hippel–Lindau disease. This is a biologically benign tumour and occurs most commonly in young and middle-aged adults; the cerebellum is the preferred site, although they can occur in any part of the CNS. Macroscopically, they are often well-circumscribed cystic tumours, and the strategic position in the posterior fossa often obstructs CSF flow. Histologically, they are highly vascular (**Figure 26**) and are composed of sheets and lobules of vacuolated large stromal cells, the histological hallmark of the tumour. This is combined with a rich capillary network (**Figure 27**). Immunohistochemically, they do not express markers of endothelial cells. The stromal cells contain lipid droplets, a feature best demonstrated by fat stains on frozen sections or electron microscopy, but do not express any marker of endothelial cells. Their nature remains a matter of dispute.

Chordoma. Chordomas are tumours of notochordal tissue that arise mainly at either end of the vertebral column. These tumours may involve the cranial base and extend into the middle and posterior fossae and penetrate the dura. Chordomas in the sacral region are more common, and rarely extend into the spinal canal. Most cases are seen after the age of 50 years (Dorfman and Czerniak, 1998). Although locally invasive, these tumours are histologically benign. Macroscopically, they are grey, bulky, soft to gelatinous masses in consistency. The histological hallmark of chordoma is the large and bubbly physaliferous cell (**Figure 28**) which has pale eosinophilic or amphophilic cytoplasm and a centrally located small-to-medium-sized nucleus. These cells form cohesive sheets or nests and are embedded in a sea of strongly metachromatic acellular substances. Ultrastructurally, the tumour cells contain numerous desmosome-type

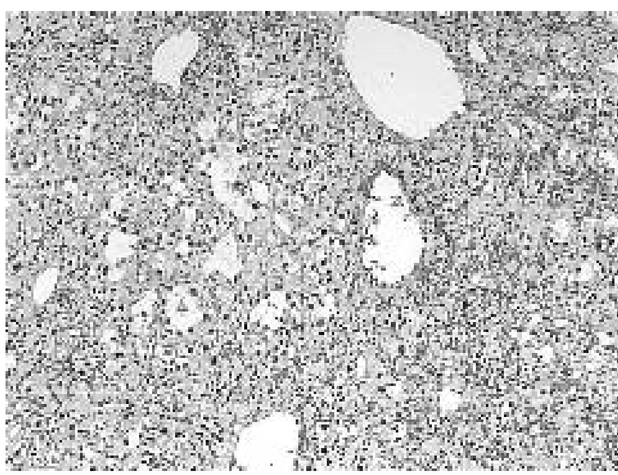


Figure 26 Haemangioblastomas are richly vascular and often have dilated blood vessels.

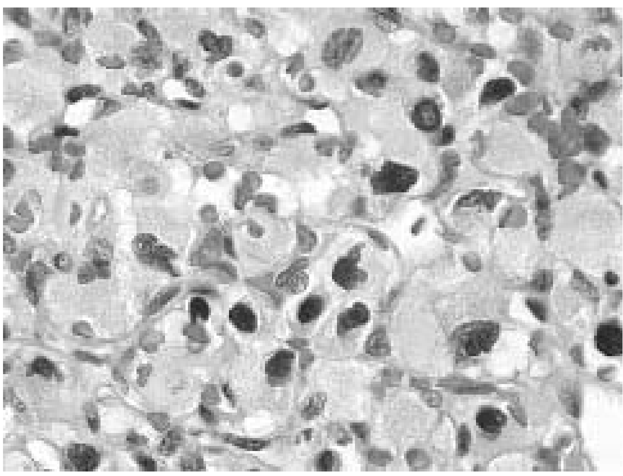


Figure 27 High-magnification photomicrograph of stromal cells in haemangioblastoma showing pale to vacuolated cytoplasm with centrally located nuclei.

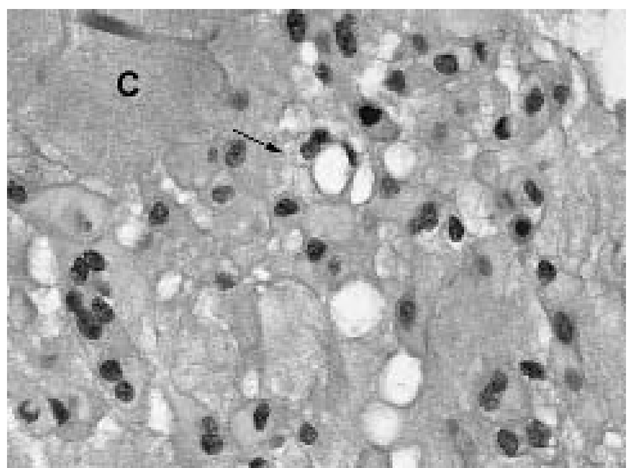


Figure 28 Photomicrograph of chordoma showing bubbly physaliferous cells (arrow) and metachromatic acellular material (C).

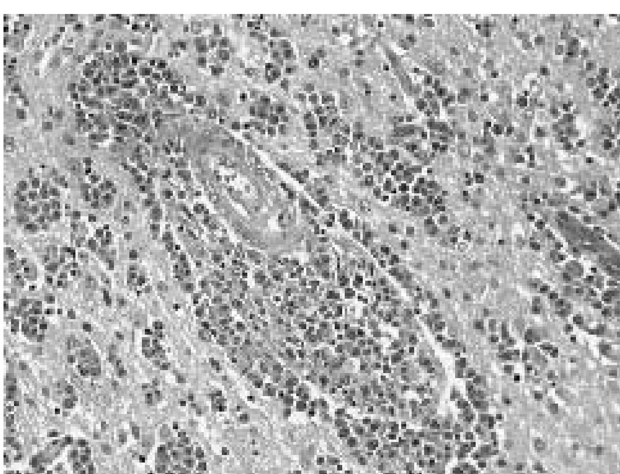


Figure 29 Neoplastic lymphoid cells demonstrating characteristic concentric perivascular arrangements.

intercellular junctions which tightly bond them to one another. These cells are immunohistochemically reactive for S-100 protein, EMA, vimentin and cytokeratin. These tumours must be distinguished from chondrosarcoma which they may mimic histologically.

Primary Tumours of Tissue Outside the CNS

These interesting primary tumours have no normal counterparts in the mature CNS. Three families of tumours comprise this group: tumours of the haematopoietic system, germ cell tumours and teratomas, and craniopharyngioma.

Primary Tumours of the Haematopoietic System

The CNS may be involved by both primary and secondary tumours of the haemotopoietic system. A variety of primary lymphoproliferative disorders, including intraparenchymal primary lymphoma, plasmacytoma and mucosal associated lymphoid tissue (MALT) lymphoma of the dura and post-transplantation lymphoproliferative disorder, can be seen in the CNS. Primary non-Hodgkin lymphoma is the most commonly seen entity. In contrast, primary Hodgkin disease of the CNS is exceedingly rare. Primary histiocytic tumours, predominantly Langerhans histiocytosis, are also seen in the CNS.

Primary non-Hodgkin lymphoma of the CNS was a rare disease until the epidemic of HIV infection and wide use of immunosuppression for organ transplantation. In immunocompetent patients, primary CNS lymphoma is seen mainly in elderly patients, but in immunodeficient patients, the age of onset is related to the time of onset of the immunodeficiency and all age groups are affected. Patients with hereditary immunodeficiency disorders have the lowest age of onset, followed by patients with acquired immunodeficiency such as AIDS, and then by patients with organ allograft transplantation. Symptoms and clinical presentation are nonspecific. MRI may reveal solitary or multiple hyperdense or isodense lesions that may or may not enhance. Cystic change is not common. Macroscopically, these tumours are preferentially found in deep-seated brain tissue such as the basal ganglia and periventricular areas. Demarcation from the surrounding tissue can be variable and the consistency varies from soft to firm.

Classification of lymphomas changes frequently. The most recent system in use is the Revised European–American Lymphoma (REAL) classification (Harris *et al.*, 1994), which defines entities on the basis of morphological, immunological, genetic and clinical information. Over 98% of all primary CNS lymphomas are B cell lymphomas, predominantly the diffuse large B lymphoma. T cell lymphomas represent the other 2% of primary CNS lymphoma. Morphologically, the lymphomas are characterized by a multifocal, perivascular (**Figure 29**) and concentric infiltration by large, atypical and monotonous lymphocytes which extensively infiltrate brain tissue. Reticulin fibres around the blood vessels are typically layered and infiltrated by the neoplastic lymphocytes giving an 'onion bulb' appearance. Immunohistochemically, B cell lymphomas express pan-B markers such as CD20 and CD79a. They express monoclonal surface or cytoplasmic immunoglobulins with IgM and kappa light chain being the most common combination. Immunocompetent patients are usually responsive to therapy and the overall 5-year survival rate is about 25–45%. The prognosis in AIDS patients is much more ominous and their median survival rate is only 10–18 months when treated with multimodal therapy.

Germ Cell Tumours

Germ cells are not normal residents of the CNS. However, it is thought that a small number of primordial germ cells are disseminated along the migration trail from the yolk sac to the gonads during embryogenesis. These cells are

the putative origin of germ cell tumours in the CNS. Some tumours are primarily seen in children and are more common in men than women. For unknown reasons, the prevalence in Asian countries is higher than that in the West. They typically occur along the midline with the pineal gland (**Figure 30**) being the most common site, followed by the suprasellar region. The third ventricle, optic chiasma and tectal plate are often affected. In consequence, they often cause hydrocephalus, precocious puberty, visual field defects and Parinaud syndrome.

Like other gonadal and extragonadal germ cell tumours, those arising in the CNS display a spectrum of differentiation. Germinoma, choriocarcinoma, yolk sac tumour, embryonal carcinoma, mixed germ cell tumour and immature and mature teratoma can all occur as primary tumours of the CNS. Pure germinoma and teratoma are, however, the most commonly seen entities. Their pathology is similar to that of other gonadal and extragonadal tumours. Germinomas are generally soft and friable. Histologically, they are composed of large, uniform cells with large vesicular nuclei containing a prominent nucleolus which resemble primitive germ cells. There is an admixture of lymphocytic cells. This is variable but sometimes intense. Placental alkaline phosphatase (PLAP) is expressed by germinoma cells and is helpful in diagnosis. Yolk sac tumours display histological features reminiscent of the embryonic yolk sac endoderm and are composed of primitive-appearing cells that arrange themselves into loosely packed, variably cellular retiform structures. When present, Schiller–Duval bodies are diagnostic. Yolk sac tumours characteristically contain eosinophilic hyaline globules that are PAS positive and immunoreactive for α-foetoprotein (AFP). The histology of embryonal carcinoma may be confused with that of yolk sac tumour. These tumours, however, are more monotonous in their patterns and are composed of sheets of undifferentiated epithelial cells. Embryonal carcinoma expresses cytokeratin and carcinoembryonic antigen (CEA). Choriocarcinoma is a germ cell tumour that differentiates along the trophoblastic lineage and is characterized by the presence of cytotrophoblastic and syncytiotrophoblastic elements. These tumours are immunoreactive for β-human chorionic gonadotrophin.

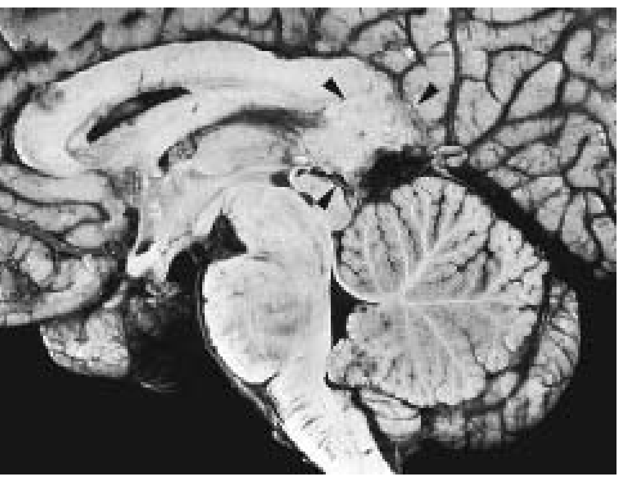


Figure 30 Sagittal section of brain showing yolk sac tumour (arrow head) arising from the pineal gland.

Teratomas, in contrast to the aforementioned germ cell tumours, differentiate along ectodermal, endodermal and mesodermal lineages. Derivatives from all three germ layers such as neural tissue, respiratory epithelium, cartilage, bone, secretory glands and others may be found. In mature teratoma, all elements are fully differentiated 'adult-type tissue.' Immature teratomas contain elements that are not fully developed. Teratoma with malignant transformation refers to tumours that have a frankly malignant component, most often a rhabdomyosarcoma or undifferentiated sarcoma.

Because of the secretion of PLAP, AFP and β-HCG, germ cell tumours currently compose the only class of primary CNS tumours in which detection of primary and recurrent tumour can be easily assayed for oncoproteins in CSF and serum.

Craniopharyngioma

Craniopharyngiomas are biologically benign but locally invasive tumours. They have a bimodal distribution; the bulk of them are seen in children and young adults and the second peak occurs around the sixth decade. They occur in the suprasellar region (**Figure 31**), often have an intrasellar component and are presumably derived from Rathke's pouch. Their proximity to the optic chiasma and pituitary gland make visual disturbance, endocrine abnormalities of the anterior pituitary and diabetes insipidus the

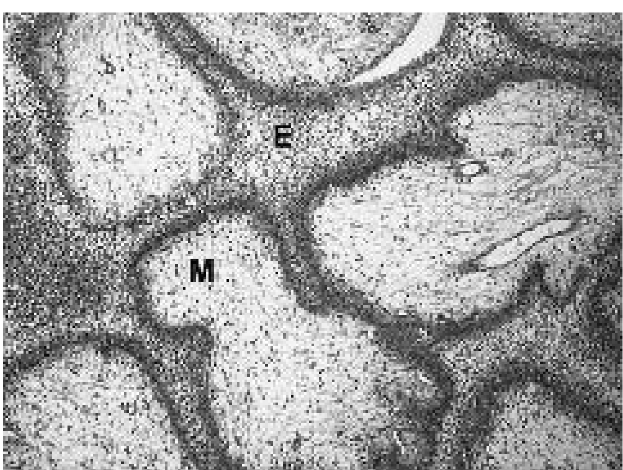


Figure 31 Photomicrograph of an adamantinomatous craniopharyngioma showing a complex pattern composed of anastomosing trabeculae of epithelial cells (E) with peripheral palisading. Supporting the epithelial cells is delicate mesenchymal tissue (M). This field is highly reminiscent of ameloblastomas that occur in the jaw.

most frequent clinical presentations. Compression of the third ventricle often causes hydrocephalus. Craniopharyngiomas are well-demarcated tumours that vary greatly in size and extent. They vary from solid to predominantly cystic. The cystic lesions often contain a cholesterol-rich, thick, brownish yellow fluid that looks like machine oil.

Two major histological variants, namely the adamantinomatous and papillary type, have been described. The adamantinomatous type is histologically very similar to if not identical with ameloblastoma or calcifying odontogenic tumour of the jaw. They are characterized by anastomosing trabeculae of stratified squamous epithelium with peripheral palisading of nuclei. Enclosed within these islands are loosely arranged stellate-shaped epithelial cells and nodular and often calcified masses of 'wet keratin,' a distinctive diagnostic feature of craniopharyngioma. These cell islands are held together and nourished by the vascularized connective tissue stroma. Substantial calcification is often associated with the components that resemble calcifying odontogenic tumour. The papillary craniopharyngioma is composed of pseudopapillae covered by keratinizing squamous cell epithelium. Craniopharygioma is a biologically benign but locally invasive tumour, often associated with substantial gliosis in the surrounding brain tissue. Surgical removal remains the most effective approach and they recur occasionally.

Secondary Tumours

Secondary neoplastic involvement of the CNS falls into two major groups, namely metastatic tumours and secondary involvement of the CNS by haematopoietic tumours such as lymphomas and leukaemias; both involve haematogenous spread. They may present as single or multiple lesions in the brain.

Carcinoma, malignant melanoma and germ cell tumours often produce metastases to the CNS. In contrast, metastatic sarcomas with the exception of alveolar soft part sarcoma are not as common. Metastatic tumours are typically associated with substantial oedema and the clinical symptoms may be disproportionate to the size of the metastases *per se*. The location and size of the metastasis but not the type of tumour dictate the type and nature of neurological abnormalities. Whereas metastases in the cerebral cortex may cause seizures, metastases to the posterior fossa structure often cause hydrocephalus by obstructing the flow of CSF. Metastasis to the spinal cord may compress the spinal cord and produce hemiplegia or quadriplegia, depending on the level of involvement.

Intracranial metastases are most frequently seen in the brain and occasionally in the dura. In the cerebrum, the tumours are most frequently located at the cortical–medullary junction where the calibre of the blood vessels abruptly changes from large to small; most of them are found in the arterial watershed areas of the cerebral hemispheres. In contrast, epidural metastases are most commonly seen in the spinal cord although leptomeningeal and intramedullary (spinal cord) metastases can also occur.

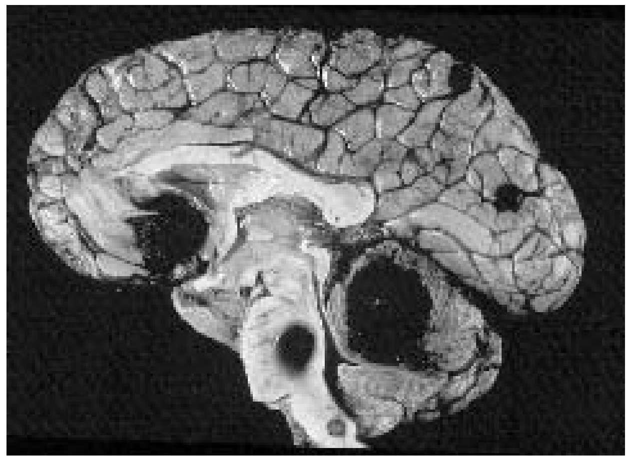


Figure 32 Multiple metastases from a peripheral malignant melanoma. Note the carbon black colour of the tumour.

Over 50% of all metastatic carcinomas originate from the respiratory tract. Adenocarcinoma, small cell and large cell carcinomas are far more likely to generate metastases than squamous cell carcinoma of the lung. Other common origins include carcinoma of the breast and malignant melanoma (**Figure 32**). Renal cell carcinoma has an unexplained high tendency to metastasize to the cerebellum and prostate carcinoma tends to generate epidural metastases that compress the spinal cord. Germ cell tumours of the testis metastasize to the brain in 15–25% of patients and in patients with choriocarcinoma the rate is 83%.

In most cases, the histopathology of the metastatic tumour is similar to the original tumour. The primary tumours in some cases are never found. In these cases, histology and immunohistochemistry may be helpful in identifying the tumour of origin, although this may not always be possible. The survival rate of patients with metastatic tumour in the brain varies greatly but those with single metastasis show the best survival.

Tumours of the PNS

Similar to the CNS, tumours of the PNS can also display features of embryonal and mature features of the PNS. The reader should refer to the Further Reading section for detailed discussion of PNS tumours.

Neuroblastic Tumours of the Adrenal Gland and Sympathetic Nervous System

Neuroblastic tumours are childhood tumours composed of immature neuroectodermal cells derived from the neural crest with or without associated mature ganglionic elements. Neuroblastic tumours are the most common

extracranial malignant tumours during the first 2 years of life and they tend to metastasize widely. About half arise from the adrenal medulla whereas the remainder are associated with ganglia along the sympathetic chain. They typically present clinically as a palpable abdominal mass, hepatomegaly or a thoracic mass. They may compress the spinal cord and/or adjacent nerve roots and present with neurological symptoms. A small number may secrete vasoactive intestinal polypeptide and cause diarrhoea.

The International Neuroblastoma Pathology Classification by Shimada *et al*. (1999) provides a comprehensive approach in correlating the histopathology and clinical outcome of neuroblastic tumours. The neuroblastic cell nests may also display various degrees of differentiation. Neuroblastoma refer to a lesion that is composed exclusively of primitive neuroectodermal cells (**Figure 33**). Ganglioneuroblastoma is composed of nests and islands of neuroblastic cells embedded within a schwannian stroma that contains ganglionic cells (**Figure 34**). Ganglioneuroma is composed predominantly of mature ganglionic cells in a schwannian background, although a small amount of differentiating neuroblasts can be present. Grossly, they are encapsulated, soft, grey–tan tumours. Ganglioneuromas are firmer than neuroblastomas. Histologically, neuroblastomas are composed of neuroblastic cells that have small, hyperchromatic nuclei and minimal cytoplasm. Some of them may display features of differentiating neural cells with neuropil and Homer Wright rosettes. Adrenergic, cholinergic or mixed neurotransmitter enzyme and intracytoplastic catecholamines may be demonstrated histochemically. These tumours may also express chromogranin, gamma-subunit of neuron-specific enolase, neurofilament proteins and neuronal microtubule-associated proteins. The schwannian stroma expresses S-100 protein. The overall prognosis of neuroblastoma correlates with clinical staging, the histological grade of tumour, expression of TrkA and N-Myc amplification.

Olfactory Neuroblastoma (Esthesioneuroblastoma)

Olfactory neuroblastoma is a malignant tumour that is composed of primitive neuroepithelial cells and is assumed to arise from olfactory receptor cells in the nasal cavity. In contrast to neuroblastoma of the adrenal gland and sympathetic ganglia, olfactory neuroblastomas are fairly uncommon and occur predominantly in adults. They are slow-growing tumours and are often associated with long-standing symptoms. Invasion into the adjacent structures and metastasis, often to the cervical lymph nodes, is common. Macroscopically, they are soft and richly vascularized polypoid lesions. Histologically, they are composed of primitive neuroectodermal cells arranged in lobules and often form Homer Wright rosettes. Immunohistochemistry is often necessary to distinguish them from other primitive tumours.

Neurofibroma

Neurofibroma is a benign tumour that may occur in all age groups. The exact incidence of neurofibroma is unclear and they are often associated with neurofibromatosis 1 (NF1). They are most frequently found in the part of the PNS that is most distant from the spinal cord and brain; they are occasionally seen in spinal nerve roots and rarely seen in the cranial nerves. When neurofibromas arise from a small peripheral nerve, they characteristically diffusely infiltrate the nerve. They may be confined by the perineurium when they arise from medium-sized nerves but they tend to infiltrate into the surrounding soft tissue. Neurofibromas usually present as cutaneous nodules (localized cutaneous

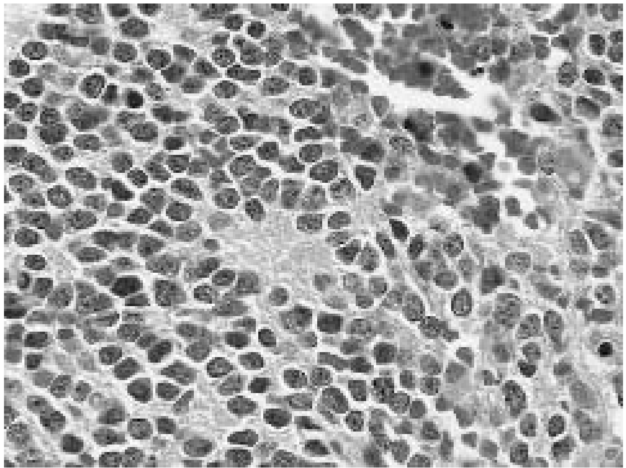


Figure 33 Neuroblastoma arising within the adrenal gland showing a Homer Wright (neuroblastic) rosette (centre of the figure). (Courtesy of Dr David Carpentieri, Children's Hospital of Philadelphia.)

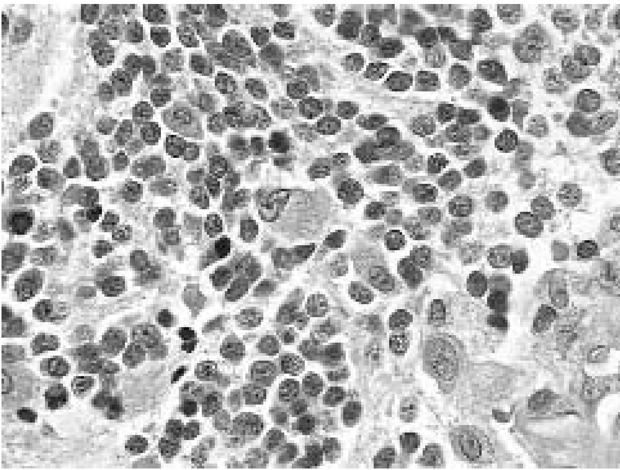


Figure 34 Ganglioneuroblastoma displaying neuronal differentiation (large cells). (Courtesy of Dr David Carpentieri, Children's Hospital of Philadelphia.)

neurofibroma) and less commonly as isolated, well-circumscribed masses (isolated intraneural neurofibroma) within nerves. In patients with NF1, they often present as plexiform neurofibromas that appear as diffuse enlargement of major nerve trunk and their branches. Macroscopically, they are firm and grey–tan tumours. While the isolated intraneural neurofibromas may be well demarcated, the plexiform neurofibroma typically appears like a string of sausages (**Figure 35**). Since they arise within the nerve and become an integrated part of it, they cannot be removed without sacrificing the nerve. Cells composing the tumour are fibroblasts, Schwann cells and perineural cells that are embedded within a collagenous and mucoid matrix (**Figure 36**). Most tumour nuclei are small and hyperchromatic; mitotic figures are rare. Axons and myelin degenerate as a consequence of pressure on nerves by neoplastic cells. S-100 is invariably expressed in these lesions. Ultrastructural examination documents the presence of Schwann cells and perineural cells.

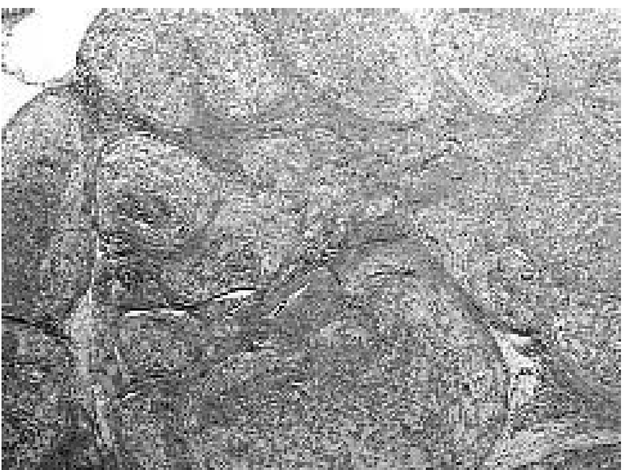


Figure 35 Plexiform neurofibroma showing variable enlargement of nerve bundles which completely obscure normal structure.

Schwannoma

Schwannomas are common benign tumours that are most frequently seen in middle-aged patients (fourth to sixth decades), although they are seen in all age groups. Unless associated with neurofibromatosis 2 (NF2) they are usually solitary. They arise most frequently from sensory branches of peripheral nerves of the head and neck region and extensor aspects of the extremities and are often asymptomatic. When they arise from the spinal nerve roots, they may have a dumbbell shape and compress the nerve root and spinal cord. The cochlear–vestibular nerve is also a common site, causing tinnitis and hearing loss. Rarely, these tumours may be found within the brain and spinal cord.

The majority are well-circumscribed, oval, solid to cystic, tan, rubbery tumours. In contrast to neurofibroma which diffusely infiltrates the nerve, schwannomas compress but do not excessively infiltrate the nerve. Thus, the nerve may be be salvaged during surgical excision of schwannomas. Histologically, schwannomas are composed of elongated spindle-shaped cells that exhibit a variety of patterns. The tumour cells often arrange into Antoni A and Antoni B patterns. The Antoni A pattern is characterized by compact spindle cells that may have nuclear palisading and whorling of cells. Verocay bodies (**Figure 37**) refer to structures formed by two compact rows of palisading nuclei separated by fibrillary cell processes. The nuclei of schwannoma are usually elongated. The Antoni B pattern consists of cells that are loosely arranged in a delicate and reticular pattern (**Figure 38**).

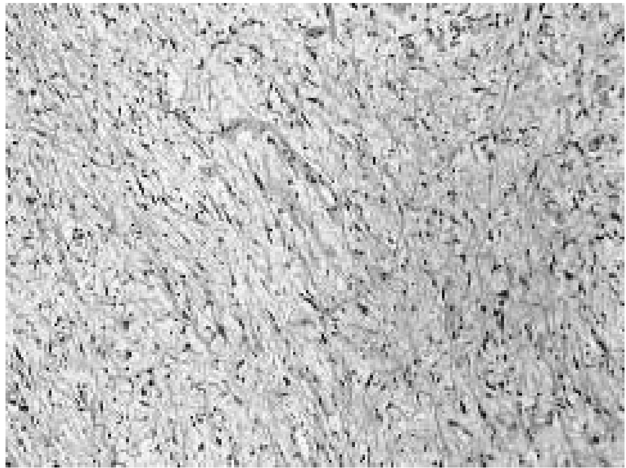


Figure 36 Neurofibromas typically contain delicate spindle cells and mucoid material.

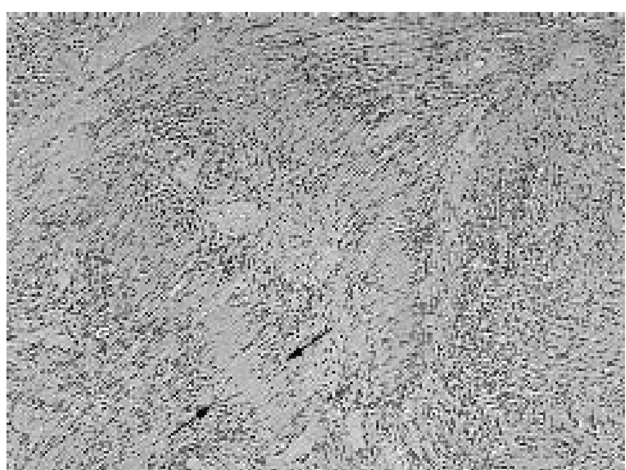


Figure 37 This photomicrograph shows the densely packed Antoni A areas and Verocay bodies formed by alternating palisading bands of nuclei and acellular tumour tissue (arrow).

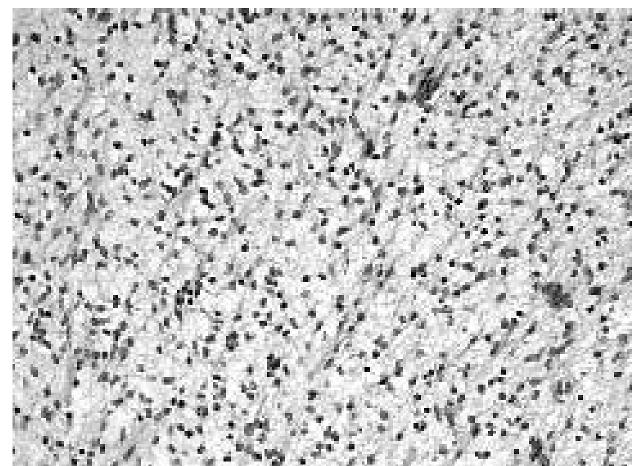


Figure 38 In Antoni B areas, the tumour cells are loosely packed and do not grow in any specific pattern.

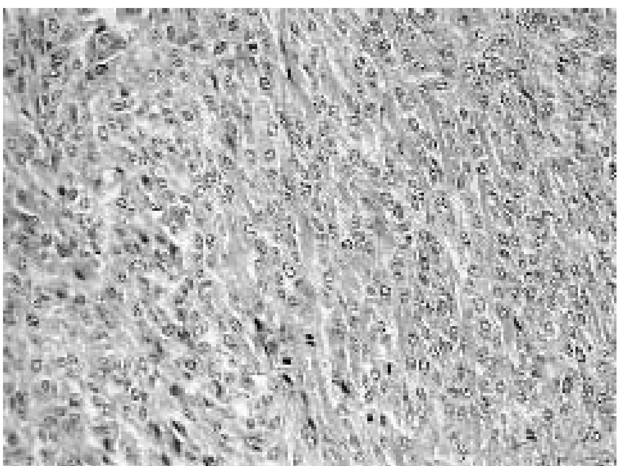


Figure 39 A malignant peripheral nerve sheath tumour can display densely packed and pleomorphic spindle cells. (Courtesy of Dr Paul Zhang, University of Pennsylvania.)

Lipid-laden macrophages and mucoid material are often present. Cellular schwannoma is a variant with increased cellularity. It is a benign variant of schwannoma, although recurrences have been reported. Melanotic schwannomas contain melanin and have immunohistochemical and ultrastructural features of both Schwann cells and melanocytes.

Malignant Peripheral Nerve Sheath Tumour (MPNST)

Malignant peripheral nerve sheath tumours (MPNSTs) are uncommon and over half occur in patients with neurofibromatosis 1. They most commonly arise from plexiform neurofibromas. Many have both features of neurofibroma and schwannoma but the preferred diagnostic classification is MPNST. This has been defined as 'any malignant tumour arising from a peripheral nerve or showing nerve sheath differentiation, with the exception of tumours originating from epineurium or the peripheral nerve vasculature,' according to the WHO Classification. The most common sites of MPNST are deep nerve branches in the buttock, thigh, brachial plexus and paraspinal areas. Cranial nerves are uncommonly involved. On gross examination they are medium-sized, firm, hard, pseudoencapsulated tumours that are adherent to the nerve of origin. Necrosis is common. MPNSTs exhibit the most diverse histological features among all tumours of the PNS and soft tissue tumours. Significant portions may be undifferentiated (**Figure 39**). Diagnosis depends either on demonstration of tumour origin from a peripheral nerve trunk, a benign or malignant peripheral nerve tumour or immunohistochemical or ultrastructural features of Schwann or perineural cell differentiation.

Paraganglioma

This is a benign neuroendocrine tumour that arises in autonomic ganglia (paraganglia) throughout the body or within the adrenal medulla where this tumour is known as a phaeochromocytoma. Within the CNS, paragangliomas are restricted to the cauda equina and filum terminale. Paragangliomas usually occur in adults and some are endocrinologically active. These well-differentiated tumours are composed of large polygonal chief cells (type I cells) arranged in lobules (zellballen architecture) delimited by reticulin fibres and a single layer of inconspicuous sustentacular cells (type II cells). The relationship between these two types of cells is progressively lost with reduction of the sustentacular component as the tumour increases in malignancy. The chief cells are immunoreactive to synaptophysin and chromogranin whereas the sustentacular cells express S-100 protein and GFAP. The chief cells may also express other neuropeptides. Ganglionic differentiation is seen in over half of the paragangliomas arising within the cauda equina.

MOLECULAR GENETICS

Recent advances in molecular genetics have greatly improved our understanding of the aetiology of brain tumours. At the same time, such knowledge has been used to screen high-risk patients, to confirm diagnosis and to plan treatment. Limitation of space allows discussion of only a few examples.

Familial Tumour Syndromes Involving the Nervous System

This is a consortium of syndromes characterized by a combination of malformations, hamartomas and tumours of the nervous system with systemic neoplastic and non-neoplastic conditions. Many are associated with other

conditions which facilitate early clinical recognition (**Table 4**). Some have characteristic skin lesions and are known as neurocutaneous syndromes, phakomatoses or ectomesodermal syndromes, the most common of which are discussed.

Tuberous Sclerosis

Tuberous sclerosis involves multiple organs but the brain is most frequently affected. It is relatively common, the incidence being between 1 in 6000–10 000. Multiple hamartomas or slowly growing tumours involve the CNS, retina, skin, heart and kidney. This is an autosomal dominant condition with high penetrance. It is unusual to observe affected siblings of apparently normal parents. The *TSC1* gene on chromosome 9q34 that encodes hamartin and and *TSC2* gene on chromosome 16p13.3 that encodes tuberin have been identified. Hamartin is strongly expressed in organs that are affected by tuberous sclerosis such as brain, kidney and heart. Tuberin is also widely expressed and its pattern of expression overlaps that of hamargin. Although mutations of both *TSC1* and *TSC2* have been described in patients with tuberous sclerosis, genotype–phenotype correlation has not been well established.

Clinical manifestations are mostly related to the slowly growing tumours or hamartomas, and are highly variable. Severity is related to the extent of involvement and age of onset. In classical cases patients have seizures, are often mentally retarded and have adenoma sebacum; however, clinical manifestations can be extremely variable. Tuberous sclerosis can be separated into definitive, provisional and suspect category on the basis of specific clinical criteria. The pathology of tuberous sclerosis is diverse (**Table 6**), and only features pertinent to the CNS will be discussed.

Table 6 Manifestations of tuberous sclerosis

Brain	Cortical tuber
	Subependymal nodule
	Subependymal giant cell tumour
	Heterotopic grey matter
Skin	Adenoma sebacum (facial angiofibroma)
	Hypomelanotic macule
	Peri- or subungal fibroma
	Shagreen patches (fibrous hamartomas of dorsal surfaces, rarely seen before puberty)
	Poliosis and leucotrichia
Eye	Retinal giant cell astrocytoma
	Hypopigmented iris spot
	White eyelashes
	Hamartomata of eyelids and conjunctivae
Kidney	Angiolipoma
Heart	Single or multiple rhabdomyoma
Other organs affected	Bone, lung, liver, adrenals, gonads, thyroid, teeth and gingival tissue

Cortical tubers consist of pale, firm, cortical nodules that expand the gyri and blur the margin between grey and white matter. Calcification is common and can transform the tuber into a stony-hard structure. The number and location of tubers varies and they are scattered throughout the brain and they often act as epileptic foci. Histologically, the normal cortical architecture is effaced by collections of large bizarre cells that are haphazardly arranged within a gliotic background. These bizarre-looking cells have a amphophilic to slightly eosinophilic homogeneous cytoplasm and well-defined cytoplasmic borders. They also have stout processes, peripheral vacuolation, eccentric nuclei and prominent nucleoli. It is often difficult to determine whether they are glial cells or neurons by routine histological criteria. These cells often express tuberin, vimentin, nestin and some may express GFAP and neurofilament.

Subependymal nodules are often described macroscopically as 'candle-gutterings' which consist of variably sized nodules that protrude into the ventricle, most often the lateral ventricle. Histologically, they are composed of large, bizarre cells with pale amphophilic cytoplasm and large nuclei, often more than one in each cell. Scattered clusters of primitive neural cells may be seen in neonates. Subependymal nodules are often calcified. Vimentin is usually strongly expressed and, since both GFAP and NFP are also expressed in an irregular manner, it is difficult to define these cells as glial cells or neuronal cells. In contrast to the bizarre cells in cortical tubers, cells in the white matter and subependymal nodules express tuberin weakly or not at all. In some cases, these nodules continue to expand and give rise to a true neoplasm called 'subependymal giant cell astrocytoma' (**Figure 40**). Although this tumour is called an astrocytoma, it displays both glial and neuronal phenotypes and should more correctly be simply termed 'subependymal giant cell tumour.'

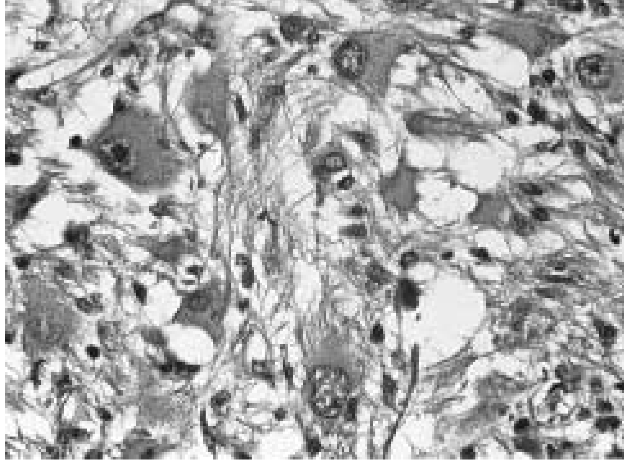


Figure 40 Large atypical cells characteristic of subependymal giant cell tumour. Note the fibrillary component and large cells with a resemblance to ganglion cells.

Table 7 Diagnostic criteria for neurofibromatosis

Neurofibromatosis 1

The criteria for making a diagnosis of NF1 are met if a patient has two or more of the following:

1. Six or more café-au-lait macules that have a maximum diameter over 5 mm in prepubertal patients and over-15 mm in postpubertal patients
2. Two or more neurofibromas of any type or one plexiform neurofibroma
3. Freckling in the axillary or inguinal regions
4. Optic nerve glioma
5. Two or more Lisch nodules (iris hamartomas)
6. A characteristic osseous lesion, such as sphenoid wing dysplasia or thinning of the long bone cortex with or without pseudoarthrosis and deformity of long bone
7. A first-degree relative (i.e. parent, sibling or offspring) with NF1 by the above criteria

Neurofibromatosis 2

The criteria for making a diagnosis of NF2 are met if a patient has one of the following:

1. Bilateral vestibular schwannomas
2. A first-degree relative with NF2 and either
   - Unilateral vestibular schwannoma or
   - Two of the following: neurofibroma, meningioma, schwannoma, glioma, juvenile posterior lenticular opacity

(Adapted from the National Institutes of Health Consensus Development Conference, 1988, Neurofibromatosis. *Archives of Neurology*, **45**, 575–578.)

Neurofibromatosis Type 1 (NF1)

NF1, also known as von Recklinghausen disease and peripheral neurofibromatosis, is a pleiotropic congenital multiple dysplasia syndrome characterized by multifocal hyperplasia and neoplasia in the supportive tissue throughout the entire nervous system. It is not a homogeneous syndrome and many clinical variants exist. The *NF1* gene is located on 17q11.2 and it encodes neurofibromin, a classical tumour-suppressor gene. Whole gene deletion, single- and multiple-exon deletions and insertions comprise about 50% of all the mutations.

The syndrome is defined by clinical criteria (**Table 7**). Clinical manifestations of NF1 are diverse (**Table 8**) and only those pertinent to the nervous system are discussed. The cardinal sign of NF1 is multiple neurofibromas. Some unusual variants of neurofibromas such as dermal neurofibroma and plexiform neurofibroma are characteristically seen in NF1 patients. Malignant transformation of neurofibromas in NF1 patients most frequently occurs in a preexisting plexiform neurofibroma and they usually do not occur until middle or late adult age. The transformation may occur as a progression in anaplasia over a period of years. The reported incidence of malignant transformation varies widely but a rate of 3–5% may be a good estimate. The resultant malignant tumour is usually a malignant peripheral nerve sheath tumour (MPNST).

Table 8 Manifestations of NF1

Skin	Café-au-lait macules
Neurofibromas	Plexiform
	Dermal
	Nodular
Sarcomas	Malignant peripheral nerve sheath tumour
	Rhabdomyosarcoma
	Triton tumour
Neoplastic CNS lesions	Optic tract/hypothalamic glioma
	Brainstem glioma
	Astrocytoma
	Anaplastic astrocytoma
	Glioblastoma
Non-neoplastic CNS conditions	Unidentified bright objects (UBO) on MRI
	Learning problem
	Epilepsy
	Neuropathy
	Hydrocephalus
	Hamatomatous glial proliferation
Osseous lesions	Sphenoid wing dysplasia
	Pseudoarthrosis
	Macrocephaly
	Scoliosis
Ocular	Lisch nodules
Other tumours	Phaeochromocytoma
	Carcinoid
	Juvenile chronic myeloid leukaemia

NF1 patients are also prone to develop gliomas. Those which develop in the optic nerve, hypothalamic region and brainstem are usually pilocytic astrocytomas and may remain static for many years and, in fact, some may even regress. Astrocytomas that arise in the cerebral and cerebellar hemispheres, however, may progress to anaplastic astrocytomas and glioblastoma multiforme.

Neurofibromatosis 2 (NF2)

Similarly to NF1, the cardinal changes of NF2 are multifocal hyperplasia and neoplasia in the supportive tissue throughout the entire nervous system. The elements being affected, however, are different.

NF2 is also an autosomal dominant disorder but its prevalence is only one-tenth of that of NF1. It is also defined by clinical criteria (**Table 7**). The gene is on chromosome 22q12 and encodes the protein merlin or schwannomin, most likely a tumour-suppressor gene. Nonsense and frameshift mutations are often associated with a more severe phenotype while missense mutations that preserve the carboxyl terminus of the protein result in milder phenotypes.

Schwannoma, particularly arising in the eighth cranial nerve, is the most frequent manifestation but malignant

transformation is extremely uncommon. Schwannomatosis, essentially pathological proliferation of Schwann cells, is often found in the dorsal spinal nerve roots and also in the perivascular space of the spinal cord. Meningiomas are seen in about half of NF2 patients and they are often multiple. Interestingly, they have a tendency to arise from the stroma of choroid plexus, an otherwise unusual site.

Meningioangiomatosis, a rare hamartomatous lesion of the leptomeninges, is far more commonly seen in NF2 patients than in general population. NF2 patients also have an increased incidence of glial tumours. Over 80% of them are in the spinal cord and of the rest 10% are in the medulla. Ependymomas are common and often occur as multiple intramedullary masses. Ependymal ectopias have also been described in NF2 patients and may represent the cell of origin of the ependymomas.

Von Hippel–Lindau Syndrome

Von Hippel–Lindau syndrome is an autosomal dominant heriditary syndrome characterized by multiple cysts and benign and malignant neoplasms involving the brain, retina, kidney, pancreas, adrenal glands, inner ear and other organs. The incidence is about 1 in 36 000 to 1 in 45 000 in the general population. This results from germ-line mutations of the *VHL* tumour-suppressor gene that is located on chromosome 3p25.3; a missense mutation is the most common. The *VHL* gene is involved in cell cycle regulation and angiogenesis.

Haemangioblastomas must be present in order to make a diagnosis of von Hippel–Lindau syndrome. Although haemangioblastomas can occur sporadically, they are seen in younger age groups when associated with von Hippel–Lindau syndrome. Most commonly, they occur in the cerebellum but they can also occur in other parts of the brain and in the retina. Other manifestations of von Hippel–Lindau syndrome include renal cysts and renal cell carcinoma, pancreatic cysts and islet cell tumours, phaeochromocytoma and endolymphatic sac tumour of the inner ear. Other organs may also be affected. The occurrence of phaeochromocytoma and renal cell carcinoma, interestingly, is correlated with specific mutations.

Basal Cell Naevus Syndrome

Basal cell naevus syndrome, also known as naevoid basal cell carcinoma syndrome, is transmitted in an autosomal dominant pattern and is typically caused by a germ-line mutation of the *PTCH* gene on chromosome 9q22.3. The majority of the mutations are frameshift or nonsense mutations that lead to truncated proteins. This gene is involved in the Sonic hedgehog signalling pathway. Mutations of the *PTCH* gene and Sonic hedgehog signal pathway are also related to holoprosencephaly.

Basal cell naevus syndrome is characterized by multiple basal cell carcinomas of the skin, odontogenic keratocysts, palmar and plantar dyskeratotic pits and CNS abnormalities including intracranial calcifications, macrocephaly and PNETs of the posterior fossa (medulloblastoma). Other less common clinical manifestations such as ophthalmic abnormalities, cardiac fibromas and cleft palate may also be present. Histological features of the PNETs that occur in patients with basal cell naevus syndrome are similar to those that occur sporadically, although the majority are the desmoplastic variant.

Cowden Disease

One of the major manifestations of Cowden disease is dysplastic gangliocytoma of the cerebellum (Lehmitte–Duclos disease). It is associated with a variety of mucocutaneous lesions that include verrucous skin changes, fibroma of the oral mucosa, multiple facial tichilemmomas, hamartomatous polyps and cancer of the colon and tumours of the breast and thyroid. It is an autosomsal dominant disorder and is frequently caused by a germ-line mutation of the *PTEN/MMCA1* gene on chromosome 10q23. In addition to dysplastic gangliocytoma of the cerebellum, Cowden disease can also be associated with heterotopic grey matter, hydrocephalus, mental retardation and seizures. (See also chapter *Inherited Predispositions to Cancer*.)

Cytogenetics and Fluorescent *In Situ* Hybridization (FISH)

Cytogenetics is the study of karyotypes and allows the detection of structural changes of chromosomes, namely deletion, inversion, translocation, trisomy and polysomy, and microsatellites. Since all chromosomes are visualized at the same time, it is a good screening test for chromosomal abnormalities. A variety of chromosomal abnormalities have been identified in human brain tumours using cytogenetics. Although cytogenetic techniques allow visualization of the entire karyotype, culture of tumour cells is required. In addition, they cannot detect microdeletions that are beyond resolution by banding techniques. Fluorescent *in situ* hybridization (FISH) can be used in paraffin-embedded tissue and in cultured cells. By using an appropriate fluorescent probe, a particular chromosome can be visualized or the visualization can be limited to a small region that allows the detection of microdeletions. Gene amplifcation in the form of double minutes can also be seen. FISH, however, cannot visualize all chromosomes at the same time and is more appropriate for confirmation of a specific chromosomal change.

Although many genetic abnormalities have been identified in brain tumours, most of them do not have a consistent pattern of chromosomal abnormalities. However, some genetic abnormalities are more often seen than others. One of the best examples is the association of isochromosome 17q with primitive neuroectodermal tumours of the posterior fossa (medulloblastoma) (**Figure 41; see colour plate section**). Another example is atypical teratoid rhabdoid tumour, which is associated with monosomy or

deletion of chromosome 22 (**Figure 42; see colour plate section**) in over 80% of cases, a useful diagnostic feature.

Molecular Genetics and Progression of Brain Tumours

Astrocytic tumours often progress from a low-grade tumour to anaplastic astrocytomas and finally become the highly malignant glioblastoma multiforme, the so-called secondary glioblastoma. This type of tumour progression is a multistep process driven by sequential alternations of specific genes or chromosomal loss. On the other hand, glioblastoma may also arise *de novo* without a precursor low-grade astrocytoma. Interestingly, the genetic changes in these tumours, also known as primary glioblastoma, are not the same as the secondary glioblastoma. Similar sequential genetics have also been observed in the progression of oligodendrogliomas. A variety of mutations have also been observed in other tumours.

PROGNOSTIC FACTORS

It cannot be overemphasized that tumours of the CNS are heterogeneous in clinical and biological behaviour. In general, the type of the tumour and its biological grade dictate the clinical outcome. In the WHO Classification (Kleihues and Cavenee, 2000), CNS tumours are separated into four grades. Grades I and II are low-grade tumours and that behave either in a benign fashion or as slow growing lesions; prolonged survival is possible. Grades III and IV are high-grade tumours, behave aggressively (**Table 9**) and are associated with short survival.

Location of a tumour in an inaccessible site such as the deeper parts of the cerebrum (such as basal ganglia and thalamus) or brainstem often precludes the benefits of surgical resection. The general condition of a patient also affects the clinical outcome. Younger patients with high-grade astrocytic tumours, in general, have a better survival than older patients. Other favourable prognostic factors for prolonged survival for patients with glioblastoma include long preoperative duration of symptoms, high Karnofsky Performance Status score, extensive surgical resection, use of postoperative radiation therapy and use of adjuvant chemotherapy.

Multiple prognostic indices are often present for one particular type of tumour, e.g. PNETs. Diagnosis at over 4 years of age is a favourable prognostic sign in PNETs arising from the cerebellum (medulloblastoma). In addition, limited extent of local disease, extensive surgical resection and no dissemination within the CNS are also good prognostic signs. The large cell variant of PNET has a worse prognosis; c-*myc* amplification and expression of GFAP are unfavourable. On the other hand, expression of Trk C receptor, a high-affinity neurotrophin receptor, is associated with improved survival.

Table 9 WHO grades of different brain tumours

Grade I
Pilocytic astrocytoma
Subependymoma
Choroid plexus papilloma
Ganglioglioma
Desmoplastic ganglioglioma/astrocytoma (DIG/DIA)
Dysembryoplastic neuroepithelial tumour (DNET)
Paraganglioma
Schwannoma
Neurofibroma
Perineuroma
Meningioma (variant dependent)
Grade II
Diffuse astrocytoma
Pleomorphic xanthoastrocytoma
Oligodendroglioma
Oligoastrocytoma
Ependymoma
Ganglioglioma
Central neurocytoma
Pineocytoma
Meningioma (variant dependent)
Grade III
Anaplastic astrocytoma
Anaplastic oligodendroglioma
Anaplastic oligodendroglioma
Anaplastic ependymoma
Choroid plexus carcinoma
Anaplastic ganglioglioma
Malignant peripheral nerve sheath tumour (MPNST)
Meningioma (variant dependent)
Anaplastic meningioma
Grade IV
Glioblastoma
Giant cell glioblastoma
Gliosarcoma
Medulloepithelioma
Primitive neuroectodermal tumour (PNETs)
Atypical teratoid/rhabdoid tumour (ATRT)
Malignant peripheral nerve sheath tumour (MPNST)

Source: Kleihues and Cavenee, 2000, *World Health Organisation Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System* (International Agency for Research on Cancer, Lyon).

OVERVIEW OF PRESENT CLINICAL MANAGEMENT

Surgery, radiation therapy and chemotherapy are the three cornerstones of treatment. Several clinical pathological parameters must be considered before a treatment plan is formulated for an individual patient. The age of the patient, preoperative medical condition, presence of other systemic disease and malignant disease and location of the tumour must all be considered. The type of tumour, as determined by pathological examination of biopsy material, dictates the direction of further treatment. Biologically benign CNS

and PNS tumours are treated surgically unless they are located in an inoperable location. Malignant tumours are often treated with chemotherapy and radiation therapy in addition to surgery. Most primary brain tumours are candidates for surgical treatment except that haematopoietic tumours are often treated with radiation and chemotherapy alone without bulk tumour resection.

Maximum tumour resection and optimum preservation of neurological functions are the two most important treatment goals for patients with brain tumours that should be surgically treated. Tumours located around or within the brainstem, diencephalon and motor cortex are particularly troublesome because damage to these areas will lead to neurological catastrophe. For these tumours, stereotactic biopsy is the diagnostic procedure of choice for the determination of the nature of the lesion. It is a computer-assisted technique that allows three-dimensional coordinated localization of an intracranial lesion. Based on these coordinates, a needle is inserted to obtain a small amount of tissue for pathological examination without producing excessive damage to the critical surrounding brain structures. Tumours located in these difficult locations are generally treated with radiation therapy and chemotherapy alone.

Since the general principles of radiation therapy are discussed elsewhere, discussion here is limited to issues specific for the CNS. In general, toxic effects produced by irradiating the human brain are categorized temporally and pathologically into three major types. The first type are acute reactions that occur during radiation therapy. They are uncommon in patients receiving conventional daily dose fractionation (1.8–2 Gy per day) but are common in patients treated with 10 Gy or above as a single fraction. Subacute reactions (early-delayed reactions) are related to radiation injury sustained by oligodendrocytes and alterations in capillary permeability and present as transient demyelination. The subacute reactions are characterized by a transient syndrome that includes somnolence, anorexia and irritability. The late-delayed reactions occur within several months or up to many years following treatment. They comprise a spectrum of clinical and pathological conditions involving multiple mechanisms. Clinical manifestations includes asymptomatic changes of the white matter and vasculature, change in cognition, hypothalamic–pituitary dysfunction, cranial neuropathy, disseminated necrotizing leucoencephalopathy, radiation necrosis and second malignancy due to radiation.

REFERENCES

Biegel, J. A., *et al*. (2000). Mutations of the INI1 rhabdoid tumor suppressor gene in medulloblastomas and primitive neuroectodermal tumors of the central nervous system. *Clinical Cancer Research*, **6**, 2759–2763.

Cairncross, J. G., *et al*. (1998). Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. *Journal of the National Cancer Institute*, **90**, 1473–1479.

Cavenee, W. K., *et al*. (2000). Turcot syndrome. In: Kleihues, P. and Cavenee, W. K. (eds), *World Health Organization Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System*. 238–239 (International Agency for Research on Cancer, Lyon).

Coons, S. W., *et al*. (1997). The prognostic significance of Ki-67 labeling indices for oligodendrogliomas. *Neurosurgery*, **41**, 878–884.

Daumas-Duport, C. (1993). Dysembryoplastic neuroepithelial tumours. *Brain Pathology*, **3**, 283–295.

Dehghani, F., *et al*. (1998). Prognostic implication of histopathological, immunohistochemical and clinical features of oligodendrogliomas: a study of 89 cases. *Acta Neuropathologica (Berlin)*, **95**, 493–504.

Dorfman, H. D. and Czerniak, B. (1998). Chordoma and related lesions. In: *Bone Tumors*. 974–1008 (Mosby, St. Louis).

Hahn, H., *et al*. (1996). Mutations of the human homolog of *Drosophila* patched in the nevoid basal cell carcinoma syndrome. *Cell*, **85**, 841–851.

Harris, N. L., *et al*. (1994). A revised European–American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*, **84**, 1361–1392.

Hamilton, S. R., *et al*. (1995). The molecular basis of Turcot's syndrome. *New England Journal of Medicine*, **332**, 839–847.

Kandt, R. S., *et al*. (1992). Linkage of an important gene locus for tuberous sclerosis to a chromosome 16 marker for polycystic kidney disease. *Nature Genetics*, **2**, 37–41.

Kleihues, P. and Cavenee, W. K. (2000). *World Health Organization Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System*. (International Agency for Research on Cancer, Lyon).

Lang, F. F., *et al*. (1993). Central nervous system gangliogliomas. Part 2. Clinical outcome. *Journal of Neurosurgery*, **79**, 867–873.

Latif, F., *et al*. (1993). Identification of the von Hippel–Lindau disease tumor suppressor gene. *Science*, **260**, 1317–1320.

Maddock, I. R., *et al*. (1996). A genetic register for von Hippel–Lindau disease. *Journal of Medical Genetics*, **33**, 120–127.

Malkin, D., *et al*. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, **250**, 1233–1238.

Onda, K., *et al*. (1996). Comparison of bromodeoxyuridine uptake and MIB 1 immunoreactivity in medulloblastomas determined with single and double immunohistochemical staining methods. *Journal of Neurooncology*, **29**, 129–136.

Packer, R. J., *et al*. (1999). Outcome for children with medulloblastoma treated with radiation and cisplatin, CCNU, and vincristine chemotherapy. *Journal of Neurosurgery*, **81**, 690–698.

Paraf, F., *et al*. (1997). Brain tumor–polyposis syndrome: two genetic diseases? *Journal of Clinical Oncology*, **15**, 2744–2758.

Pollack, I. F. and Mulvihill, J. J. (1997). Neurofibromatosis 1 and 2. *Brain Pathology*, **7**, 823–836.

Pratt, C. B., *et al*. (1994). Outcome for patients with constitutional 13q chromosomal abnormalities and retinoblastoma. *Pediatric Hematology and Oncology*, **11**, 541–547.

Ritter, A. M., *et al*. (1998). Ependymomas: MIB-1 proliferation index and survival. *Journal of Neurooncology*, **40**, 51–57.

Robinson, S. and Cohen, A. R. (2000). Cowden disease and Lhermitte–Duclos disease: characterization of a new phakomatosis. *Neurosurgery*, **46**, 371–378.

Ron, E., *et al*. (1988). Tumors of the brain and nervous system after radiotherapy in childhood. *New England Journal of Medicine*, **319**,1033–1039.

Schiffer, D., *et al*. (1991). Ependymoma: internal correlations among pathological signs: the anaplastic variant. *Neurosurgery*, **29**, 206–210.

Schild, S. E., *et al*. (1996). Histologically confirmed pineal tumors and other germ cell tumors of the brain. *Cancer*, **78**, 2564–2571.

Seizinger, B. R., *et al*. (1987). Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell*, **49**, 589–594.

Shaw, E. G., *et al*. (1992). Oligodendrogliomas: the Mayo Clinic experience. *Journal of Neurosurgery*, **76**, 428–434

Shaw, E. G., *et al*. (1994). Mixed oligoastrocytomas: a survival and prognostic factor analysis. *Neurosurgery*, **34**, 577–582.

Shimada, H., *et al.* (1999). The International Neuroblastoma Pathology Classification (the Shimada system). *Cancer*, **86**, 364–372.

Sopta, M., *et al*. (1992). The retinoblastoma protein and the cell cycle. *Seminars in Cancer Biology*, **3**, 107–113.

Sutphen, R., *et al*. (1999). Severe Lhermitte–Duclos disease with unique germline mutation of PTEN. *American Journal of Medical Genetics*, **82**, 290–293.

Tachibana, I., *et al*. (2000). Investigation of germline PTEN, p53, p16(INK4A)/p14(ARF), and CDK4 alterations in familial glioma. *American Journal of Medical Genetics*, **92**, 136–141.

Trofatter, J. A., *et al*. (1993). A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. *Cell*, **72**, 791–800.

Vajtai, I., *et al*. (1996). MIB-1 immunoreactivity reveals different labelling in low-grade and in malignant epithelial neoplasms of the choroid plexus. *Histopathology*, **29**, 147–151.

van Slegtenhorst, M., *et al*. (1997). Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science*, **277**, 805–808.

Von Deimling, A., *et al*. (2000). Neurofibromatosis type 1. In: Kleihues, P. and Cavenee, W. K. (eds), *World Health Organization Classification of Tumours – Pathology and Genetics, Tumours of the Nervous System*. 216–218 (International Agency for Research on Cancer, Lyon).

Vorechovsky, I., *et al*. (1999). The patched/hedgehog/smoothened signalling pathway in human breast cancer: no evidence for H133Y SHH, PTCH and SMO mutations. *European Journal of Cancer*, **35**, 711–713.

FURTHER READING

Adams, E. D., *et al*. (1997). Intracranial neoplasms and paraneoplastic disorders. In: *Principles of Neurology*, 6th edn. 642–698 (McGraw-Hill, New York).

Aicardi, J. (1998). Tumors of the central nervous system and other space-occupying lesions. In: *Diseases of the Nervous System in Childhood*, 2nd edn. 491–533. (Cambridge University Press, Cambridge).

Berger, M. S. and Wilson, C. B. (1999). The Gliomas (W. B. Saunders, Philadelphia).

Bigner, D. D., *et al*. (eds) (1998). *Russel and Rubinstein's Pathology of Tumours of the Nervous System*, 6th edn (Arnold, London).

Enzinger, F. M. and Weiss, S. W. (2001). *Soft Tissue Tumors*, 4th edn (Mosby, St. Louis).

Greenberg, H. S., *et al*. (1999). *Brain Tumors* (Oxford University Press, Oxford).

Jacobson, M. (1991). *Developmental Neurobiology*, 3rd edn. (Plenum Press, New York).

Norman, M. G., *et al*. (1995). Embryology of the central nervous system. In: *Congenital Malformations of the Brain*. 9–51 (Oxford University Press, Oxford).

Pannese, E. (1994). *Neurocytology. Fine structure of Neurons, Nerve Processes, and Neuroglial Cells*. 156–168 (George Thieme, Stultgant).

Russell, D. S. and Rubinstein, L. J. (1989). *Pathology of Tumours of the Nervous System* (Williams and Wilkins, Baltimore).

Sidman, R. L. and Rakic, P. (1982). Development of the central nervous system. In: Hayemarker, W. and Adams, R. D. (eds), *Histology and histopathology of the Central Nervous System*, Vol. 1, 3–145 (Charles C. Thomas, Springfield).

Zülch, K. J. (1986). *Brain Tumors. Their Biology and Pathology*, 3rd edn (Springer, Berlin).

Websites

Neuroanatomy and neuropathology on the internet: http://www.neuropat.dote.hu/document.htm

Chapter 52

Eye and Ocular Adnexa

Arun Jain and Frederick A. Jakobiec
Massachusetts Eye and Ear Infirmary, Boston, MA, USA

CONTENTS

- Normal Development and Structure
- Tumour Pathology

NORMAL DEVELOPMENT AND STRUCTURE

The globe and the ocular adnexa represent a unique compartment, where terminal differentiation from all three germinal layers of the embryo has given rise to very specialized structures not found elsewhere in the body. Owing to this extraordinary collection of different tissues, one encounters tumours unique to this location as well as others found more commonly in distant parts of the body.

The eye is a sensory end organ that develops from the interaction of the primordial optic vesicle with the overlying ectoderm during the fourth week of gestation. The optic vesicle, one on each side, is an out pouching from the rostral end of the neural tube. The neural tube originates from the neural plate, which is a derivative of embryonic ectoderm and precursor of the entire central nervous system. The lens and most of the cornea develop from the inductive interaction of the optic vesicle with the surface ectoderm. The neural crest contributes extensively to the connective tissue of the head and neck region, which is due to the absence of paraxial somites in the head and neck region. These connective tissue contributions of the neural crest are collectively referred to as mesoectoderm or ectomesenchyme. Therefore, in the orbit, the fibrous and fibroadipose tissue, meninges of the optic nerve, sclera and episclera, vascular pericytes and striated extraocular muscle satellite cells, peripheral nerve cellular elements, and osteocytes and cartilaginous elements are all progeny of neural crest anlage. **Table 1** demonstrates the origin of the various structures of the eye from different germinal layers.

The eye is composed of three layers: the outermost sclera with its anterior extension as the cornea, the uveal tract and the innermost layer being the retina. The six extraocular muscles are composed of striated muscle fibres and aid in the fine and gross control of eye movements. Four of these muscles arise from the orbital apex in a concentric fashion surrounding the annulus of Zinn. Intraocularly the iris and ciliary body are composed of smooth muscle and are innervated by autonomic nerve fibres. The retina is a complex multilayered structure, innermost layer of which is formed by the axons of the ganglion cell layer (nerve fibre layer) and the outmost

Table 1 Origin of different structures from germinal layers

Ectoderm		Mesoderm	Endoderm	Neural crest
Surface ectoderm	Neuroectoderm			
Anterior cornea	Neurosensory retina	Vascular endothelium	None	Anterior chamber
Eyelids	Retinal pigment epithelium	Extraocular muscles		(Posterior cornea)
(Holocrine glands)	Pars plana epithelium	Corneal stroma		(Trabecular meshwork)
(Apocrine glands)	(Secondary vitreous)			(Iris stroma)
(Eccrine glands)	(Lens zonules)			Uveal tract
(Pilosebaceous units)	Ciliary epithelium			(Dendritic melanocytes)
Conjunctiva	Iris pigment epithelium			Sclera
(Lachrymal glands)	(Dilator muscle)			Meninges
(Mucous glands)	(Sphincter muscle)			Orbital soft tissues
Crystalline lens	Optic nerve			Facial bones
				Trochlea

layer is formed by the photoreceptors. The neuro-sensory retina and the retinal pigment epithelium (RPE) continue anteriorly as the pigmented and non-pigmented epithelium of the ciliary body and iris.

The conjunctiva is lined by non-keratinized stratified squamous epithelium about 4–5 cell layers thick and has numerous goblet cells. Accessory lachrymal glands of Krause and Wolfring are present in the superior fornix of the conjunctiva. At the mucocutaneous junction of the eyelid, where conjunctival epithelium changes into keratinized stratified squamous epithelium of the skin, there are openings for about 30 sebaceous glands known as meibomian glands in the upper lid and 20 in the lower lid. These glands are embedded in a matrix of dense connective tissue known as tarsus. Anterior to this row of meibomian glands is a row of cilia (eye lashes) which have their own associated sebaceous glands known as glands of Zeiss and apocrine glands known as glands of Moll.

The orbit is compartmentalized into intraconal and extraconal spaces by the intermuscular septa of the extraocular muscles. The intraconal space is by far the larger of the two and is further divided by numerous thin fibrous septa into spaces that are all filled with fat. Through these compartments traverse the nerves and vessels supplying the extraocular muscles and the autonomic nerves and vessels to the choroid, ciliary body and iris. The ophthlamic artery, which later becomes the central retina artery, supplies the optic nerve and the retina. The orbit and the intraocular tissues are devoid of lymphatics, which serves to explain the metastatic pattern of some tumours. Conjunctiva and eyelids, on the other hand, are richly supplied with lymphatics.

The lachrymal gland is present in the supero-lateral recess of the orbit and is divided into two lobes by the lateral extension of the levator palpebrae superioris. The fine ductules of the lachrymal gland then open into superior fornix for the egress of tears. The tears then drain though two canaliculi into the lachrymal sac present in the lachrymal fossa formed by the lachrymal bone and frontal process of the maxilla.

TUMOUR PATHOLOGY

Eye and adnexal tumours can be subdivided into intraocular tumours, adnexal tumours of the eyelid, lachrymal sac and conjunctiva and orbital tumours. There is considerable overlap in this classification scheme, since intraocular tumours can sometimes invade the orbit and adnexal tumours on rare occasions can extend intraocularly.

Eyelid and conjunctival tumours have various counterparts in the skin from other parts of the body. Primary acquired melanosis, although analogous to lentigo maligna in the skin, has different prognostic and therapeutic implications. One of the most important malignancies confined almost exclusively to the skin of the eyelids is sebaceous cell carcinoma.

The differential diagnosis of orbital tumours in children is strikingly different from that of adults. Capillary haemangioma and dermoid cysts are probably the most frequent orbital tumours encountered in children, whereas cavernous haemangioma is more often seen in adults and lymphoproliferative lesions are confined almost exclusively to the sixth and seventh decades of life.

The most important intraocular tumours are retinoblastoma in children and uveal melanoma in adults. However, there is a long list of other intraocular tumours that can arise *de novo* or from metastasis. In the following sections, emphasis will be placed on tumours that are unique to the eye and ocular adnexa.

Epidemiology

Eyelid and conjunctival tumours usually affect the elderly. The median age at diagnosis of sebaceous carcinoma of the eyelid is 64 years and 60–70% of them are females.

Retinoblastoma is the most common intraocular tumour of childhood and the most common tumour of the retina. It is a rare malignant tumour with a prevalence of about 1 in 23 000 live births in the UK, 1 in 16 000 in The Netherlands and 1 in 20 000 in Japan (Albert and Jakobiec, 2000a)

Uveal melanoma is by far the most clinically important intraocular malignancy in adults, owing to its risk of metastasis and subsequent death, even though uveal metastasis has a higher incidence when autopsy eyes are studied for occult intraocular tumours. It is the most common primary intraocular malignancy, occurring in approximately 6–7 cases per million people in the USA.

Aetiology

A variety of known and unknown factors contribute to neoplastic transformation of cells in the eye and ocular adnexa. The risk associated with UV exposure in the development of skin tumours is well established. Conjunctival lesions such as pterygia, pinguecula, and actinic keratosis are also related to prolonged exposure to UV light as documented by their higher frequency in more tropical areas.

Conjunctival papillomas and dysplasias are akin to skin papillomas, and human papillomaviruses 16 and 18 have been implicated in the pathogenesis of some of these lesions.

The presence of genetic mutations is well established in retinoblastoma. The role of specific gene alterations in the pathogenesis of uveal melanoma is less well defined than it is in retinoblastoma (see section on Molecular Genetic Findings).

Certain lesions in the conjunctiva and the uveal tract predispose to the development of malignant melanoma.

These include congenital melanosis and naevi. Congenital melanosis can involve only the sclera (ocular melanocytosis) or it can affect both the sclera and the skin of the lids, in which case it is referred to as Naevus of Ota (oculodermal melanocytosis) (**Figure 1; see colour plate section**). The risk of progression to malignant melanoma in congenital melanosis is primarily in the uveal tract. Naevi, however, in the conjunctiva or the uvea can give rise to malignant melanoma in either location.

The role of racial pigmentation is well established in a variety of skin, conjunctival and uveal malignancies. Individuals with lightly pigmented skin also have less melanin in the RPE and choroid, which leads to less protection from the harmful UV rays of the sun. The incidence of uveal malignant melanoma is 8–9 times higher in whites than in blacks in the USA.

Radiation used for the treatment of childhood retinoblastoma has led to an increase in the development of sarcomas in the field of radiation. Patients with the heritable form of retinoblastoma, however, have an increased incidence of sarcomas within and outside the field of radiation, with osteogenic sarcoma the most prevalent.

Screening and Prevention

Better understanding of the pathophysiology of certain tumours of the eye and ocular adnexa has greatly improved the prognosis of these tumours owing to the increased emphasis on screening and prevention. The risk of development of retinoblastoma in the siblings and progeny of patients with retinoblastoma can be accurately assessed and appropriate screening can be deployed to detect early lesions. Patients with uveal and conjunctival naevi should be evaluated on a yearly basis to detect progression of these lesions to malignant melanoma.

The association of certain tumours of the eye and ocular adnexa with other systemic neoplasms (both benign and malignant) has led to an increased awareness of these genetic syndromes and underscores the importance of the ophthlamologist in making an early diagnosis. Some of these will be discussed in the sections below.

Gross/Histopathology/ Preinvasive Lesions/ Ultrastructure/ Immunohistochemistry

Tumours of the Eyelid and Conjunctiva

Tumours of the eyelid skin are the same as found in other parts of the body and are detailed elsewhere in this book. Basal cell carcinoma of the eyelids is the most frequent of the skin malignancies followed by squamous cell carcinoma (Albert and Jakobiec, 2000b). The lower lid is the most frequently affected area followed by the medial canthus and then the upper lid and lateral canthus. The basal cell naevus (Gorlin–Goltz) syndrome is characterized by the presence of multiple basal cell epitheliomas on the face, and less frequently on the trunk, neck and axilla. The tumours may not become invasive for many years. Associated anomalies include pitted areas in the skin of the palms and soles, odontogenic cysts of the jaw, spina bifida occulta, rib anomalies, skull anomalies, cleft lip and palate and hypogonadism in males. It is inherited in an autosomal dominant fashion with variable penetrance. The specialized glands of Moll, Zeiss, Wolfring and Krause can give rise to adenocarcinoma or a variety of benign lesions such as hydrocystoma, syringoma, eccrine acrospiroma and pleomorphic adenoma. Benign tumours of hair follicle origin include trichoepithelioma, trichofolliculoma, trichilemmoma and pilomatrixoma (calcifying epithelioma of Malherbe). Multiple trichilemmomas are seen in Cowden's syndrome. Melanocytic lesions of the eyelid skin and conjunctiva such as naevi and melanomas are also frequently encountered.

Pinguecula, pterygium and actinic keratosis result from UV damage to the conjunctiva. These lesions are exclusively found in the interpalpebral conjunctiva exposed to sunlight. The characteristic hallmark of all three is the elastotic degeneration present in the substantia propria of the conjunctiva.

Dysplasia and conjunctival intraepithelial neoplasia (carcinoma-*in-situ*) may occur anywhere on the conjunctival surface, but most often starts at the limbus (**Figure 2**). Most lesions have an opalescent papillary rather than a leukoplakic appearance due to lack of keratinization observed microscopically. A benign condition but often confused for malignancy seen in certain Haliwa Indian tribes of North Carolina is benign hereditary intraepithelial dyskeratosis appearing as patches of thickened and injected plaques of conjunctiva on the bulbar surface.

Other rare tumours include choristomas of the conjunctiva and Merkel cell tumours and myxomas of the

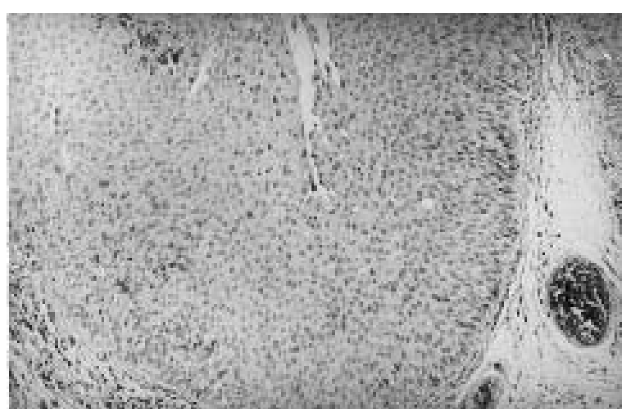


Figure 2 Conjunctival intraepithelial neoplasia demonstrating full thickness involvement of the epithelium with malignant cells.

eyelid. The latter are associated with the Carney complex, a genetic syndrome comprised of cardiac and cutaneous myxomas, spotty skin pigmentation and endocrinopathy. Choristomas such as the limbal dermoid are associated with lid coloboma in the Goldenhar syndrome. Two categories of tumours deserve special focus owing to their exclusive localization to the ocular adnexa and often challenging management. These are the sebaceous cell carcinoma of the eyelid and primary acquired melanosis of the conjunctiva.

Sebaceous Gland Tumours

Sebaceous adenomas are rare and usually solitary lesions not exceeding 1 cm in size. Histologically, they are composed of sebaceous lobules that show an orderly maturation from germinative cells located on the periphery towards the more central sebaceous cells. Muir–Torre syndrome, an autosomal dominant cause of familial cancer, is an association of multiple sebaceous gland neoplasms, other cutaneous neoplasms and multiple visceral carcinomas, especially of the colon. Often, only a single sebaceous adenoma of the eyelid may be associated with internal malignancy.

Most sebaceous carcinomas of the eyelid arise in the meibomian (tarsal) glands, followed by the glands of Zeiss, and, less frequently, sebaceous glands in the caruncle or the skin of the eyebrow. A distinctive clinical feature of many sebaceous gland carcinomas is a persistent unilateral blepharo-conjunctivitis, referred to as the 'masquerade syndrome' (**Figure 3; see colour plate section**). They can also arise in a nodular fashion at the lid margin where they appear as waxy yellowish lesions, sometimes simulating a chalazion. Histologically, the most common pattern is lobular with the degree of differentiation determining the type of cells present. The lobules exhibit basaloid features but lack the peripheral palisading of basal cell carcinoma. Well-differentiated tumours usually have areas of easily identifiable sebaceous cells with the characteristic finely vacuolated cytoplasm. Poorly differentiated neoplasms have anaplastic cells with hyperchromatic nuclei and typical mitotic figures. Frozen section and oil red O stains for lipid may be necessary to establish the diagnosis.

Sebaceous carcinoma can spread by either intraepithelial or direct extension. Intraepithelial spread can occur in a carcinoma-*in-situ* like fashion with the entire epithelium being replaced by neoplastic cells, or individual cells can migrate through intercellular spaces to distant areas in the eyelid skin or conjunctiva (**Figure 4**). This latter phenomenon is termed pagetoid spread, resembling Paget disease of the breast.

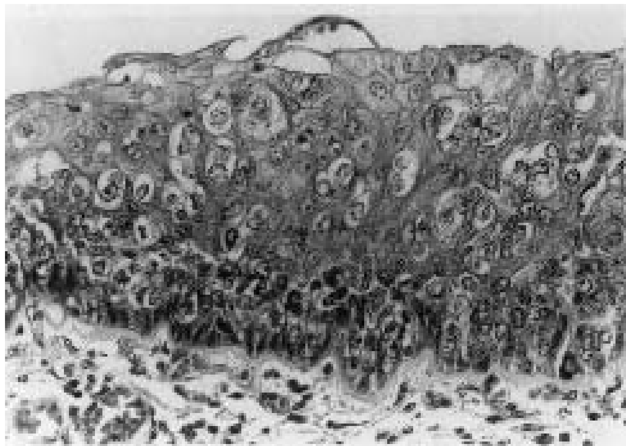


Figure 4 Pagetoid spread of sebaceous carcinoma (cells with clear cytoplasm) in the conjunctival epithelium (courtesy AFIP).

Primary Acquired Melanosis (PAM)

PAM is a unilateral neoplastic melanocytic proliferation within the conjunctival epithelium, observed most often in white patients. They appear as flat areas of increased yellow–brown pigmentation in middle age, mostly on the bulbar conjunctiva, which does not fade towards the fornices like racial melanosis (**Figure 5; see colour plate section**). When the palpebral conjunctiva near the lid margin is involved, the melanosis often extends on to the adjacent epidermis. Elevated lesions developing in PAM is usually an indication of malignancy, i.e. melanoma.

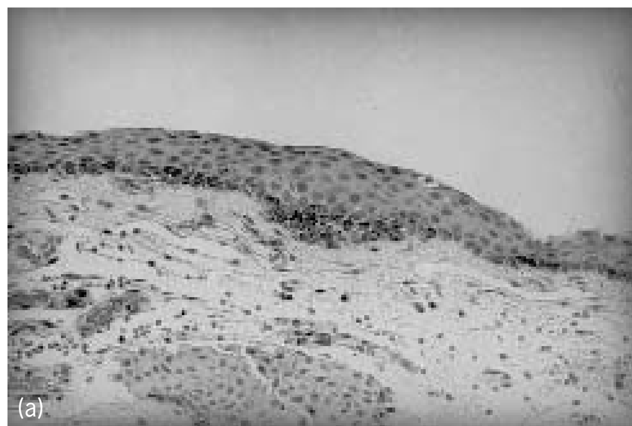
Histologically, their appearance can be variable. There may be small polyhedral cells with no atypia, spindle-shaped cells with moderate atypia or large highly atypical epithelioid cells. The pattern of invasion of the conjunctival epithelium can vary from involvement of the basal layers to pagetoid extension or even full thickness replacement by atypical melanocytes resembling melanoma-*in-situ* (**Figure 6**).

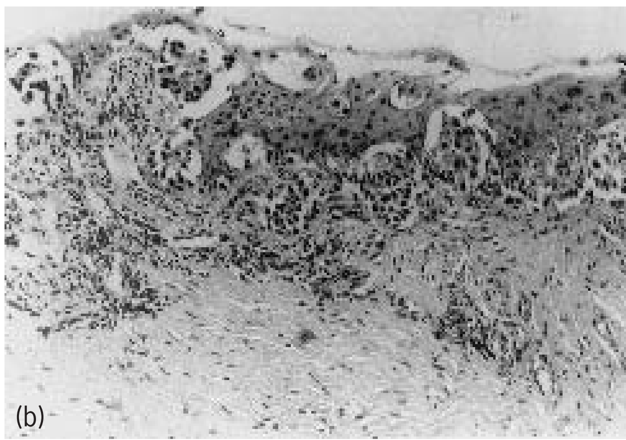
Tumours of the Retina

Retinoblastoma

Five growth patterns are recognized for intraocular retinoblastoma: endophytic, exophytic, mixed endophytic–exophytic, diffuse infiltrating and complete spontaneous regression. The tumour can either arise from one location or be multicentric. Endophytic retinoblastomas grow mainly from the inner nuclear layer of the retina into the vitreous. Endophytic tumours can shed tumour cells into the vitreous, as they become large and friable, where they grow into separate tiny spheroidal masses that appear as cotton balls. Tumour cells in the vitreous may seed on to the inner surface and invade into the retina. It is important to distinguish multicentric retinoblastoma from retinal seeding because the presence of multiple tumours indicates a germ-line mutation. This distinction is frequently difficult or impossible to make.

Exophytic retinoblastomas grow from the outer nuclear layer towards the choroid, producing first an elevation and





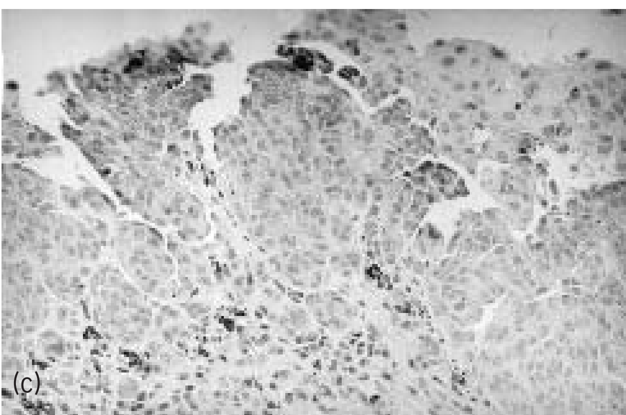


Figure 6 (a) Primary acquired melanosis without atypia. The melanotic hyperplasia is confined to the basilar layer. (b) Primary acquired melanosis with severe atypia. Nests and individual epithelioid melanocytes extend throughout the thickness of the epithelium in a pagetoid fashion (courtesy AFIP). (c) Primary acquired melanosis with invasive melanoma. Besides full thickness involvement of the epithelium with melanoma cells, some have invaded the substantia propria through the basement membrane.

then a detachment of the retina (**Figure 7; see colour plate section**). On ophthalmoscopic examination, the tumour is seen through the retina with vessels coursing over it (**Figure 8; see colour plate section**). Tumour cells can then seed the outer retinal surface or invade into the RPE and Bruch's membrane with subsequent infiltration of the choroid. From the choroid, the tumour cells can escape along the ciliary vessels and nerves into the orbit and conjunctiva and then gain access to blood vessels and lymphatics.

Mixed endophytic–exophytic tumours are the most common form of presentation, especially among larger tumours. They have features of both endophytic and exophytic growth. Diffuse infiltrating retinoblastomas often present the greatest challenge in clinical diagnosis, but fortunately are the least common of all forms. These tumours grow diffusely within the retina without thickening it much. Tumour cells are often shed into the vitreous, which can reach the anterior chamber to create a pseudohypopyon. Because of the absence of a mass, this type of retinoblastoma masquerades as a retinitis, vitritis or Toxocara endophthalmitis.

Complete spontaneous regression is believed to occur more frequently in retinoblastoma than in any other malignant neoplasm. Typically, there is a severe inflammatory reaction followed by phthisis bulbi. The mechanisms by which regression occurs are unknown.

Table 2 illustrates the differential diagnosis of retinoblastoma.

Histologically, retinoblastomas are essentially malignant neuroblastic tumours that may arise in any of the nucleated retinal layers. The predominant cell has a large basophilic nucleus of variable size and shape and scanty cytoplasm. Mitotic figures are typically numerous. The tumour cells have a striking propensity to outgrow their blood supply. Especially in large tumors, cuffs of cells surrounding central dilated vessels form areas of viable tumour that have a characteristic appearance (**Figure 9**). The tumour cells undergo ischaemic necrosis if they are displaced more than 90–110 μm from the vessel. A cuff thickness of approximately 100 μm represents the approximate distance that oxygen can diffuse before it is completely consumed in rapidly growing neoplasms.

The ability of retinoblastomas rapidly to outstrip their blood supply leads to areas of coagulative necrosis. Foci of calcification are frequently seen in these areas. The DNA liberated from necrotic cells can become absorbed preferentially in the walls of blood vessels and by the internal limiting membrane of the retina, giving a deep blue (haematoxylinophilic) or Feulgen-postive stain to these tissues.

Retinoblastomas characteristically form rosettes and flurettes, which are attempts at photoreceptor differentiation. Flexner–Wintersteiner rosettes are highly specific for retinoblastomas, but are also observed in pineoblastoma and medulloepithelioma. Even though these rosettes represent areas of differentiation by the

Table 2 Differential diagnosis of retinoblastoma: conditions simulating retinoblastoma based on clinical diagnoses in a study consisting of 500 patients (From Shields *et al.*, 1996, *Archives of Ophthalmology*, **114**, 1330–1338.)

Condition	%
PHPV	27.8
Coat disease	16.0
Ocular toxocariasis	15.6
Retinopathy of prematurity	4.7
Combined hamartoma	4.2
Coloboma	4.2
Vitreous haemorrhage	3.8
Astrocytic hamartoma	2.8
Familial exudative vitreoretinopathy	2.4
Idiopathic retinal vascular hypoplasia	1.9
Rhegmatogenous retinal detachment	1.9
X-linked retinoschisis	1.9
Medulloepithelioma	1.9
Congenital cataract	1.9
Retinal capillary haemangioma	1.4
Cicumscribed choroidal haemangioma	1.4
Diffuse choroidal haemangioma	1.4
Peripheral uveoretinitis	1.4
Toxoplasmic retinitis	0.9
Idiopathic endophthalmitis	0.9
Norrie disease	0.5
Incongentia pigmenti	0.5
Optic nerve dysplasia	0.5

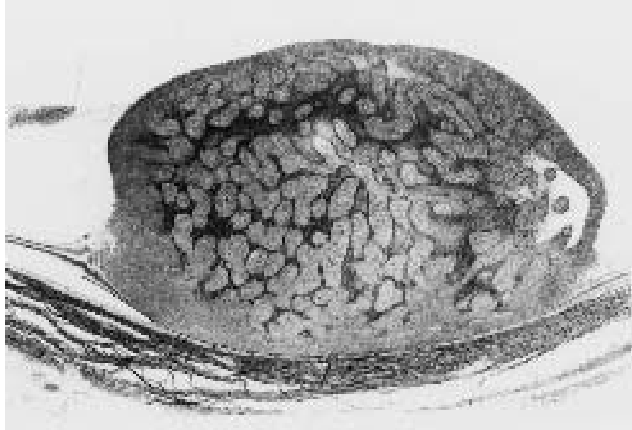


Figure 9 Retinoblastoma. Viable cells forming sleeves around blood vessels with areas of necrosis.

tumour, they are composed of malignant cells and occur in areas of mitotic activity. The typical Flexner–Wintersteiner rosette (**Figure 10**) is composed of tall cuboidal cells that encircle an apical lumen. The apical ends of the cuboidal cells are held together by terminal bars and the cells may have apical cytoplasmic projections into the lumen of the rosette. Electron microscopy has demonstrated that these projections represent primitive inner and

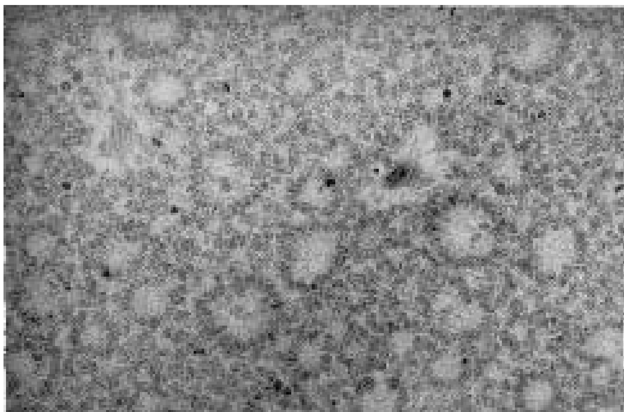


Figure 10 Flexner-Wintersteiner rosettes in retinoblastoma.

outer segments and, therefore, these tumour cells are attempting to form photoreceptor cells. Alcian blue stain reveals the presence of hyaluronidase-resistant glycosaminoglycans in the lumina of the rosettes that have similar staining characteristics to the extracellular matrix of rods and cones. Cells forming Flexner–Wintersteiner rosettes share several other ultrastructural features with retinal photoreceptors: zonula occludens that form a luminal limiting membrane analogous to the external limiting membrane of the retina, cytoplasmic microtubules, cilia with the $9+2$ pattern and lamellated membranous structures resembling the discs of rod outer segments. Immunohistochemical and lectin histochemical studies have also supported the concept that retinoblastomas arise from undifferentiated retinal cells that may differentiate into photoreceptor cells. Homer Wright rosettes are the other type of rosettes found in retinoblastomas, but are less common. They are found in a variety of neuroblastic tumours and are less specific for retinoblastoma. In these rosettes, the cells are not arranged about a lumen but send out cytoplasmic processes that form a tangle in the centre of the rosette. Finally, fleurettes probably represent the highest degree of differentiation of the tumour cells into photoreceptors. Fleurette represents a collection of benign appearing cells with long cytoplasmic processes that stain brightly with eosin. The cytoplasmic processes project through a fenestrated membrane and fan out like a bouquet of flowers (hence the term fleurette).

Extraocular extension and metastasis Retinoblastoma, if left untreated, has a propensity for rapid invasive growth typical of blastic tumours in children. The most common routes for local spread are through the optic nerve (**Figure 11**) and invasion of the choroid. Extraocular extension occurs via scleral canals or by massive replacement of the sclera.

Glial Tumours

Astrocytomas of the retina are rare and almost always benign tumours. The majority of them (over 50%) occur in

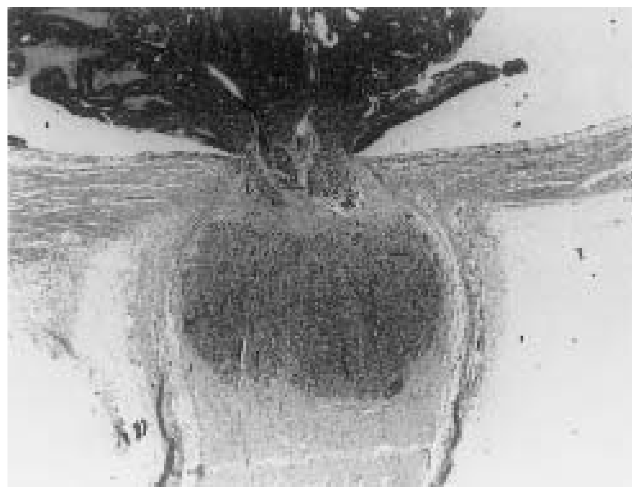


Figure 11 Retinoblastoma. Tumour is invading the optic nerve but not reaching the surgical margin (courtesy AFIP).

patients with tuberous sclerosis and a small proportion develops in patients with neurofibromatosis. They are also described as astrocytic hamartomas. On ophthalmoscopic appearance, early tumours appear flat and translucent. Older lesions tend to calcify and may be confused with retinoblastoma. Histologically, they are composed of elongated fibrous (pilocytic) astrocytes containing small oval nuclei. Rare giant cell (gemistocytic) astrocytomas have also been described.

Vascular Tumours

Retinal capillary haemangiomas are the hallmark of von Hippel–Lindau and are similar to cerebellar haemangioblastomas. Clinically, the tumours progress from small red to greyish appearing lesions with no abnormal vessels, to moderate-sized pink or yellowish lesions with large feeder vessels and retinal haemorrhage and exudation. Fluorescein angiography reveals leakage in the early phase of the angiogram and is diagnostic. Histologically, the tumour is composed of small vascular channels that have the appearance of capillaries. Large foamy cells which are debated to be histocytic, endothelial or astrocytic in origin are present in the stroma. Gliosis of the adjacent retina with massive exudation, haemorrhage and retinal detachment is frequently present.

Cavernous haemangioma of the retina is a rare congenital malformation that differs from capillary haemangiomas in several aspects. The affected retina shows isolated clusters of aneurysmal dilation of retinal vessels with absent feeder vessels. There is no arteriovenous shunting owing to the low flow state of these tumours and there is no disturbance of permeability on clinical examination or fluorescein angiography. Histologically, large vascular channels with normal walls thicken the retina. The inner retinal layers may be discontinuous in the area of the vascular lesions.

Lymphoid Tumours

Retinal involvement may be the initial and only manifestation of primary CNS lymphoma. The retina is usually not involved in a generalized systemic lymphoma and conversely the reticuloendothelial system is not affected by a primary CNS lymphoma. The vitreous is often involved by the presence of lymphoma cells and can masquerade as uveitis. In fact, the the presence of glaucoma, uveitis and neurological symptoms in an elderly individual should raise the suspicion of primary CNS lymphoma and appropriate imaging studies of the brain should be performed. The diagnosis may be confirmed microscopically by cytological examination of cells from the cerebral spinal fluid, vitrectomy, vitreous aspiration, or from biopsy of brain or retina.

The gross findings include placoid or hemispherical mound-like elevations of the retinal pigment epithelium (**Figure 12; see colour plate section**). The optic nerve may be involved with neoplastic cells as opposed to the uvea, which is rarely affected except for the presence of a reactive inflammatory infiltrate. The tumour cells possess varying amounts of cytoplasm with polyhedral or amoeboid outlines. Most of the nuclei exhibit single or multiple nucleoli. Mitotic figures are frequently present along with areas of retinal necrosis.

Neuroepithelial Tumours

Tumours of the neuroepithelium can be divided into congenital and acquired lesions. Congenital lesions usually arise from the medullary epithelium and are called medulloepitheliomas. The medullary epithelium has the capacity to differentiate into retinal pigment epithelium (RPE), nonpigmented and pigmented ciliary epithelium and neurons. They have also been termed diktyomas, since they are composed of a network of neuroepithelial bands. When they contain heterologous tissue elements such as cartilage, skeletal muscle and brain, they are referred to as teratoid.

Clinically, these tumours usually arise from the ciliary body, but on rare occasion may arise from the retina or optic disc. Some ciliary body medulloepitheliomas are cystic and may become detached to be carried into the anterior chamber through the pupil. Histologically, the tumour is composed of cords and sheets of cells ranging from a single layer of columnar epithelium to stratified multilayered structure that resembles embryonic retina. The apical side of the cords has several fenestrations resembling the external limiting membrane of the retina. Malignant medulloepitheliomas contain totally undifferentiated cells and Flexner–Wintersteiner rosettes may occasionally be observed, making differentiation from childhood retinoblastoma difficult.

Acquired lesions arising from the retinal pigment epithelium, non-pigmented and pigmented ciliary pigment epithelium can be categorized either as adenomas or adenocarcinomas. However, a more frequent lesion arising from

the non-pigmented ciliary epithelium is a Fuchs adenoma, which is a hyperplasia rather than a true neoplasm. Grossly, Fuchs adenomas are small, white nodules that usually measure less than 1 mm in diameter and are located on the pars plicata. Histologically, the lesion consists of a nodule of hyperplastic, nonpigmented ciliary epithelium arranged in sheets and tubules, embedded in a matrix of PAS-positive basement membrane-like material. Congenital lesions of the RPE include combined hamartoma of the retina and RPE and congenital hypertrophy of the retinal pigment epithelium (CHRPE). Bilateral and multiple CHRPE lesions have been associated with Gardner's syndrome, an autosomal dominant condition characterized by adenomatous polyps of the colon, which may progress to adenocarcinoma and musculoskeletal anomalies.

Tumours of the Uveal-Tract

Melanocytic naevi and malignant melanoma are by far the most frequent and clinically important tumours of the iris, ciliary body and the choroid. In autopsy studies, metastasis is the most frequent tumour observed in the choroid; however, choroidal melanoma is the most frequent primary intraocular malignancy of adults.

Iris

Tumours of the iris can arise from the iris pigment epithelium, smooth muscle or stromal melanocytes. These include cysts and adenomas of the iris pigment epithelium, leiomyomas, naevi and malignant melanomas.

An iris naevus is a benign tumour that arises from the stromal melanocytes. They appear flat to slightly elevated under slit lamp examination and usually remain stationary in size, although they have the potential for malignant transformation (**Figure 13; see colour plate section**). Histologically, they are composed of low-grade spindle-type cells. Occasionally, when the naevus consists of heavily pigmented, plump polyhedral cells, it is referred to as a melanocytoma (**Figure 14**). In neurofibromatosis type I, the stromal melanocytes undergo a hamartomatous proliferation into small, tan-coloured nodules on the surface of the iris known as Lisch nodules. Juvenile xanthogranuloma, a benign dermatological disorder of children and young adults, can lead to thickening of the iris with neovascularization and haemorrhage. Histologically, it is composed of mildly atypical histiocytes.

An iris melanoma is a malignant neoplasm of the stromal melanocytes. They are nodular, circumscribed and variably pigmented. An iris melanoma can be darkly pigmented or amelanotic or a mixture of the two. An important variant of the nodular type of iris melanoma is diffuse iris melanoma. This type of melanoma presents with diffuse thickening of the iris, acquired hyperchromic heterochromia and glaucoma (**Figure 15**). The histology of

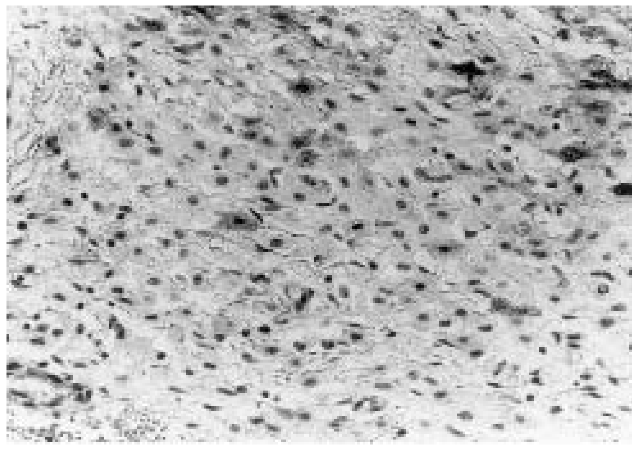


Figure 14 Melanocytoma. Bleached preparation showing plump polyhedral melanocytes with small ovoid nuclei (courtesy AFIP).

Figure 15 Diffuse iris melanoma. The angle of the eye is well shown in the low-power photomicrograph. The higher power photograph shows discohesive epithelioid melanoma cells studded on the anterior surface of iris and the angle structures.

iris naevus and melanoma will be discussed with their choroidal counterparts.

Ciliary Body

Ciliary body tumours include neoplasms of the pigmented and non-pigmented ciliary epithelium as discussed in the section on retinal tumours. Ciliary body melanomas are less frequent than choroidal melanomas but more frequent than iris malignant melanomas. The growth pattern of ciliary body melanomas differs from that of the choroidal variety. The absence of Bruch's membrane in the ciliary body leads to early endophytic growth of the tumour in the posterior chamber, with occasional subluxation or indentation of the lens. Ciliary body melanoma can grow on to the iris, in which case a small hyperpigmented lesion on the iris might appear on the periphery of the iris, representing the tip of the iceberg. Occasionally, they can erode through the sclera and appear on the surface of the globe near the limbus, simulating a conjunctival melanoma. Diffuse ciliary body melanomas have a tendency to grow in a ring-like fashion.

Choroid

Choroidal naevi are benign melanocytic tumours that can occasionally degenerate into malignant melanomas. Ophthalmoscopically, a choroidal naevus is variably pigmented (it may rarely be amelanotic) and is usually less than 2 mm in height. The surface of the lesion often has small yellowish–white deposits known as drusen. Subretinal fluid is not a hallmark of choroidal naevi as is the case with melanomas. Histologically, choroidal naevi are composed of low-grade spindle cells. The proliferation of benign melanocytes throughout the uvea is associated with certain systemic malignancies such as ovarian and uterine carcinoma (Gass *et al*., 1990). This is referred to as bilateral diffuse uveal melanocytic proliferation (BDUMP) and is a paraneoplastic syndrome.

Choroidal melanomas are the most frequent of all the uveal (iris, ciliary body and choroid) melanomas. The use of the indirect ophthalmoscope has dramatically improved the detection and diagnosis of small choroidal tumours. The gross pathological features of choroidal melanoma are characteristic. The tough fibrous sclera prevents the outward growth of the melanoma into the orbit. However, the weaker Bruch's membrane on the inner aspect of the tumour can become stretched and eventually rupture from the expanding sub-retinal pigment epithelial mass. The tumour herniates through the rupture and grows into the subretinal space, giving a 'collar button' or mushroom-shaped configuration (**Figure 16; see colour plate section**). The retina overlying the uveal melanoma undergoes atrophy or cystoid degeneration while the retina surrounding the tumour is detached by the accumulation of serous exudate between the retina and the RPE. Some choroidal melanomas can invade the orbit by growing along the course of perforating nerves and vessels.

In 1913, Calender classified uveal malignant melanoma based on cytological and histological features. Since then, this scheme has been widely accepted owing to its prognostic significance, but has also undergone some modification from the original description of cell types. The cells of the uveal malignant melanoma are now divided into spindle and epithelioid cell types. Spindle-type cells are fusiform shaped and usually arranged in a fascicular pattern (**Figure 17**). They have cigar-shaped nuclei, with prominent nucleoli and frequent mitotic figures. The epithelioid-type cells are larger and more pleomorphic than the spindle type cells. They have abundant cytoplasm with distinct cell boundaries (**Figure 18**). Their characteristic loose intercellular cohesion leads to extracellular space between adjacent cells. The nucleus is larger and rounder with irregular indentations when compared with the spindle cell counterpart. Mitotic activity is also greater than in spindle cell tumours. Based on the presence or absence of each cell type, uveal melanomas can be classified as

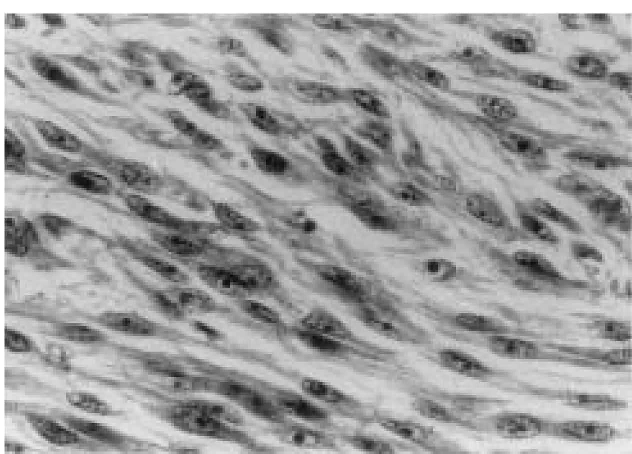


Figure 17 Malignant melanoma. Spindle-type cells (courtesy AFIP).

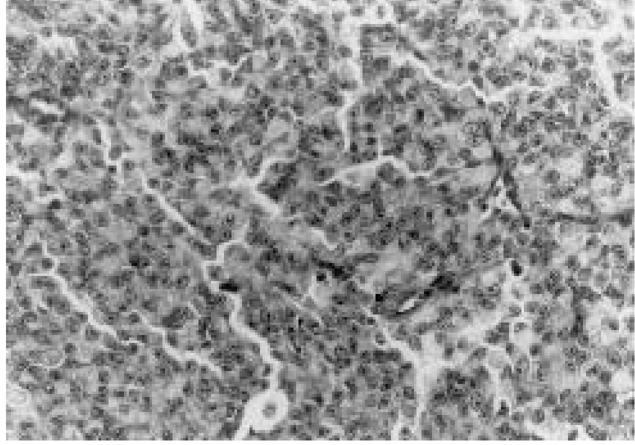


Figure 18 Malignant melanoma. Epithelioid-type cells (courtesy AFIP).

spindle, epithelioid or mixed cell-type tumours, each with different prognostic significance (discussed later).

Since uveal melanocytes are derived from neural crest, they stain positive for HMB-45, S-100 and neuron-specific enolase. Uveal naevi stain with less intensity for all three immunohistochemical markers.

Other tumours of the uvea include leukaemic infiltration, uveal lymphoma with extraocular extension, leiomyosarcoma and neuroepithelial tumours of the ciliary body (discussed under Tumours of the Retina). Choroidal osteoma and haemangioma are relatively rare tumours. The latter when seen in association with the Sturge–Weber syndrome often gives the fundus a 'tomato ketchup' appearance.

Tumours of the Orbit

A variety of benign and malignant tumours of the orbit have been described in adults and children. Capillary haemangioma, lymphangioma, dermoid cysts and rhabdomyosarcoma are most common in children whereas fibrous histiocytoma, cavernous haemangioma, schwannoma and lymphoid neoplasms occur more commonly in adults.

Capillary Haemangioma

In most cases, capillary haemangioma is located in the anterior aspect of the orbit and frequently involves the lid. The appearance of the lesion varies from a striking red 'strawberry naevus' located in the dermis to a more diffuse and light blue discoloration of the skin present in deeper orbital lesions. They have an infiltrative growth pattern and can involve any of the orbital structures. The tumour is composed of plump endothelial cells, which proliferate in solid lobules during the rapid growth phase of the tumour. Although these are benign tumours, they may contain numerous mitotic figures. Electron microscopy has revealed pericytes surrounding the endothelial lobules, indicating the capillary origin of these tumours.

Cavernous Haemangioma

Grossly, cavernous haemangiomas have an intense violaceous hue due to the presence of partially oxygenated venous blood in the vascular spaces. The tumour has a well-formed capsule surrounding vascular channels, which measure 0.5–1 mm in diameter and are filled with blood. Electron microscopy has revealed the vascular spaces to be lined by a monolayer of endothelial cells.

Lymphangioma

Lymphangioma is considered a choristoma since there are no endothelial-lined lymphatic channels in the orbit. In contrast to other encapsulated orbital tumours, a lymphangioma does not possess a capsule and therefore has a tendency to diffusely infiltrate soft tissues of the orbit. It is composed of endothelial lined channels with the ultrastructural features of lymphatics.

Haemangiopericytoma

Haemangiopericytoma is an encapsulated tumour with a grey to red–brown colour on cut section. They are composed of small polyhedral and spindle-shaped cells tightly packed around blood vessels. These vessels range from small capillaries to large sinusoidal spaces, some of which have a typical 'staghorn'configuration. Haemangiopericytoma have been divided into benign and malignant varieties based on histological criteria: cytological atypia, mitotic activity and necrosis.

Fibrous Histiocytoma

Fibrous histiocytoma is one of the most common tumours of the orbit. Grossly, they are encapsulated, rubbery to firm and vary from greyish white to yellow–tan. Microscopically they are composed of a mixture of spindle-shaped fibroblast-like cells and more ovoid, sometimes lipidized histiocytes. Detailed descriptions about the histological variations can be found elsewhere.

Neurofibroma

Of the three types of neurofibromas found in the orbit, plexiform, diffuse and isolated, only the plexiform variant is consistently observed in patients with neurofibromatosis type I. It appears within the first decade of life and grows in an infiltrating fashion that may involve any orbital structure. The enormous overgrowth of the orbital peripheral nerves may sometimes cause excessive redundancy of the lid skin, leading to elephantiasis neuromatosa. This tumour-like lymphangioma is not encapsulated. Histologically, it is composed of numerous proliferating units of the terminal branches of peripheral nerves in a loose stroma of axons, Schwann cells and endoneural fibroblasts.

Schwannoma

Like neurofibroma, schwannoma originates from the peripheral nerves of the orbit, most often the supraorbital nerve. Histopathologically, they are encapsulated by the perineurium of the nerve of origin. The classical feature of a schwannoma is the presence of solid cellular areas, referred to as the Antoni A pattern and other more myxoid areas with stellate or ovoid-appearing cells referred to as Antoni B areas. In Antoni A areas, nuclear palisading is common, which often leads to formation of fascicles of nuclei in a highly regimented fashion known as Verocay bodies.

Rhabdomyosarcoma

Rhabdomyosarcoma is the most common primary malignant tumour of childhood. The average age at time of diagnosis is 7 years. The superonasal aspect of the orbit is the site of predilection for this tumour. When it originates in a submucosal site such as the conjunctival stroma, it is

referred to as the botryoid type. The tumour is not encapsulated and usually invades adjacent structures. Based on histological features, rhabdomyosarcomas are divided into embryonal, alveolar and pleomorphic (differentiated) types.

Dermoid Cyst

Dermoid cysts result from congenital arrest of ectoderm within sutures of the orbital bones. The majority of dermoid cysts become symptomatic within the first decade of life. Nasally located dermoids are often lined by conjunctival epithelium, whereas temporal dermoids are lined by skin and have dermal appendages such as hair and sebaceous glands present in the wall of the cyst.

Lymphoid Neoplasms

Lymphoid neoplasms of the orbit are among the most prevalent of orbital tumours. In patients above age 60 years, they represent >50% of all orbital tumours. Clinically and grossly, lymphoid tumours have a fish-flesh ('salmon patch') or creamy to yellow appearance. They are very friable owing to the absence of fibrous stroma. Histologically, lymphoid tumours should be divided into lymphomas and benign lymphoid hyperplasia. However, this distinction is not always possible and tumours exhibiting intermediate characteristics are referred to as atypical lymphoid hyperplasia (Knowles *et al*., 1990). Benign lymphoid hyperplasia has a follicular organisation with benign-appearing small, dark lymphocytes. However, the follicles can be of irregular size and shape and have high mitotic activity.

Orbital and adnexal lymphomas are almost always of the non-Hodgkins's type. Small cell lymphomas of B cell type originating in the conjunctival substantia propria are referred to as Maltoma (mucosal associated lymphoid tissue).

Other Tumours

Other tumours found in the orbit include teratomas, exclusively in newborns, histiocytic disorders such as unifocal and multifocal eosinophilic granuloma (Langerhans cell histiocytosis), malignant melanoma, blue naevus, fibrous dysplasia, granulocytic sarcoma, osteogenic sarcoma and metastasis. This is by no means an exhaustive list, since other rare tumours in the orbit have been described.

Tumours of the Optic Nerve

Glioma

Most gliomas of the optic nerve are benign pilocytic astrocytomas. They are usually seen in association with neurofibromatosis type I. They are very slow-growing tumours, often enlarging in an episodic fashion and then lying dormant for many years. On imaging studies such as MRI and CT, they appear as fusiform enlargements of the optic nerve. The dura is always intact surrounding the tumour and it is often difficult to determine the margin of the tumour from the normal optic nerve. Histologically, pilocytic astrocytomas of the optic nerve are composed of elongated, spindle-shaped, hair-like astrocytes. They have a benign appearance with the notable exception of mitotic figures. These tumours show positive immunohistochemical staining for glial fibrillary acidic protein (GFAP). Rosenthal fibres, which represent degenerative changes within the astrocytic cell processes, are also frequently seen in these tumours. They appear as cylindrical or spherical eosinophilic bodies. Ultrastructurally, Rosenthal fibres are composed of electron-dense granular material and glial filaments.

Meningioma

Almost all orbital meningiomas either arise from the meninges of the optic nerve or invade the orbit secondarily from an adjacent intracranial site such as the sphenoid wing. The clinical features of meningioma are much different from those of glioma. Meningiomas are associated with type II and not type I neurofibromatosis. The age at diagnosis is considerably greater in patients with meningioma. However, it should be considered in the differential diagnosis of any optic nerve lesion causing exophthalmos and visual loss even in the first few years of life.

Unlike gliomas, meningiomas usually extend through the dura. The tumour does not invade the optic nerve and grows in a sleeve-like manner within the meninges. Meningiomas may infiltrate all of the orbital structures but intraocular invasion through the sclera is rare. Histologically, they are similar to intracranial meningiomas.

Tumours of the Lachrymal Gland and Sac

Lachrymal gland tumours bear a striking resemblance to those of the salivary gland. Approximately half of all lachrymal gland tumours are epithelial and the rest are non-epithelial. Of the epithelial tumours, half are benign and half malignant. They constitute 10% of all orbital lesions in some series.

Pleomorphic Adenoma

Pleomorphic adenomas or benign mixed tumours are the most common tumours of the lachrymal gland. They present as a firm, painless and slowly developing mass in the superolateral orbit. They are pseudoencapsulated and may contain cystic areas. Histological sections reveal an admixture of epithelial and mesenchymal elements that can take a variety of patterns.

Pleomorphic Carcinoma

Also known as malignant mixed tumours, pleomorphic carcinomas arise from malignant transformation of pleomorphic adenomas. Histologically, they can take the appearance of adenocarcinoma, adenoid cystic carcinoma,

squamous cell carcinoma, undifferentiated carcinoma or sebaceous carcinoma in which a pre-existing pleomorphic adenoma can be identified.

Other Tumours

Adenoid cystic carcinoma, mucoepidermoid and adenocarcinoma of the lachrymal gland can arise *de novo* in the lachrymal gland. Their histopathology is similar to that of tumours found in the salivary glands. Lymphomas and benign reactive lymphoid hyperplasia are the most common non-epithelial tumours of the lachrymal gland. **Table 3** lists the different types of tumours that can arise in the lachrymal sac.

Molecular Genetic Findings

Retinoblastoma represents a paradigm in understanding the molecular basis of inherited predisposition to developing malignancy. The retinoblastoma gene, a tumour suppressor gene, is located on the long arm (q14 locus) of chromosome 13 (Dryja *et al*., 1984). Hence each cell carries two copies of this gene and in normal individuals their gene product serves to inhibit the development of retinal tumours. The retinoblastoma gene product (pRB) normally inactivates E2F (a nuclear transcription factor) and halts the normal cell cycle between G1 and S phases. However, in each case of retinoblastoma, an initial mutation inactivates one copy of the gene. The mutation can be caused by a single-base change of DNA, a point mutation, a chromosomal translocation, a small deletion or a large deletion that is observable by karyotyping. The mutation can be germ-line or somatic leading to the hereditary (multiplex) and non-hereditary (simplex) forms of retinoblastoma. The homozygous form of the mutated retinoblastoma gene is probably fatal *in utero*. In the heterozygous form (hereditary form), the individual inherits a mutated gene from one parent and acquires the second mutation sometime after birth. This was Knudson's initial 'two hit hypothesis' and postulates the necessity for both genes to be inactivated for development of retinoblastoma. Loss of all pRB due to homologous deficiency at 13q14 eliminates the 'brakes' and uncontrolled proliferation of retinoblastoma ensues. In the non-hereditary form, the individual acquires both mutations after birth.

The majority of retinoblastomas are of the inherited variety. The defective gene in this instance can be acquired from a parent who is a known carrier or more often they may represent a new germline mutation. Multiple and especially bilateral retinoblastomas almost always signify a germinal mutation. Median age at diagnosis is significantly lower for bilateral retinoblastoma than for unilateral retinoblastoma. An important facet of providing care for patients with retinoblastoma is the assessment of risk of developing the same malignancy in the siblings or progeny of the affected individual. This can be done accurately using modern molecular biology techniques.

Phakomatoses ('birthmarks') are comprised of a group of disparate clinical syndromes, which share certain features in common. These include the presence of multisystem hamartomas, predominantly in ocular, cutaneous and intracranial locations. They may arise at birth or later in life, and can undergo malignant transformation. The term is usually applied to include von Hippel–Lindau (angiomatosis retinae, capillary haemangioma), turberous sclerosis (Bournveille disease), neurofibromatosis (type I and II) and Sturge–Weber sydnromes (encephalo-trigeminal angiomatosis), although some authors also include Wyburn–Mason and ataxia-telangectasia (Louis–Bar syndrome) in this group of diseases. Of the first four conditions, all except Sturge–Weber are inherited in an autosomal dominant fashion. The target locus for von Hippel–Lindau was recently mapped to 3p25–26 and the gene product (pVHL) balances mRNA transcription. The same locus is home to the gene which codes for vascular endothelial growth factor (VEGF). Mutations in 3p25–26 lead to unchecked transcription and upregulation of VEGF. Tuberous sclerosis complex (TSC) has two known disease

Table 3 Lachrymal sac tumours

Lachrymal sac tumours
Epithelial tumours
Benign papillomas—exophytic (grow into lumen of sac) and endophytic (inverted).
Squamous
Transitional
Mixed
Carcinomas—*de novo* or within papillomas
Squamous
Transitional
Mixed
Mucoepidermoid
Glandular tumours
Benign
Pleomorphic adenoma
Oncocytoma
Malignant
Oncocytic adenocarcinoma
Adenocarcinoma
Adenoid cystic
Mesenchymal tumours
Fibrous histiocytoma
Haemangiopericytoma
Fibroma/fibromyxoma
Haemangioma
Malignant melanoma
Lymphoid infiltrates (benign, atypical, malignant)
Inflammatory pseudotumour
Secondary tumours (extension from nose, paranasal sinuses, skin and conjunctiva)
Metastasis (rare)

loci, on chromosomes 9q34 (*TSC1*) and 16p13.3 (*TSC2*). The TSC1 and TSC2 gene products are known as hamartin and tuberin, respectively, and are expressed in neurons and astrocytes where they physically interact. Neurofibromatosis (NFT), the most common of the phakomatoses, has two forms. The gene responsible for NFT type I (peripheral form) is located on chromosome 17 and NFT type II (central form) on chromosome 22.

The majority (85%) of follicular B cell lymphomas, such as primary ocular–CNS lymphoma, harbour a characteristic translocation (t(14;18)) which is felt to represent an error occurring during physiological gene rearrangement. Genes located at the translocation breakpoint (*bcl-2/ IgH*) are considered putative oncogenes that are deregulated via juxtaposition with an antigen receptor gene. Southern blot analysis is helpful in confirming both the translocation as well as monoclonal IgH expression.

Prognostic Factors

Mortality from choroidal melanoma depends on a variety of factors but has been reported in larger series to be 35% at 5 years and 50% at 10 years. There have been multiple studies to correlate gross and cytopathological features of uveal melanoma with prognosis. Location of the tumour, largest tumour diameter (LTD), height of tumour have all been found to correlate with future mortality. Patient survival is better for LTD less than 10 mm. On the histological side, optic nerve or scleral invasion, cell type (spindle, mixed or epithelioid), standard deviation of the nucleolar area and mean of the largest nucleoli are positively correlated with death from metastatic melanoma. Spindle cell tumours carry the most favourable prognosis, followed by mixed cell lesions. Pure epithelioid cell malignant melanomas have the poorest prognosis, but fortunately are also the most rare.

Retinoblastoma has a much better prognosis than choroidal melanoma. There has been a remarkable shift in the long-term survival of these patients since the early part of the century, when it was uniformly a fatal disease. The overall long-term survival in retinoblastoma in the USA and UK is now 85%. Even though patients with retinoblastoma survive the initial disease, those that have the genetic abnormality are subject to other tumours, most notably osteosarcoma. About two-thirds of the second tumours are in the radiation field used to treat retinoblastoma and the other third are in areas remote from the radiation field.

Retinoblastomas metastasize in four ways: direct infiltration, dispersion, haematogenous and lymphatic dissemination. Direct infiltrative spread occurs along the optic nerve from the eye to the brain. Once the orbital soft tissues are invaded, the tumour spreads directly into the orbital bones, through the sinuses into the nasopharynx, or via the various foramina into the cranium. Dispersion of tumour cells occurs after cells in the optic nerve have invaded the leptomeninges and gained access to the subarachnoid fluid. Flow of cerebrospinal fluid spreads the tumour cells to the brain and spinal cord. The most common sites for haematogenous spread are lungs, bones, brain and other viscera. Extraocular invasion and, to a lesser degree, choroidal invasion increase the risk of haematogenous spread. Lymphatic spread occurs in those tumours that invade the anterior orbital tissues including the conjunctiva, since the orbit is devoid of lymphatics. The spread of retinoblastoma into the optic nerve and into the orbit through the sclera is associated with less favourable prognosis.

When metastasis occurs, it is generally within the first year or two following treatment. The median time to death in patients with metastatic retinoblastoma is 6.4 months in unilateral cases and 14.2 months in bilateral cases. In contrast, the estimated median time to death in uveal melanoma is 7.2 years. Late death from metastasis, which occurs frequently following enucleation for melanoma, is so rare after treatment for retinoblastoma that when metastasis is suspected, the question of an independent new primary tumour must be considered.

All malignant neoplasms of the lachrymal gland with the possible exception of mucoepidermoid carcinoma carry a poor prognosis. Death usually results within 3–5 years from tumour extension along nerves through the superior orbital fissure and invasion through the orbital bones into the middle cranial fossa. Haematogenous metastasis, particularly to the lungs, is the second most common cause of death.

Overview of Present Clinical Management

Tumours of the Eyelid and Conjunctiva

Local excision is usually curative for the majority of skin and conjunctival tumours. Notable exceptions to this are sebaceous cell carcinoma and primary acquired melanosis. Besides growing locally in a nodular fashion, both of these lesions can spread to distant sites of the conjunctiva and epidermis of the skin intraepithelially termed pagetoid extension, rendering local excision unfeasible. Map biopsies are therefore required to assess the extent of neoplastic involvement of the ocular adnexa. In such cases cryotherapy can be used to treat wider areas, but excessive application can lead to dry eye and symblepharon formation.

Squamous cell carcinoma and to a lesser extent basal cell carcinoma can invade the deeper tissues in the orbit, requiring orbital exenteration. Extensive pagetoid spread of sebaceous cell carcinoma to the cornea, all four quadrants of the conjunctiva and the skin of the eyelid may also necessitate orbital exenteration. A variety of composite and myocutaneous flaps and skin grafting techniques from adjacent areas have been described for repair of large defects of the lids in cases where wide local excision is performed.

Primary lymphoid neoplasms of the ocular adnexa may be observed if they are asymptomatic and without evidence of systemic dissemination. Radiation or cryotherapy can be used if they involve extensive areas of the conjunctival fornices or the patient is symptomatic. A thorough evaluation to rule out a systemic lymphoma is mandatory in every patient in whom a lymphoma is suspected.

Tumours of the Retina

It is imperative to make as accurate a diagnosis as possible when a patient in whom retinoblastoma is suspected is encountered. The long list of benign conditions causing leukocoria, one of the cardinal signs of retinoblastoma (see **Table 2**), can cause confusion and lead to misdiagnosis with disastrous consequences. A variety of modalities are available for the treatment of retinoblastomas; however, the type of treatment depends on many factors, such as age of the patient, unilateral or bilateral disease, size and intraocular location of the tumour, and most importantly local or distant spread of the tumour. Enucleation is probably the most widely employed treatment. Other forms of treatment include external beam radiation, cryotherapy and photocoagulation therapy. Invasion of the optic nerve and extraocular extension significantly alter prognosis and mandate the use of chemotherapy in addition to local treatment (Gallie *et al*., 1996; Murphree *et al*., 1996; Shields *et al*., 1996).

Capillary haemangiomas of the retina are most often treated with cryotherapy or photocoagulation, depending upon the location. Treatment of haemangiomas near the optic disc or macula can be challenging. Intraocular lymphoma is treated with radiation and chemotherapy, but it carries a poor prognosis.

Tumours of the Uveal Tract

Iris melanomas can be locally excised by performing a sectoral iridectomy or iridocyclectomy with good results. Juvenile xanthogranuloma responds dramatically to topical steroids and is a self-limited disorder.

The management of choroidal melanoma is controversial (Zimmerman and McLean, 1979). The collaborative ocular melanoma study (COMS) now in progress is comparing enucleation versus plaque radiotherapy for medium-sized tumours (6–10 mm in height) and enucleation versus pre-enucleation radiotherapy for large tumours (>10 mm in height) with 10-year survival as outcome measure. Initial data suggest no statistical difference between enucleation and plaque radiotherapy for medium-sized choroidal melanomas (Collaborative Ocular Melanoma Study Group 1998). Enucleation is generally recommended for large tumours with little hope of salvageable vision (Gragoudas *et al*., 1980). Radiotherapy, with either radioactive plaque or charged particles (proton beam), is recommended for medium-sized and large melanomas when there is a possibility of salvaging some vision. More recently, the use of transpupillary thermotherapy either alone or combined with plaque radiotherapy has been advocated.

Tumours of the Orbit

Capillary haemangiomas grow rapidly after birth and usually appear during the first 6 months of life. They enter a stationary stage during the next 1–2 years followed by spontaneous involution during the next 5–6 years. If the tumour is not judged to be a threat to vision and subsequent amblyopia, it can be observed without treatment. In case of significant functional or cosmetic impairment, intralesional corticosteroid injection is the most common form of treatment.

The surgical management of encapsulated and nonencapsulated tumours of the orbit is different. Encapsulated tumours more easily lend themselves to removal without too much dissection. Even then surgery is usually reserved for progressively enlarging tumours that threaten vision due to compromise of the optic nerve or the ocular surface from progressive proptosis. Lymphangioma and plexiform neurofibroma, on the other hand, are rarely approached surgically with the intent of completely excising the tumour, owing to their infiltrative growth pattern. Rhabdomyosarcoma is essentially treated with systemic chemotherapy.

The mainstay of treatment for ocular and periocular lymphoid tumors is radiotherapy (Char *et al*., 1998). Cryotherapy is often very effective for localized and low-grade conjunctival lymphomas and has much less morbidity when compared with radiotherapy. Local radiotherapy and systemic chemotherapy are recommended for high clinical grade lymphomas, multifocal processes such as angiotropic large cell lymphoma and systemic leukaemia.

Tumours of the Optic Nerve

Both optic nerve glioma and meningioma are slow-growing tumours. Management of these tumours is again controversial. These can often be observed for many years until such time when they extend intracranially to involve the optic chiasm or the contralateral optic nerve. In these instances the tumour can be surgically resected or radiation can be used to retard its growth.

Tumours of the Lachrymal Gland and Sac

Complete excision of pleomorphic adenomas of the lachrymal gland is absolutely essential and usually curative. However, incomplete resection can lead to local recurrence or even malignant transformation in the future. All malignant neoplasms of the lachrymal gland carry a poor prognosis despite aggressive surgical attempts at complete resection. Adenoid cystic and adenocarcinoma of the lachrymal gland are almost uniformly fatal at 5 years.

REFERENCES

Albert, D. M. and Jakobiec, F. A. (2000a). Ocular oncology. In: *Principles and Practice of Ophthalmology*, Sec. XVIII, Vol. 6. 5003–5178 (W. B. Saunders, Philadelphia).

Albert, D. M. and Jakobiec, F. A. (2000b). Lids and orbit. In: *Principles and Practice of Ophthalmology*, Sec. XIII, Vol. 4. 3045–3571 (W. B. Saunders, Philadelphia).

Char, D. H., *et al*. (1988). Primary intraocular lymphoma (ocular reticulum cell sarcoma) diagnosis and management. *Ophthalmology* **95**, 625.

Collaborative Ocular Melanoma Study Group. (1998). The Collaborative Ocular Melanoma Study (COMS) randomized trial of pre-enucleation radiation of large choroidal melanoma II: Initial mortality findings. COMS Report No. 10. *American Journal of Ophthalmology* **125**, 779–796.

Dryja, T. P., *et al*. (1984). Homozygosity of chromosome 13 in retinoblastoma. *New England Journal of Medicine* **310**, 550–553.

Gallie, B. L., *et al*. (1996). Chemotherapy with focal therapy can cure intraocular retinoblastoma without radiotherapy. *Archives of Ophthalmology* **114**, 1321–1328.

Gass, J., *et al.* (1990). Bilateral diffuse uveal melanocytic proliferation in patients with occult carcinoma. *Archives of Ophthalmology* **108**, 527–533.

Gragoudas, E. S., *et al*. (1980). Proton beam irridiation. An alternative to enucleation for intraocular melanomas. *Ophthalmology*, **87**, 571–581.

Knowles, D. M., *et al*. (1990). Lymphoid hyperplasia and malignant lymphoma occurring in the ocular adnexa (orbit, conjunctiva, and eyelids): a prospective multiparametric analysis of 108 cases during 1977 to 1987. *Human Pathology* **21**, 959.

Murphree, A. L., *et al*. (1996). Chemotherapy plus local treatment in the management of intraocular retinoblastoma. *Archives of Ophthalmology* **114**, 1348–1356.

Shields, C. L., *et al*. (1996). Chemoreduction in the initial management of intraocular retinoblastoma. *Archives of Ophthalmology* **114**, 1330–1338.

Zimmerman, L. E. and McLean I. W. (1979). An evaluation of enucleation in the management of uveal melanomas. *American Journal of Ophthalmology* **87**, 741–760.

FURTHER READING

McLean (1994). *Tumors of the Eye and Ocular Adnexa*, Third Series, Fascicle 12. (Armed Forces Institute of Pathology, Washington, DC).

Shields, J. A. and Shields, C. L. (1999). *Atlas of Eyelid and Conjunctival Tumors*. (Lippincott Williams & Wilkins, Baltimore).

Shields, J. A. and Shields, C. L. (1999). *Atlas of Intraocular Tumors*. (Lippincott Williams & Wilkins, Baltimore).

Shields, J. A. and Shields, C. L. (1999). *Atlas of Orbital Tumors*. (Lippincott Williams & Wilkins, Baltimore).

Spencer (1996). *Ophthalmic Pathology–An Atlas and Textbook*. (W. B. Saunders, Philadelphia).

Chapter 53
Ear

Bruce M. Wenig
The Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, NY, USA

CONTENTS

- Normal Development and Structure
- Neoplasms of the Ear and Temporal Bone
- Neoplasms of the Middle Ear and Temporal Bone

NORMAL DEVELOPMENT AND STRUCTURE

The ear is represented by three distinct compartments, the external ear, the middle ear and temporal bone, and the inner ear. The external ear develops from the first branchial groove. The external auricle (pinna) forms from the fusion of the auricular hillocks or tubercles, a group of mesenchymal tissue swellings from the first and second branchial arches, that lie around the external portion of the first branchial groove (Moore, 1988). The external auditory canal is considered a normal remnant of the first branchial groove. The tympanic membrane forms from the first and second branchial pouches and the first branchial groove (Moore, 1988). The ectoderm of the first branchial groove gives rise to the epithelium on the external side, the endoderm from the first branchial pouch gives rise to the epithelium on the internal side and the mesoderm of the first and second branchial pouches gives rise to the connective tissue lying between the external and internal epithelia (Moore, 1988). The middle ear space develops from invagination of the first branchial pouch (pharyngotympanic tube) from the primitive pharynx. The eustachian tube and tympanic cavity develop from the endoderm of the first branchial pouch; the malleus and incus develop from the mesoderm of the first branchial arch (Meckel's cartilage) while the incus develops from the mesoderm of the second branchial arch (Reichert's cartilage) (Moore, 1988). The first division of the ear to develop is the inner ear that appears toward the end of the first month of gestation (Dayal *et al*., 1973; Moore, 1988). The membranous labyrinthe, including the utricle, saccule, semicircular ducts and cochlear duct, arises from the otic vesicle (otocyst). The otic vesicle forms from the invagination of the surface ectoderm, located on either side of the neural plate, into the mesenchyme. This invagination eventually loses its connection with the surface ectoderm. The bony labyrinthe, including the vestibule, semicircular canals and cochlea arises from the mesenchyme around the otic vesicle (Dayal *et al*., 1973; Moore, 1988).

The outer portion of the external ear includes the auricle or pinna leading into the external auditory canal with its medial limit being the external aspect of the tympanic membrane. Histologically, the auricle is essentially a cutaneous structure composed of keratinizing, stratified squamous epithelium with associated dermal adnexal structures that include hair follicles, sebaceous glands and eccrine sweat glands. The subcutaneous tissue is composed of fibroconnective tissue, fat and elastic-type fibrocartilage which gives the auricle its structural support. In addition to the dermal adnexal structures, the outer third of the external canal is noteworthy for the presence of modified apocrine glands called ceruminal glands that replace the eccrine glands seen in the auricular dermis. Ceruminal glands produce cerumen and are arranged in clusters composed of cuboidal cells with eosinophilic cytoplasm often containing a granular, golden yellow pigment. These cells have secretory droplets along their luminal border. In the inner portion of the external auditory canal, ceruminal glands, and also the other adnexal structures, are absent. Similarly to the auricle, the external auditory canal is lined by keratinizing squamous epithelium that extends to include entire canal and covers the external aspect of the tympanic membrane. The inner two-thirds of the external auditory canal contains bone rather than cartilage.

The middle ear or tympanic cavity contents include the ossicles (malleus, incus and stapes), eustacian tube, tympanic cavity proper, epitympanic recess, mastoid cavity and the chorda tympani of the facial (VII) nerve. The middle ear, and also the external ear, function as conduits for sound conduction for the auditory part of the internal ear. The anatomic limits of the middle ear include (1) lateral or internal aspect made up by the tympanic membrane and squamous portion of the temporal bone, (2) medial aspect bordered by the petrous portion of the temporal bone, (3) superior (roof) delimited by the tegmen tympani, a thin plate of bone which separates the middle ear space from the cranial cavity, (4) inferior (floor) aspect bordered by a thin plate of bone separating the tympanic cavity from the superior bulb of the internal jugular vein,

(5) anterior aspect delimited by a thin plate of bone separating the tympanic cavity from the carotid canal housing the internal carotid artery and (6) posterior aspect delimited by the petrous portion of the temporal bone containing the mastoid antrum and mastoid air cells. Histologically, the lining of the middle ear is a respiratory epithelium varying from ciliated epithelium in the eustacian tube to a flat, single, cuboidal epithelium in the tympanic cavity and mastoid. The epithelial lining the eustacian tube becomes pseudostratified as it approaches the pharyngeal end. Under normal conditions, there are no glandular elements within the middle ear. The eustachian tubes contain a lymphoid component, particularly in children, that is referred to as Gerlach's tubal tonsil. The ossicular articulations are typical synovial joints.

The internal ear is embedded within the petrous portion of the temporal bone and consists of the structures of the membranous and osseous labyrinthe, and the internal auditory canal in which the vestibulocochlear (VIII) nerve runs. The internal ear is the sense organ for hearing and balance. The anatomy and histology of this region is complex and beyond the scope of this chapter and the reader is referred to specific texts detailing the inner ear anatomy and histology.

NEOPLASMS OF THE EAR AND TEMPORAL BONE

The classification of neoplastic lesions of the ear and temporal bone are listed in **Table 1**. Owing to limitations of space, this chapter will detail those lesions unique to the ear. The most common lesions of the external ear are of cutaneous origin. The reader is referred to the chapter on dermatopathology for a more complete discussion of cutaneous pathology. Numerous non-neoplastic lesions occur in the ear and temporal bone. Although this chapter is limited to neoplasms, for completion the classification of non-neoplastic lesions of the ear and temporal bone are listed in **Table 2**. Along this line, no chapter dealing with lesions of the ear would be complete without a discussion on cholesteatomas. The section on cholesteatoma precedes the discussion on the neoplasms of this region.

Table 1 Classification of neoplasms of the ear

External ear
Benign
Keratoacanthoma
Squamous papilloma
Seborrhaeic keratosis
Ceruminal gland neoplasms
Melanocytic nevi
Dermal adnexal neoplasms
Neurilemmoma/neurofibroma
Osteoma
Chondroma
Others
Malignant
Basal cell carcinoma
Squamous cell carcinoma
Verrucous carcinoma
Ceruminal gland adenocarcinomas
Malignant melanoma
Merkel cell carcinoma
Atypical fibroxanthoma
Others
Middle and inner ear
Benign
Middle ear adenoma
Epithelial papilloma
Jugulotympanic paraganglioma
Meningioma
Acoustic neuroma
Indeterminant biological behaviour
Endolymphatic sac papillary tumour
Malignant
Middle ear adenocarcinoma
Primary squamous cell carcinoma
Rhabdomyosarcoma
Osteosarcoma, chondrosarcoma
Haematolymphoid (e.g. malignant lymphoma)
Others
Secondary tumours

Cholesteatoma (Keratoma)

Cholesteatoma is a pseudoneoplastic lesion of the middle ear characterized by invasive growth and the presence of stratified squamous epithelium that forms a sac-like accumulation of keratin within the middle ear space. Despite their invasive growth, cholesteatomas are not considered to be true neoplasms. The term cholesteatoma is a misnomer in that it is not a neoplasm and it does not contain cholesterol (Ferlito *et al*., 1997). Perhaps the designation of keratoma would be more accurate, but the term cholesteatoma is entrenched in the literature.

Cholesteatomas tend to be more common in men than in women and are most common in the third and fourth decades of life. The middle ear space is the most common site of occurrence. Initially, cholesteatomas may remain clinically silent until extensive invasion of the middle ear space and mastoid occurs. Symptoms include hearing loss, malodorous discharge and pain and may be associated with a polyp arising in the attic of the middle ear or perforation of the tympanic membrane. Otoscopic examination may reveal the presence of white debris within the middle ear which is considered diagnostic.

The majority of cholesteatomas are acquired and either arise *de novo* without a prior history of middle ear disease or arise following a middle ear infection; a small percentage of cases are congenital. The latter have also been referred to as epidermoid cysts (Schuknecht, 1993a). The pathogenesis is thought to occur via migration of

Table 2 Classification of non-neoplastic lesions of the ear

External ear
- Developmental (accessory tragi; first branchial cleft anomalies, others)
- Infectious diseases
- Keloid
- Epidermal and sebaceous cysts
- Idiopathic cystic chondromalacia
- Chondrodermatitis nodularis helicis chronicus
- Angiolymphoid hyperplasia with eosinophilia/Kimura disease
- Autoimmune/systemic diseases (relapsing polychondritis; gout)
- Exostosis
- Others

Middle and inner ear, including temporal bone
- Developmental and congenital anomalies
- Infectious (otitis media)
- Otic or aural polyp
- Cholesteatoma
- Otosclerosis
- Langerhans cell histiocytosis (eosinophilic granuloma)
- Heterotopias (central nervous system tissue; salivary gland)
- Teratoma
- Others

squamous epithelium from the external auditory canal or from the external surface of the tympanic membrane into the middle ear. The mechanism by which the epithelium may enter the middle ear probably is by a combination of events, including perforation of the tympanic membrane (particularly in its superior aspect referred to as the pars flaccida or Shrapnell's membrane following an infection) coupled with invagination or retraction of the tympanic membrane into the middle ear as a result of long-standing negative pressure on the membrane secondary to blockage or obstruction of the eustachian tube. Other theories by which cholesteatomas are thought to occur include traumatic implantation, squamous metaplasia of the middle ear epithelium or congenital.

Cholesteatomas appear as a cystic, white to pearly appearing mass of varying size containing creamy or waxy granular material. The histological diagnosis of cholesteatoma is made in the presence of a stratified keratinizing squamous epithelium, subepithelial fibroconnective or granulation tissue and keratin debris (**Figure 1; see colour plate section**). The essential diagnostic feature is the keratinizing squamous epithelium and the presence of keratin debris alone is not diagnostic of a cholesteatoma. The keratinizing squamous epithelium is cytologically bland and shows cellular maturation without evidence of dysplasia. In spite of its benign histology, cholesteatomas are 'invasive' and have widespread destructive capabilities. The destructive properties of cholesteatomas result from a combination of interrelated reasons, including mass effect with pressure erosion of surrounding structures from the cholesteatoma, the production of collagenase which has osteodestructive capabilities by its resorption of bony structures and bone resorption (Abramson *et al*., 1984). Collagenase is produced by both the squamous epithelial and the fibrous tissue components of the cholesteatoma.

The histological diagnosis of cholesteatomas is relatively straightforward in the presence of keratinizing squamous epithelium. In contrast to cholesteatomas, squamous cell carcinoma shows dysplastic or overtly malignant cytological features with a prominent desmoplastic stromal response to its infiltrative growth. Cholesteatomas do not transform into squamous cell carcinomas. DNA analysis on human cholesteatomas showed the majority of cases analysed to be euploid (Desloge *et al*., 1997). As such, owing to a lack of overt genetic instability, cholesteatomas could not be considered to be malignant neoplasms. Cholesterol granuloma is not synonymous with cholesteatoma. These entities are distinctly different pathological entities and should not be confused with one another.

Complete surgical excision of all histological components of the cholesteatoma is the treatment of choice. If not completely excised, cholesteatomas can have progressive and destructive growth, including widespread bone destruction which may lead to hearing loss, facial nerve paralysis, labrynthitis, meningitis, epidural and/or brain abscess.

Ceruminal Gland Neoplasms

Ceruminal gland tumours arise from the cerumen-secreting modified apocrine glands (ceruminal glands) of the external auditory canal. Ceruminal glands are located in the dermis of the cartilaginous (outer) portion of the external auditory canal. In general, ceruminal gland neoplasms are uncommon but represent one of the more common tumours of the external auditory canal. The generic designation of ceruminoma should be avoided. Ceruminal gland neoplasms should be specifically diagnosed according to tumour type. The classification of ceruminal gland neoplasms includes benign and malignant tumours. The benign ceruminal gland tumours include ceruminal gland adenoma (ceruminoma), pleomorphic adenoma and syringocystadenoma papilliferum (Hyams *et al*., 1988a). The malignant ceruminal gland tumours include ceruminal gland adenocarcinoma, adenoid cystic carcinoma and mucoepidermoid carcinoma.

Ceruminal gland neoplasms tend to affect men more than women and occur over a wide age range but are most frequently seen in the fourth to sixth decades of life. Symptoms include a slow-growing external auditory canal mass or blockage, hearing difficulty and infrequently otic discharge (Wetli *et al*., 1972; Pulec, 1977; Hyams *et al*., 1988a). The gross appearance of ceruminal gland neoplasms includes skin-covered, circumscribed, polypoid or rounded mass ranging in size from 1 to 4 cm in diameter. Ulceration is uncommon and may suggest a malignant neoplasm.

Histologically, ceruminal gland adenomas are unencapsulated but well-demarcated glandular proliferations. The glands vary in size and may have various combinations of growth patterns, including solid, cystic and papillary. A cribriform or back-to-back glandular pattern is commonly seen. The glands are composed of two cell layers, the inner or luminal epithelial cell is cuboidal or columnar appearing with an eosinophilic cytoplasm and a decapitation-type secretion (apical 'snouts') characteristic of apocrine-derived cells; the outer cellular layer is a spindle cell with a hyperchromatic nucleus and represents myoepithelial derivation (**Figure 2; see colour plate section**). A golden yellow–brown granular-appearing pigment can be seen in the inner lining cells and represents cerumen. Cellular pleomorphism and mitoses can be seen but are not prominent. Diastase-resistant, PAS-positive and/or mucicarmine-positive intracytoplasmic and/or intraluminal material may be seen.

Ceruminal gland pleomorphic adenomas are uncommon tumours. The histology is similar to that of pleomorphic adenomas of salivary gland origin. Syringocystadenoma papilliferum is a benign tumour of apocrine gland origin that usually occurs on the scalp and face area. Syringocystadenoma papilliferum may originate in the external auditory canal from ceruminal glands. The histology is similar to that of tumours of the more common cutaneous sites.

The treatment of choice for benign ceruminal gland neoplasms is complete surgical excision, which is curative. Recurrence of the tumour can occur and relates to inadequate surgical excision.

In contrast to ceruminal gland adenoma, patients with ceruminal gland adenocarcinomas more often have associated pain (Wetli *et al*., 1972; Pulec, 1977; Hyams *et al*., 1988a). Histologically, features that may assist in differentiating ceruminal gland adenocarcinomas from the adenomas include a loss of the glandular double cell layer with identification of only the inner or luminal epithelial cell, the presence of cellular pleomorphism with nuclear anaplasia, increased mitotic acitvity and invasive growth. However, well-differentiated ceruminal gland adenocarcinomas may appear similar to their benign counterparts and are differentiated only on the basis of invasive growth. At the other end of the spectrum, poorly differentiated ceruminal adenocarcinomas occur and are recognized on the basis of their localization in external auditory canal. Other types of ceruminal gland malignant tumours include adenoid cystic carcinoma and mucoepidermoid carcinoma. These tumours are morphologically similar to their salivary gland counterparts.

For ceruminal gland adenocarcinoma, *en bloc* surgical resection is the treatment of choice. Middle ear or temporal bone involvement necessitates more radical surgery (Hicks, 1983). Supplemental radiotherapy is recommended (Hicks, 1983). Metastases are rare and include regional lymph nodes and the lung (Pulec, 1977; Hicks, 1983). For ceruminal gland adenoid cystic carcinoma and mucoepidermoid carcinoma, wide surgical resection is the recommended treatment with or without supplemental radiotherapy. The prognosis for ceruminal gland adenoid cystic carcinoma generally is similar to that for their salivary gland counterparts, including relatively good short-term (i.e. 5-year) survival but poor long-term (i.e. 10–20 years) survival (Perzin *et al*., 1982).

NEOPLASMS OF THE MIDDLE EAR AND TEMPORAL BONE

Benign Tumours

Middle Ear Adenomas (MEA)

MEA are benign glandular neoplasms originating from the middle ear mucosa (Hyams *et al*., 1988b; Batsakis, 1989). MEA occur equally in both genders and occur over a wide age range but are most common in the third to fifth decades of life. MEA are found in any portion of the middle ear including the eustachian tube, mastoid air spaces, ossicles and chorda tympani nerve. The most common symptom is unilateral conductive hearing loss but fullness, tinnitus and dizziness may also occur. Pain, otic discharge and facial nerve paralysis rarely occur and, if present, may be indicative of a malignant process. Otoscopic examination in the majority of cases will identify an intact tympanic membrane with tumour confined to the middle ear space with possible extension to the mastoid. Occasionally, the adenoma will perforate through the tympanic membrane with extension into and presentation as an external auditory canal mass. There are no known aetiological factors related to the development of MEA. MEA are not associated with a prior history of chronic otitis media. Concurrent cholesteatomas may be seen with MEA but there is no known association between these two lesions.

An MEA is a grey–white to red–brown, rubbery to firm mass free of significant bleeding on manipulation. Histologically, MEA are unencapsulated lesions with glandular or tubule formation, as well as solid, sheet-like, trabecular, cystic and cribriform growth patterns (**Figure 3; see colour plate section**). Rarely, MEA may show a predominant papillary growth. The neoplastic glands occur individually or have back-to-back growth. The glands are composed of a single layer of cuboidal to columnar cells with a varying amount of eosinophilic cytoplasm and a round to oval hyperchromatic nucleus. Nucleoli may be seen and are generally eccentrically located. The cells may have a prominent plasmacytoid appearance, particularly evident in the more solid areas of growth but also in the cells forming the glandular structures (**Figure 4; see colour plate section**). A paranuclear clear zone is not present. Often, adjacent to or intimately admixed with the

glands is a more solid or sheet-like growth of similar-appearing neoplastic cells. The cells may have a more dispersed or stippled nuclear chromatin with the 'salt and pepper' pattern suggestive of neuroendocrine differentiation. Cellular pleomorphism may be prominent but mitoses are uncommon. The stromal component is sparse and may appear fibrous or myxoid.

Histochemical stains show the presence of intraluminal but not intracytoplasmic mucin-positive material. Periodic acid Schiff (PAS)-positive material is not present. By immunohistohemical evaluation, the neoplastic cells are cytokeratin positive but are not reactive with chromogranin, synaptophysin, S-100 protein, desmin, actin or vimentin. Some MEA may have immunoreactivity with one or more neuroendocrine markers, including chromogranin and synaptophysin. These MEA with neuroendocrine differentiation have been termed carcinoid tumours of the middle ear (Latif *et al.*, 1987; Stanley *et al.*, 1987; Manni *et al*., 1992). However, these 'carcinoid tumours' are better viewed as part of the histologic spectrum of MEA (El Naggar *et al*., 1994), albeit one with neuroendocrine differentiation, rather than representing a distinct middle ear neoplasm separate from MEA.

MEA should be differentiated from glandular metaplasia that may occur in the setting of chronic otitis media (COM). These metaplastic glands may be misdiagnosed as neoplastic. In contrast to MEA, the glandular proliferation in COM is focal or haphazardly arrayed and occurs in the presence of histological features of COM, including chronic inflammation with fibrosis and calcifications (tympanosclerosis). MEA may perforate the tympanic membrane and appear to represent a neoplasm of the external auditory canal, such as a ceruminal gland adenoma. The histological features of these two tumour types are distinctly different and should allow for easy distinction. In contrast to the rare middle ear adenocarcinoma, MEA lack marked cellular pleomorphism, increased mitotic activity, necrosis or invasion of bone and other soft tissue structures.

The treatment for all MEA is complete surgical excision. Surgery may be conservative if the lesion is small and confined to the middle ear or more radical (mastoidectomy) for larger lesions associated with more extensive structural involvement. Recurrent tumour may occur and is a function of inadequate excision. Some MEA may be locally aggressive and rarely may invade vital structures causing death, but metastatic disease does not occur. In general, the clinical, radiological and pathological findings are indicative of a benign tumour. Nevertheless, the histological appearance is not always predictive of the clinical behaviour.

Jugulotympanic Paragangliomas (JTP)

JTP are benign neoplasms that arise from the extraadrenal neural crest-derived paraganglia specifically located in the middle ear or temporal bone region. Synonyms include glomus jugulare tumour and glomus tympanicum tumour. JTP are considered the most common tumour of the middle ear (Hyams *et al*., 1988b). JTP affect women more often than men and are most common in the fifth to seventh decades of life. The majority (85%) of the JTP arise in the jugular bulb resulting in a mass lesion in the middle ear or external auditory canal (Hyams *et al*., 1988b). Approximately 12% take origin from Jacobson's nerve (tympanic branch of the glossopharyngeal nerve) and present as a middle ear tumour (Hyams *et al*., 1988b). Approximately 3% arise from Arnold's nerve (posterior auricular branch of the vagus nerve) and arise in the external auditory canal (Hyams *et al*., 1988b). The most common symptom is conductive hearing loss. Other symptoms include tinnitus, fullness, otic discharge, pain, haemorrhage, facial nerve abnormalities and vertigo. JTP are often locally invasive neoplasms with extension into and destruction of adjacent structures including the temporal bone and mastoid. CT scan will show a soft tissue mass often with evidence of extensive destruction of adjacent structures. JGP are vascularized lesions that are fed by branches of nearby large arteries.

The gross appearance of JTP include a polypoid, red, friable mass identified behind an intact tympanic membrane or within the external auditory canal, measuring from a few millimetres to a large mass completely filling the middle ear space. Irrespective of the site of origin, the histological appearance of all extra-adrenal paragangliomas is the same. The hallmark histological feature is the presence of a cell nest or 'zellballen' pattern (**Figure 5; see colour plate section**). The stroma surrounding and separating the nests is composed of prominent fibrovascular tissue. While this pattern is characteristic of paragangliomas it is not unique to paragangliomas and can be seen in other tumours, including neuroendocrine carcinomas (carcinoid and atypical carcinoid tumours), melanomas and carcinomas. Paragangliomas are predominantly composed of chief cells which are round or oval cells with uniform nuclei, dispersed chromatin pattern and abundant eosinophilic, granular or vacuolated cytoplasm. The sustentacular cells, represent modified Schwann cells, are located at the periphery of the cell nests as spindle-shaped, basophilic appearing cells but are difficult to identify by light microscopy. Cellular and nuclear pleomorphism can be seen but these features are not indicative of malignancy. Mitoses and necrosis are infrequently identified. Paragangliomas lack glandular or alveolar differentiation.

Paragangliomas are often readily identified by light microscopic evaluation. However, in certain instances paragangliomas may be difficult to differentiate from other tumours that have similar histomorphological features. Not infrequently, middle ear and temporal paragangliomas do not show the characteristic cell nest appearance that occurs in other sites. This 'loss' of the organoid growth may be

artifactually induced by surgical manipulation ('squeezing') of the tissue during removal. The absence of the typical growth pattern may result in diagnostic confusion with other middle ear tumours.

Histochemical stains may be of assistance in the diagnosis of paragangliomas. Reticulin staining may better delineate the cell nest growth pattern with staining of the fibrovascular cores surrounding the neoplastic nests. In addition, the tumour cells are argyrophilic (Churukian–Schenk). Argentaffin (Fontana), mucicarmine and PAS stains are negative. The diagnosis of JTP is facilitated by immunohistochemical stains. The immunohistochemical antigenic profile of paragangliomas includes chromogranin and synaptophysin positivity in the chief cells and S-100 protein staining localized to the peripheral located sustentacular cells. Vimentin is variably reactive in both the chief cells and sustentacular cells. In general, epithelial markers, including cytokeratin, and also HMB-45 and mesenchymal markers (desmin and other markers of myogenic differentiation), are negative. Ultrastructural evaluation shows the presence of neurosecretory granules (Kliewer *et al*., 1989). DNA ploidy studies by image analysis are not predictive of the behaviour of paragangliomas (Barnes and Taylor, 1990).

Complete surgical excision is the treatment of choice, but the location and invasive nature of these lesions often preclude the ability to resect JTP completely. In such cases, radiotherapy is a useful adjunct to surgery. Radiotherapy results in a decrease or ablation of vascularity and promotes fibrosis. Preoperative embolization is useful in decreasing the vascularity of the tumour and facilitating surgical resection. Local recurrence of the tumour can be seen in as high as 50% of the cases. The histological appearance of paragangliomas does not correlate with the biological behaviour of the tumour. JTP are slow-growing tumours but may be locally invasive with extension into and destruction of adjacent structures, including the temporal bone and mastoid (Larson *et al*., 1987). Intracranial extension may occur in up to 15% of cases (Spector *et al*., 1975). Neurological abnormalities, including cranial nerve palsies, cerebellar dysfunction, dysphagia and hoarseness, may be seen and correlate with the invasive capabilities of this neoplasm. Functioning JTP as evidenced by endocrinopathic manifestations occur, but are extremely uncommon. Malignant JTP occur, are associated with histological criteria of malignancy, including increased mitotic activity, necrosis usually seen within the centre of the cell nests and vascular invasion, and may metastasize to cervical lymph nodes, lungs and liver.

Acoustic Neuroma (AN)

AN is a benign neoplasm that originates from Schwann cells specifically from the VIIIth cranial nerve. Synonyms include neurilemmoma, acoustic Schwannoma and benign peripheral nerve sheath tumour. AN accounts for up to 10% of all intracranial neoplasms and represent up to 90% of all cerebellopontine angle tumours (Hyams *et al*., 1988b). AN are more common in women than in men and may affect any age but are most common in the fourth to seventh decades of life. The majority of AN involve the superior or vestibular portion of the VIIIth nerve as compared with involvement of the cochlear portion of the VIIIth nerve. Symptoms include progressive (sensorineural) hearing loss, tinnitus and loss of equilibrium; with progression the tumour enlarges and may compress adjacent cranial nerves (V, VII, IX, X, XI), the cerebellum and the brainstem leading to facial paraesthesia and numbness, headaches, nausea, vomiting, diplopia and ataxia. Up to 8% of ANs may be bilateral (Erickson *et al*., 1965; Kasantikul *et al*., 1980; Martuza and Ojemann, 1982; Anand *et al*., 1993; Moffat and Irving, 1995). Bilaterality of ANs may represent a potential indicator of neurofibromatosis type 2 (Moffat and Irving, 1995; Rietz *et al*., 1983). Symptoms of neurofibromatosis may be seen in up to 16% of patients and those with neurofibromatosis who develop AN generally are symptomatic at an earlier age (second decade). The radiological appearance of AN include flaring, asymmetric widening or erosion of the internal auditory canal by CT or MRI. Tumours as small as 1 cm or less are capable of being detected by CT or MRI analysis.

The gross appearance of AN includes a circumscribed, tan–white, rubbery to firm mass which may appear yellow and have cystic change. Tumour sizes range from a few millimetres up to 4—5 cm in greatest diameter. Histologically, the tumours are unencapsulated and similar in appearance to benign Schwannomas of all other locations (**Figure 6; see colour plate section**). The cellular component includes elongated and twisted nuclei with indistinct cytoplasmic borders. The cells are arranged in short, interlacing fascicles and whorling or palisading of nuclei may be seen. Nuclear palisading with nuclear alignment in rows called Verocay bodies can be seen. The cellularity may vary and some benign Schwannomas can be very cellular (so-called cellular Schwannoma). Mitoses are usually sparse in number, and cellular pleomorphism with hyperchromasia can be identified but are not features of malignancy. Retrogressive changes, including cystic degeneration, necrosis, hyalinization, calcification and haemorrhage may be seen. Schwannomas have prominent vascularity composed of large vessels with thickened (hyalinized) walls.

Immunohistochemistry shows the presence of diffuse and intense S-100 protein reactivity. There is no immunoreactivity with cytokeratin or the neuroendocrine markers chromogranin and synaptophysin.

Complete surgical excision is the treatment of choice. Complete removal usually is curative. AN may result in death secondary to herniation of the brainstem in untreated and/or large neoplasms. Malignant AN are exceedingly rare and, if present, neurofibromatosis should be suspected.

Meningiomas

Meningiomas are benign neoplasms arising from arachnoid cells forming the arachnoid villi seen in relation to the dural sinuses. Meningiomas represent 13–18% of all intracranial tumours and are the second most common tumour to AN of the cerebellopontine angle (Hyams *et al.*, 1988b). Meningiomas are more common in women than in men and are most commonly seen in the fifth decade of life. Meningiomas infrequently occur in children. The occurrence of a meningioma outside the central nervous system is considered ectopic and can be divided into those meningiomas with no identifiable CNS connection (primary) and those with CNS connection (secondary). The development of primary meningiomas in the middle ear and temporal bone results either from direct extension or from the presence of arachnoid cells ectopically located. The most common sites of occurrence of the 'ectopically' located meningiomas is the head and neck region, specifically the middle ear and temporal bone, including the internal auditory canal, jugular foramen, geniculate ganglion, roof of the eustachian tube and sulcus of the greater petrosal nerve (Hyams *et al.*, 1988b). The clinical presentation of middle ear meningiomas includes progressive hearing loss, loss of equilibrium, headaches, cerebellar dysfunction and cranial nerve abnormalities. Patients with neurofibromatosis have an increased incidence of developing a meningioma. In addition, patients with neurofibromatosis also experience increased incidence of multiple, separate occurring meningiomas in intra- and extracranial sites. Radiological findings include a soft tissue mass with variable vascularity. A pathognomonic feature for meningioma in this location is the presence of speckled calcification in a soft tissue mass.

The histological features of middle ear and temporal bone meningioma are similar to those of their intracranial counterparts (**Figure 7; see colour plate section**). The immunohistochemical antigenic profile of meningiomas includes reactivity with epithelial membrane antigen (EMA) and vimentin. In contrast to middle ear adenomas, meningiomas are generally nonreactive with cytokeratin, and in contrast to jugulotympanic paragangliomas, meningiomas are nonreactive with neuroendocrine markers (e.g. chromogranin and synaptophysin).

Complete surgical excision is the treatment of choice and is curative. Local recurrence relates to inadequate excision. Malignant change rarely, if ever, occurs. A diagnosis of middle ear meningioma should be made only after clinical evaluation is made to exclude secondary extension from an intracranial neoplasm (Rietz *et al.*, 1983).

Endolymphatic Sac Papillary Tumour

The endolymphatic sac papillary tumour (ESPT) is an uncommon but distinct neoplasm possibly representing a manifestation of von Hippel–Lindau (VHL) syndrome (Megerian *et al.*, 1995; Manski *et al.*, 1997). ESPT has been referred to by a variety of names, including adenoma of endolymphatic sac, adenoma/adenocarcinoma of temporal bone or mastoid, low-grade adenocarcinoma of probable endolymphatic sac origin, papillary adenoma of temporal bone, aggressive papillary tumour of temporal bone, aggressive papillary middle ear tumour and, more recently, as the Heffner tumour (Batsakis and El-Naggar, 1993; Wenig and Heffner, 1996). An endolymphatic sac origin for these tumours is supported by a combination of findings, including the early clinical manifestations of vestibular disease (e.g. sensorineural hearing loss, tinnitus and episodic vertigo), radiographic features showing the tumour to grow in the region site where the endolymphatic sac is located (i.e. posterior–medial petrous ridge), intraoperative identification of an *in situ* tumour originating from within the endolymphatic sac and morphological similarities and shared immunohistochemical and ultrastructural features of the tumour with the normal endolymphatic sac epithelium (Wenig and Heffner, 1996). The diagnosis of this tumour is based on clinical, radiographic and pathological correlation. A diagnosis of ESPT should prompt the clinician to exclude the possibility that the patient has VHL syndrome (Megerian *et al.*, 1995; Manski *et al.*, 1997).

The histopathological appearance of ESPT is variable. ESPTs are papillary and focally cystic tumours. The papillary structures are generally not complex in their growth. The neoplastic cells vary in appearance from flattened or attenuated appearing cells to columnar appearing cells (**Figure 8; see colour plate section**). Most often there is only a single row of cells. Occasionally, the surface epithelial cells may have the appearance suggesting a double layer of cells (epithelial and myoepithelial); however, the 'outer' row of cells, in all probability, represent a stromal element as they have not been shown to be immunoreactive with epithelial markers (Heffner, 1989). The epithelial cells have uniform nuclei that are usually situated either in the centre of the cells or toward the luminal aspect, and have a pale eosinophilic to clear appearing cytoplasm. The latter may predominate in any given tumour. Cell borders may be seen but, not infrequently, the neoplastic cells lack a distinct cell membrane. In some cases, there are hypercellular areas with crowded, variably sized cystic glandular spaces that contain eosinophilic (colloid-like) material (**Figure 9; see colour plate section**). The latter appear remarkably similar to thyroid tissue. In all cases, pleomorphism is minimal, and mitotic activity and necrosis are rarely present.

A granulation tissue reaction is seen in association with the neoplastic cells and includes small vascular spaces lying in close proximity to the surface epithelium and/or within the stroma of the papillary fronds. Owing to the absence of a distinct cell membrane around the neoplastic cells, a sharp demarcation separating the neoplastic cells from the subjacent granulation tissue is not present. This appearance may create diagnostic confusion so that the

neoplastic proliferation is not appreciated, and the entire process is viewed as reactive. This interpretation is further enhanced by the presence in the stroma of a mixed inflammatory cell infiltrate, fibrosis, vascular proliferation, fresh haemorrhage and/or haemosiderin (within the neoplastic cells or within macrophages), cholesterol granulomas and dystrophic calcification. The latter does not include laminated calcific concretions (psammomatoid bodies).

Intracytoplasmic diastase-sensitive, PAS-positive material can be seen. The colloid-like luminal material stains strongly with PAS reagent with and without diastase digestion. Intracytoplasmic and intraluminal mucin staining is rarely positive. Iron stains are positive. ESPTs are diffusely cytokeratin positive and also show variable reactivity with epithelial membrane antigen (EMA), S-100 protein, vimentin, neuron-specific enolase (NSE), glial fibrillary acidic protein (GFAP), Ber-EP4, synaptophysin and Leu-7. Thyroglobulin immunoreactivity is not seen. Ultrastructurally, ESPT shows the presence of intercellular junctional complexes, microvilli, basement membrane material, rough endoplasmic reticulum and intracytoplasmic glycogen and secretory granules (Heffner, 1989).

The differential diagnosis includes middle ear adenoma. However, the clinical, radiographic and pathological features that are unique to ESPT should allow for its distinction from middle ear adenoma. The same would apply for the other common neoplasms of the middle ear and temporal bone. The differential diagnosis also includes choroid plexus papilloma and metastatic carcinoma of thyroid gland or renal origin. Choroid plexus papillomas are intracranial (i.e. intraventricular) tumours with histological features different from those of ESPT (Wenig and Heffner, 1996). The absence of thyroglobulin reactivity would differentiate ESPT from metastatic thyroid papillary carcinoma. Metastatic renal cell carcinoma would not have the immunohistochemical antigenic features seen in ESPT.

Radical surgery, including mastoidectomy and temporal bone resection that may necessitate sacrifice of cranial nerves, is the treatment of choice, and is potentially curative. Local recurrence will result following inadequate surgical removal; operative morbidity may be high. Despite their relatively slow growth, these neoplasms are capable of widespread infiltration and destruction, and may be lethal (Heffner, 1989). The prognosis is dependent on the extent of disease and the adequacy of resection. Earlier detection when the tumours are relatively small and confined may decrease the operative-associated morbidity and be curative.

Malignant Neoplasms of the Middle Ear

Primary malignant neoplasms of the middle ear are extremely rare. **Table 1** lists some of these malignant neoplasms. The discussion of malignant neoplasms of the middle ear will be limited to epithelial malignancies, including squamous cell carcinoma and adenocarcinoma, and a brief discussion on rhabdomyosarcoma of this region.

Middle Ear Squamous Cell Carcinoma (ME-SCC)

Primary malignant neoplasms with squamous differentiation originating from the middle ear mucosal epithelium are rare (Kenyon *et al*., 1985; Hyams *et al*., 1988c). ME-SCC is most common in the sixth and seventh decades of life. The majority of patients have a long history of chronic otitis media usually greater than 20 years in duration. The development of ME-SCC is also linked to radiation treatment for intracranial neoplasms and, although no longer used, radiotherapy for middle ear inflammatory conditions. Concomitant cholesteatomas can be seen in up to 25% of cases but there is no correlation between cholesteatomas and the development of a middle ear squamous cell carcinoma (Hyams *et al*., 1988c). ME-SCC should be suspected in patients with a long-standing chronic otitis media who present with sudden onset of pain out of proportion to the clinical extent of disease, onset or increase of otorrhea which is often haemorrhagic and/or a lack of clinical resolution following therapeutic doses of antibiotics.

The histology of ME-SCC is similar to that of squamous carcinomas of other sites. The tumours vary from well to poorly differentiated and include infiltrative malignant cells with associated keratinization and/or intercellular bridges.

The differential diagnosis includes a cholesteatoma and metastatic squamous cell carcinoma. Cholesteatomas do not have the dysplastic cytological changes seen in squamous carcinoma. Secondary involvement of this area by squamous cell carcinoma may originate from a distant site and metastasize to the middle ear and temporal bone. Alternatively, a cutaneous squamous cell carcinoma from an adjacent site (external ear, nasopharynx, parotid gland or skin) can directly invade into the middle ear or temporal bone. Detailed clinical history or physical examination would assist in identifying a squamous carcinoma that is metastatic to this site or extends to the middle ear from an adjacent primary tumour.

Radical surgery with radiotherapy is the treatment of choice. In advanced disease, chemotherapy may be of benefit. Prognosis is poor with 5- and 10-year survival rates of 39% and 21%, respectively (Hyams *et al*., 1988c). Metastases may occur but are considered uncommon.

Middle Ear Adenocarcinoma

Middle ear adenocarcinoma is a malignant glandular neoplasm arising from the middle ear mucosa. Middle ear adenocarcinomas are rare (Hyams *et al*., 1988d). In the presence of a malignant glandular neoplasm of the middle

ear and temporal bone, secondary metastasis to this region should be excluded. Middle ear adenocarcinomas may attain large sizes, filling the middle ear space and encasing the ossicles. Symptoms are typically present for many years and include progressive hearing loss and a unilateral draining ear; pain and vestibular manifestations are uncommon. Otoscopic examination in the majority of cases will identify an intact tympanic membrane with tumour confined to the middle ear space with possible extension to the mastoid. Occasionally, the adenocarcinoma will perforate through the tympanic membrane with extension into and presentation as an external auditory canal mass. There is no known association between chronic otitis media and the development of these adenocarcinomas.

Histologically, middle ear adenocarcinomas are in many respects similar to adenomas. In contrast to adenomas, adenocarcinomas have increased cellular pleomorphism, increased mitotic activity and extensive infiltration of surrounding soft tissue structures involving nerves, lymph-vascular spaces and bone.

Complete surgical excision is the treatment of choice. In general, these are slow-growing neoplasms that are locally aggressive but do not metastasize. Death may occur as a result of direct intracranial extension. Confinement to the middle ear space and association with the middle ear mucosa are supportive evidence of origin from the middle ear; nevertheless, metastatic adenocarcinoma from a separate site must be excluded prior to treatment.

Rhabdomyosarcoma (RMS)

In the head and neck, RMS is primarily but not exclusively a disease of the paediatric population. In children and adolescents, RMS represents the most common aural-related malignant neoplasm. There is no gender predilection. RMS of the middle ear and mastoid presents as painless unilateral otitis media unresponsive to antibiotic therapy. According to the WHO classification, RMS is divided into six histological subtypes, including embryonal, botryoid, spindle cell, alveolar, pleomorphic and RMS with ganglionic differentiation (so-called ectomesenchymoma) (Weiss, 1995). The International classification of RMS proposed four groups based on prognosis: I, superior prognosis (botryoid RMS and spindle cell RMS); II, intermediate prognosis (embryonal RMS); III, poor prognosis (alveolar RMS and undifferentiated RMS); and IV, subtypes whose prognosis is not presently evaluable (RMS with rhabdoid features) (Newton *et al*., 1995).

The majority of RMS of the middle ear and mastoid are of the embryonal type that includes botryoid RMS. The next most common histological type is alveolar RMS. The other histological types may occur in the head and neck but are considered uncommon. RMS of the middle ear and mastoid most often appears as an aural (external or middle ear) polypoid lesion similar in appearance to an aural polyp.

RMS is treated by a combination of surgery, radiation and chemotherapy. This combined therapeutic approach has greatly enhanced survival with an overall 5-year survival of 74% for paediatric head and neck RMS (Kraus *et al*., 1997). A problem with middle ear and mastoid RMS is the delay in diagnosis due to misinterpretation of the biopsy specimen as inflammatory polyps or as granulation tissue. This delay in diagnosis may result in more advanced stage disease, placing patients at greater risk of treatment failure owing to uncontrollable local disease. Poor prognostic findings include meningeal involvement (Raney *et al*., 1987). Regional lymph node metastasis and distant haematogenous metastasis to the lungs and bones may also occur.

Secondary Tumours

Metastatic tumours secondarily involving the middle ear and temporal bone originate from virtually every site. The more common malignant tumours to metastasize to this region originate from the breast, lungs and kidneys (Hill and Kohut, 1976; Schuknecht, 1993b). Other tumours that may metastasize to this region include malignant melanoma and prostatic adenocarcinoma. While metastases to the temporal bone often occur late in the disease course, metastatic involvement of the temporal bone may represent the initial presentation of a distant malignant disease. Metastatic disease to the temporal bone occurs via haematogenous spread but may also occur by direct extension from a nearby primary tumour (e.g. squamous cell carcinoma), meningeal carcinomatosis or leptomeningeal extension from an intracranial primary neoplasm (Berlinger *et al*., 1980).

REFERENCES

Abramson, M., *et al*. (1984). Histology, pathogenesis, and treatment of cholesteatoma. *Otolaryngology, Rhinology and Laryngology*, **112**, 125–128.

Anand, T., *et al*. (1993). Bilateral acoustic neuromas. *Clinical Otolaryngology*, **18**, 365–371.

Barnes, L. and Taylor, S. R. (1990). Carotid body paragangliomas: a clinicopathologic and DNA analysis of 13 cases. *Archives Otolaryngology and Head and Neck Surgery*, **116**, 447–453.

Batsakis, J. G. (1989). Adenomatous tumors of the middle ear. *Annals of Otolaryngology, Rhinology and Laryngology*, **98**, 749–752.

Batsakis, J. G., and El-Naggar, A. K. (1993). Papillary neoplasms (Heffner's tumors) of the endolymphatic sac. *Annals of Otolaryngology, Rhinology and Laryngology*, **102**, 648–651.

Berlinger, N. T., *et al*. (1980). Patterns of involvement of the temporal bone in metastastic and systemic malignancy. *Laryngoscope*, **90**, 619–627.

Dayal, V. S., *et al*. (1973). Embryology of the ear. *Canadian Journal of Otolaryngology*, **2**, 136–142.

Desloge, R. B., *et al*. (1997). DNA analysis of human cholesteatomas. *American Journal of Otolaryngology*, **18**, 155–159.

El-Naggar, A. K., *et al*. (1994). Tumors of the middle ear and endolymphatic sac. *Pathology Annual*, **29**, 199–231.

Erickson, L. S., *et al*. (1965). A review of 140 acoustic neurinomas (neurilemmoma). *Laryngoscope*, **75**, 601–627.

Faverly, D. R. G. S., *et al*. (1992). Adenocarcinoid or amphicrine tumors of the middle ear. *Pathology Research and Practice*, **188**, 162–171.

Ferlito, A., *et al*. (1997). Ear cholesteatoma versus cholesterol granuloma. *Annals of Otolaryngology, Rhinology and Laryngology*, **106**, 79–85.

Heffner, D. K. (1989). Low-grade adenocarcinoma of probable endolymphatic sac origin. A clinicopathologic study of 20 cases. *Cancer*, **64**, 2292–2302.

Hicks, G. W. (1983). Tumors arising from the glandular structures of the external auditory canal. *Laryngoscope*, **93**, 326–340.

Hill, B. A. and Kohut, R. I. (1976). Metastatic adenocarcinoma of the temporal bone. *Archives of Otolaryngology*, **102**, 568–571.

Hyams, V. J., *et al*. (1988a). Adenomatous neoplasms of ceruminal gland origin. In: Hartman, X. X. and Sobin, L. H. (eds), *Tumors of the Upper Respiratory Tract and Ear. Atlas of Tumor Pathology*, Fascicle, 25. 285–291 (Armed Forces Institute of Pathology, Washington, DC).

Hyams, V. J., *et al*. (1988b). Neoplasms of the middle ear. In: Hartman, X. X. and Sobin, L. H. (eds), *Tumors of the Upper Respiratory Tract and Ear. Atlas of Tumor Pathology*, Fascicle 25. 306–330 (Armed Forces Institute of Pathology, Washington, DC).

Hyams, V. J., *et al*. (1988c). Squamous cell carcinoma of the middle ear. In: Hartman, X. X. and Sobin, L. H. (eds), *Tumors of the Upper Respiratory Tract and Ear. Atlas of Tumor Pathology*, Fascicle 25. 326–327 (Armed Forces Institute of Pathology, Washington, DC).

Hyams, V. J., *et al*. (1988d). Adenocarcinoma of the middle ear. In: Hartman, X. X. and Sobin, L. H. (eds), *Tumors of the Upper Respiratory Tract and Ear. Atlas of Tumor Pathology*, Fascicle 25. 302–323 (Armed Forces Institute of Pathology, Washington, DC).

Kasantikul, V., *et al*. (1980). Acoustic neurilemmoma. Clinicoanatomical study of 103 patients. *Journal of Neurosurgery*, **52**, 28–35.

Kenyon, G. S., *et al*. (1985). Squamous cell carcinoma of the middle ear; a 25-year retrospective study. *Annals of Otolaryngology, Rhinology and Laryngology*, **94**, 273–277.

Kliewer, K. E., *et al*. (1989). Paragangliomas: assessment of prognosis by histologic, immunohistochemical, and ultrastructural techniques. *Human Pathology*, **20**, 29–39.

Kraus, D. H., *et al*. (1997). Pediatric rhabdomyosarcoma of the head and neck. *American Journal of Surgery*, **174**, 556–560.

Larson, T. C., *et al*. (1987). Glomus tympanicum chemodectomas: radiographic and clinical characteristics. *Radiology*, **163**, 801–806.

Latif, M. A., *et al*. (1987). Carcinoid tumour of the middle ear associated with systemic symptoms. *Journal of Laryngology and Otolaryngology*, **101**, 480–486.

Manni, J., *et al*. (1992). Primary carcinoid tumor of the middle ear: report of four cases and a review of the literature. *Archives of Otolaryngology and Head and Neck Surgery*, **118**, 1341–1347.

Manski, T. J., *et al*. (1997). Endolymphatic sac tumors: the basis of morbid hearing loss in von Hippel–Lindau disease. *Journal of American Medical Association*, **277**, 1461–1466.

Martuza, R. L. and Ojemann, R. G. (1982). Bilateral acoustic neuromas: clinical aspects, pathogenesis and treatment. *Neurosurgery*, **10**, 1–12.

Megerian, C. A., *et al*. (1995). Endolymphatic sac tumors: histopathologic confirmation, clinical characterization, and implication in von Hippel–Lindau disease. *Laryngoscope*, **105**, 801–808.

Moffat, D. A. and Irving, R. M. (1995). The molecular genetics of vestibular schwannomas. *Journal of Laryngology and Otolaryngology*, **109**, 381–384.

Moore, K. L. (1988). The ear. In: Moore, K. L. (ed.), *The Developing Human: Clinically Oriented Embryology*. 412–440 (W. B. Saunder, Philadelphia).

Newton, W. A., *et al*. (1995). Classification of rhabdomyosarcomas and related sarcomas: pathologic aspects and proposal for a new classification – an intergroup rhabdomyosarcoma study. *Cancer*, **76**, 1073–1085.

Perzin, K. H., *et al*. (1982). Adenoid cystic carcinoma involving the external auditory canal. A clinicopathological study of 16 cases. *Cancer*, **50**, 2873–2883.

Pulec, J. L. (1977). Glandular tumors of the external auditory canal. *Laryngoscope*, **87**, 1601–1612.

Raney, R. B., Jr, *et al*. (1987). Improved prognosis with cranial soft tissue sarcomas arising in nonorbital parameningeal sites. A report from the intergroup rhabdomyosarcoma study. *Cancer*, **59**, 147–155.

Rietz, D. R., *et al*. (1983). Significance of apparent intratympanic meningiomas. *Laryngoscope*, **93**, 1397–1404.

Schuknecht, H. F. (1993a). Cholesteatoma. In: Schuknecht H. F. (ed.), *Pathology of the Ear*. 204–206 (Lea & Febiger, Philadelphia).

Schuknecht, H. F. (1993b). Neoplastic growths. In: Schuknecht, H. F. (ed.), *Pathology of the Ear*. 447–448 (Lea & Febiger, Philadelphia).

Spector, G. J., *et al*. (1975). Glomus tumors in the head and neck. III. Analysis of clinical manifestations. *Annals of Rhinology, Otolaryngology and Laryngology*, **84**, 73–79.

Stanley, M. W., *et al*. (1987). Carcinoid tumors of the middle ear. *American Journal of Clinical Pathology*, **87**, 592–600.

Weiss, S. W. (1995). World Health Organization international histological classification of tumours. In: Weiss, S. W. (ed.), *Histological Typing of Soft Tissue Tumours*. (Springer, Berlin).

Wenig, B. M. and Heffner, D. K. (1996). Endolymphatic sac tumors: fact or fiction ? *Advances in Anatomical Pathology*, **3**, 378–387.

Wetli, C. V., *et al*. (1972). Tumors of ceruminous glands. *Cancer*, **29**, 1169–1178.

FURTHER READING

Hollinshead, W. H. (1982). The ear. In: Hollinshead, W. H. (ed.), *Anatomy for Surgeons*, Vol. 1. 159–221 (Harper and Row, Philadelphia).

Nager, G. T. (1993). *Pathology of the Ear and Temporal Bone* (Williams & Wilkins, Baltimore).

Schuknecht, H. F. (1993). *Pathology of the Ear* (Lea & Febiger, Philadelphia).

Chapter 54

Advantages and Limitations of Models for Cancer and Malignant Cell Progression

Garth L. Nicolson and Marwan Y. Nasralla
Institute for Molecular Medicine, Huntington Beach, CA, USA

CONTENTS

- Introduction to Cancer Models
- The Use of Breast Cancer Models to Study Tumour Progression
- Oncogenes and Suppressor Genes in Tumour Progression
- Host Surveillance Mechanisms and Tumour Progression
- Cancer Progression and Breast Cancer Gene Expression
- Intercellular Communication and Tumour Cell Diversification
- Cancer Progression and Cellular Diversification

INTRODUCTION TO CANCER MODELS

Since the beginning of studies on the pathogenesis of various cancers, it was recognized that little progress would be made in understanding the molecular and genetic events that define the precancerous states, the inception of cancers, sometimes called transformation, and its further progression to invasive and metastatic phenotypes, unless these events could be duplicated in some reproducible way. Thus the need for cancer models was born. There are many different types of cancer models, and our purpose here is not to review them. Rather, we will summarize the various scientific and medical needs for cancer models and what has been accomplished with them and use as an example some models for breast cancer. This in no way is meant to overemphasize the importance of breast cancers over other types of cancers. It simply reflects an example of an important need to understand various molecular and genetic aspects of an important class of cancers because it is a leading cause of death of women in the Western world.

Cancer models are useful for accurately and reproducibly studying a variety of pathogenic, biochemical, immunological, pharmacological, genetic and other properties of cancer cells. Although it would be useful if we all agreed on using one appropriate model system, with the use of only one model it is extremely difficult to study accurately all of the above properties. This is due to the obvious fact that cancer models are ‘models’ and not ‘real’ cancers obtained directly from cancer patients. As such, they are only approximations of reality, and they should not be confused with the real world, although they may have been directly derived from real cancers growing in patients. In addition, certain models are particularly useful for certain types of experiments but not for others, and it is particularly difficult to find one model that can substitute for the number of models that are currently in use to study the properties of breast cancers or any other type of cancer. For hundreds of years pathologists have warned us of the diversity and heterogeneity of clinical cancers, and this diversity cannot be easily reflected in one or even a few models of clinical cancers (Nicolson and Poste, 1983).

The worth of a cancer model is reflected in how accurately and reproducibly the model is able to mimic the ‘real world,’ with the understanding that it may not accurately reflect all of the properties of the *in situ* cancer in its natural human host. In addition, one of the commonly found properties of malignant cancers is that they tend to be quite heterogeneous at the cellular level, and the known instabilities of cancers and cancer cells can result in the continuous diversification and evolution of cancers and, in particular, the cells that comprise them (Nowell, 1976; Poste, 1982; Nicolson, 1986, 1987). Fortunately, or unfortunately depending on the need for exacting reproducibility, most if not all cancer models share with legitimate clinical cancers the properties of cellular instability and diversification into heterogeneous cellular populations during growth, and this is even more apparent in malignant cancers that have progressed to the metastatic phenotype. This instability, although a critical property for any model of human cancers, is also the bane of the experimentalist.

Thus some of the most important properties of cancers, such as their abilities to undergo change, progression, diversification and evolution, are also properties that make the reproducible study of cancers and especially cancer cells so difficult (Poste, 1982; Nicolson, 1987).

Notwithstanding the problems associated with the development and use of cancer models and also the interpretation of data obtained using them, they have proved their worth over the years. In fact, we would know considerably less about the detailed biochemical and immunological and other properties of cancers if we were entirely dependent on using clinical biopsies for experimental studies. Thus cancer models have proved to be an essential element in our research programmes, and without them our knowledge of human cancers, particularly their molecular properties, would be significantly less than it is today. But even with that it should be obvious even to the casual observer that knowing the limitations of cancer models is essential, and cancer models are likely to be as heterogeneous and thus problematic in their utility and usefulness as the cancers from which they were derived (**Table 1**).

It has been said that not all cancer models are useful because they do not closely mimic all of the properties of human cancers. On the other hand, even a model that does not closely mimic human cancer can be useful for some experimental uses, particularly those that are not designed to be applied eventually to clinical use (**Table 1**). Thus it depends greatly on the use for which the model is to be employed. One model may be useful as a cellular repository of certain critical molecules or biochemical platform for certain types of experiments but not entirely useful for other purposes. Hence it depends greatly on the use of the model. The trend in recent years, however, has been to try to develop and use models that reflect as accurately as possible the *in vivo* and *in vitro* characteristics of human cancers so that the information gathered may at some later date be applied to clinical cancers. Of critical need is the use of models to develop new therapeutic approaches for treating highly malignant tumour cells (Frei, 1982).

Cancer models are especially useful for studying certain specific aspects of the individual steps of the cancer progression process, and these models for the most part are *in vitro* and animal models. For example, various models have been developed to study specific properties of cancer cells that are related to progression to the metastatic phenotype. For example, malignant cell adhesion, invasion, survival and growth have been successfully studied *in vitro* and in animal models (**Table 2**). The utility of these *in vitro* models is that detailed biochemical studies can be performed that could not be easily done with *in vivo* systems. Although such models do not exactly mimic the events that occur *in vivo* during the process of metastasis, they are extremely useful for obtaining detailed biochemical information on cellular processes that are related to cancer progression and metastasis.

Models for human cancers have proved useful in a variety of settings where comparisons are necessary. For example, we developed an animal model for breast cancer that was based on a transplantable rat mammary 13762NF adenocarcinoma (Neri *et al*., 1982). This model not only fulfills the criteria that it displays similar pathogenic properties to human breast adenocarcinomas, but it also possesses similar biochemical, immunological and enzymatic properties, cytocellular structures and other characteristics, such as dynamic heterogeneity in drug, radiation and heat sensitivities associated with tumour cell heterogeneity and instability (review: Nicolson, 1994). (See also chapter on *Mammary Tumour Induction in Animals as a Model for Human Breast Cancer*.) When cells from this adenocarcinoma model were injected into the mammary fat pads of syngeneic rats, they first metastasized regionally to axial lymph nodes and then to lungs in a relatively reproducible pattern (Neri and Nicolson, 1981). Interestingly, when the properties of metastatic clones derived from this model

Table 1 Models for cancer and cancer progression: which models are best?

1. What questions are being asked? Do these involve eventual application to the clinic?
2. Should human cancers or cancer cells be used, or will animal tumours or cells suffice?
3. If human cancers/cells are being used, are these representative of the original cancers?
4. If animal tumours/cells are being used, are they of similar histology, natural history, etc., to the human tumours/cells?
5. If animal tumours/cells are being used, do they grow and undergo progression similar to human tumours/cells?
6. If animal tumours/cells are being used, do they have similar changes in molecular genetics?
7. If the tumours/cells are to be assayed *in vivo*, is a syngeneic tumour/host system required?
8. Can human cancers/cells be assayed effectively in immune-compromised animal hosts?
9. Are counterpart normal, preneoplastic, benign or invasive/metastatic tumours/cells necessary for comparison?
10. Are counterpart tumours/cells of similar genetic background necessary for comparison?
11. Will *in vivo* or *in vitro* assays, or both, be used? Is *in vitro* growth required?
12. Are the cancer/tumour cells relatively stable during growth *in vivo* and/or *in vitro*?
13. Are the cancer or tumour cells of similar biological phenotype to the original tumour?
14. Are the cancer or tumour cells of similar biochemistry, immunology, drug sensitivity and other properties to the original tumour?
15. Is the pathogenesis of progression and metastasis similar to the original tumour?
16. Can the model be used to develop relevant new therapies that can be applied to the clinic?

Table 2 Use of models to study individual steps of the metastatic process

1. Tumour growth at a primary site
    - A. Cell cultures (single or mixed cell)
    - B. Cell aggregates (single or mixed cell)
    - C. Implantation at a primary site (animal)
2. Primary tumour invasion
    - A. Invasion of extracellular matrix
    - B. Invasion of thrombin clot
    - C. Invasion of tissue fragments
    - D. Invasion of cell aggregates
    - E. Implantation at a primary site (animal)
3. Tumour cell intravasation
    - A. Invasion of isolated blood vessel
    - B. Invasion of endothelial cell monolayer
    - C. Intravasation of blood vessels (*in vitro*)
    - D. Intravasation of blood vessels (animal)
4. Tumour cell transport
    - A. Transport in a flow cell
    - B. Transfilter migration in a Boyden chamber
    - C. Dynamic stress in a parallel plate chamber
    - D. Transport after i.v. injection (animal)
5. Tumour cell arrest
    - A. Adhesion to endothelial cells
    - B. Adhesion to endothelial cell monolayers
    - C. Adhesion to blood vessel (*in vitro*)
    - D. Arrest after i.v. injection (animal)
6. Tumour cell extravasation
    - A. Invasion of endothelial cell monolayers
    - B. Invasion of blood vessels (*in vitro*)
    - C. Invasion of blood vessels (animal)
7. Tumour cell survival
    - A. Survival in depleted medium
    - B. Survival in organ-conditioned medium
    - C. Survival in immune cell cultures
    - D. Survival after i.v. injection (animal)
8. Tumour cell growth at secondary sites
    - A. Growth in depleted medium
    - B. Growth in organ-conditioned medium
    - C. Growth in immune cell cultures
    - D. Growth after i.v. injection (animal)
    - E. Growth after injection at secondary organ site (animal)

were compared with nonmetastatic cell clones, differences in various biochemical and immunological properties were found to be similar to the molecular differences found in human breast carcinomas derived from primary and metastatic sites (Nakajima *et al*., 1987; Steck *et al*., 1987a,b). Hence this model can be used for a variety of experiments that cannot be done easily with human cancers or biopsies. Similarly, rat mammary tumour models have proved to be very useful in studying the process of development and differentiation of mammary gland tissue (Rudland *et al*., 1998). In this case the interactions of epithelial, myoepithelial and other cell types in the developmental formation of glandular structures were studied *in vitro* and *in vivo* and then related to similar events that occur in human mammary tissues.

Models for breast cancer have also been developed based on murine mammary tumour cell lines (Heppner, 1989). Using various cloned cell lines developed from the same mammary tumour clonal interactions influenced such properties as growth, drug sensitivity and malignancy (Miller and Heppner, 1990). Although these mouse models do not spread regionally before haematogenous colonization, they have other attributes that make them useful as models for breast cancer. The rat and mouse mammary tumour models show many similarities to human breast cancer cells that have been sequentially selected for invasive properties (Thompson *et al*., 1993).

Of course, not all of the properties of human and animal mammary cells will be the same, and not all properties of *in situ* tumours can be mimicked by the same tumour cells growing *in vitro*. The reasons for this are probably related to differences in the cellular compositions of tumours compared with cell populations growing *in vitro* and differences in environment that can affect many of the cellular properties of tumour and normal cells (Miller and Heppner, 1990; Rudland *et al*., 1998). (See also chapter on *Modelling Tumour–Tissue Interactions*.) Important in this consideration is when human cancer cells are grown in animals (usually in nude mice) that they are implanted into the appropriate site (orthotropic implantation) (Fidler *et al*., 1990). For breast cancer cells or tissue, this would usually be the mammary fat pads of rodents, but for other cancers arising in other tissues the tissue of origin would apply. Unfortunately, for human breast cancer cells growing in the mammary fat pads of rodents the microenvironment is not the same as in the human breast, and differences may ensue. In this brief introduction we will discuss the use of breast cancer and other models in ascertaining the critical molecular determinants of various steps of the metastatic process and some of the general properties of malignant breast tumours.

THE USE OF BREAST CANCER MODELS TO STUDY TUMOUR PROGRESSION

There are many important questions in cancer research that can best be answered with the use of tumour models. One of most important is explaining the process of tumour progression or the tendency of tumours to undergo gradual change to more malignant phenotypes, and the use of cancer models to study this process will be summarized. Tumour progression is characterized by specific genetic changes, and these can be studied using animal tumour models that undergo progression. The accumulation of rare qualitative genetic changes is thought to drive tumour progression (Nowell, 1976; Nicolson, 1986, 1991; Moustafa and Nicolson, 1998). These changes include but are not exclusive to DNA sequence alterations, such as gene amplifications, mutations, deletions, translocations and other changes. The accumulation of genetic changes or alterations

may initiate important events that eventually lead to tumour progression, but it is likely that other events are also involved that do not depend on genomic DNA sequence changes (Nicolson, 1986, 1991; Moustafa and Nicolson, 1998). One of the uses of cancer models is to be able to determine the importance of specific genetic events in transformation and progression by replicating these events in an experimental setting. In addition, once candidate genes have been identified, their role in the pathogenesis of cancer can be confirmed using transgenic and gene knockout mice to model the changes associated with specific genes (review: Viney, 1995). (See also chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*.)

The progression of breast cells to the malignant phenotype is mainly typified by quantitative changes in gene expression rather than many qualitative changes (mutations, translocations, truncations, etc.) in gene structures, although the latter changes are known to occur with progression (Nicolson, 1991; Moustafa and Nicolson, 1998). In breast cancer models such qualitative changes have also been documented, especially in genes associated with hereditary breast cancer, such as *BRCA-1* and *BRCA-2* (Lynch *et al*., 1998). Using breast cancer models, parallel studies can be used to examine the role of multiple qualitative cytogenetic and genetic changes and determine if they are important in the process of transformation and progression (Pearce *et al*., 1984).

Qualitative genomic modifications may occur only rarely in tumours, and the rates of appearance of qualitative alterations may vary widely among different tumours. Hence it is important to determine if there are (multiple) critical changes in specific genes that must occur in order for transformation and progression to occur. If qualitative genomic changes are critical to tumour transformation, progression and clonal dominance, then it is likely that the most successful cells in a tumour that have these changes will overgrow the other cells and eventually almost all of the tumour cells within a tumour would display similar genomic modifications (Nowell, 1976). Probably not all genes that undergo change are important in the process of tumour progression, and the types of genes that are most likely to be important have been found to be the genes that encode oncogenes, suppressor genes, differentiation genes and genes associated with growth, invasion, survival and metastasis (Nicolson, 1991; Moustafa and Nicolson, 1998). These are also the genes that are likely to also be changed in their expression (Volpe, 1988; Moustafa and Nicolson, 1998). Often important qualitative changes that are associated with early states of oncogenesis, such as cellular transformation and tumour initiation, are not the same changes that are associated with progression to more malignant states (Nicolson, 1986, 1991; Volpe, 1988; Moustafa and Nicolson, 1998).

The changes in gene expression and gene structure that occur in breast cancer cells as they progress are not the only important events that typify tumour progression. There are also signals from the host tissue microenvironment (cells, extracellular matrix, soluble glyoproteins and other signals) that are important in this process (Nicolson, 1986; 1991; Volpe, 1988; Moustafa and Nicolson, 1998). These signals provide tumours with soluble and insoluble mediators that can modulate tumour cell properties and responses to host tissue, cell and extracellular matrix signals (Bissell and Barcellos-Hoff, 1987). In some tissues, such as breast eipthelium, there are also unique regulators, such as the ionic and metabolic signals that are presented to adjacent cells, either at their surfaces or through junctional communication in epithelium. There are both external and cellular signals that elicit both positive and negative signals that usually control the normal aspects of epithelial tissue differentiation, proliferation and death, and these are important in tumour progression and growth (Nicolson, 1989, 1993; Trosko *et al*., 1993).

As tumours progress, particularly breast and mammary tumours (Foulds, 1975), they are thought to be less responsive to host microenvironments and cellular controls. The end result of this is that they eventually gain autonomy from cell and tissue regulation as they progress to the late or end stages of their life histories (Nicolson, 1986, 1987, 1993). An important property of malignant tumours is their ability to undergo cellular diversification into heterogeneous phenotypes. Heterogeneity in cellular properties is found in virtually all malignant tumour cell populations, but in highly malignant tumours, such as invasive breast carcinomas, it is usually more pronounced than that found in the cells of benign or normal breast tissues. In normal tissues intercellular, cellular and matrix interactions probably combine to stabilize cellular phenotypes into more narrow states of diversity than seen in isolated single cells or tumour cells derived from the same tissue (Nicolson, 1986, 1987, 1993). Once normal cells have been removed from their normal interactions, they show increased diversity in their cellular properties (Rubin, 1990). Such diversity may be due to adoptive changes that individually affect each cell and result in individual quantitative differences in gene expression. In malignant cell populations diversification occurs irrespective of or at least less dependent on the host microenvironment. Eventually, heterogeneous cellular phenotypes arise that are less regulated by normal host cell, matrix and other interactions (Nicolson, 1987, 1993).

In mammary as in other tumour cell populations, single cells or a few cells undergo diversification to more heterogeneous phenotypes, and as this happens these cell populations also face host immune and nonimmune selection pressures. Eventually differences in cellular responses to host mediators, growth stimulators or inhibitors allow certain subpopulations to become dominant in the tumour cell population, or by active processes, such as immune or nonimmune host responses, certain subpopulations arise in the population to become the predominant cellular type (Nowell, 1976; Nicolson, 1987, 1991; Kerbel *et al*., 1988;

Frost and Chernajovsky, 1990). This can result in multiple cycles of diversification and subsequent host selection of tumour cells until dominant malignant cell subpopulations emerge that display highly malignant and autonomous phenotypes. Thus tumour progression probably results in waves of cellular diversification and then restriction of diversity (clonal dominance) (Kerbel *et al*., 1988; Nicolson, 1991) until malignant cell subpopulations become dominant that contain the essential properties to be highly self-sufficient and malignant (**Figure 1**).

The malignant cell progression process, although ominous, is not abnormal. It is probably a normal adaptive process that is important in homeostasis. This may explain why malignant cell characteristics are not restricted to cancer cells. In normal tissues there are some highly motile, invasive normal cells, usually embryonic in origin, that are capable of autonomous survival and growth in different tissues (Armstrong, 1984). In normal adult tissues, moreover, there is evidence that injury can initiate the events necessary for converting sessile, quiescent cells into motile, invasive cells capable of autonomous cellular division, e.g. during angiogenesis or the development of a new vascular system to feed the injured tissue (Folkman and Shing, 1992).

ONCOGENES AND SUPPRESSOR GENES IN TUMOUR PROGRESSION

Qualitative changes in certain oncogenes and suppressor genes are often found in mammary and other tumours. For example, in breast and colorectal cancers the accumulation of multiple, different qualitative genetic changes in oncogenes and suppressor genes typifies cancer progression (Fearon and Vogelstein, 1990), but the range of genetic alterations found in each state suggests that other changes, among them quantitative differences in gene expression, may also be important. This is especially apparent in the most progressed malignant states (Moustafa and Nicolson, 1998). Oncogenes encode proteins that function

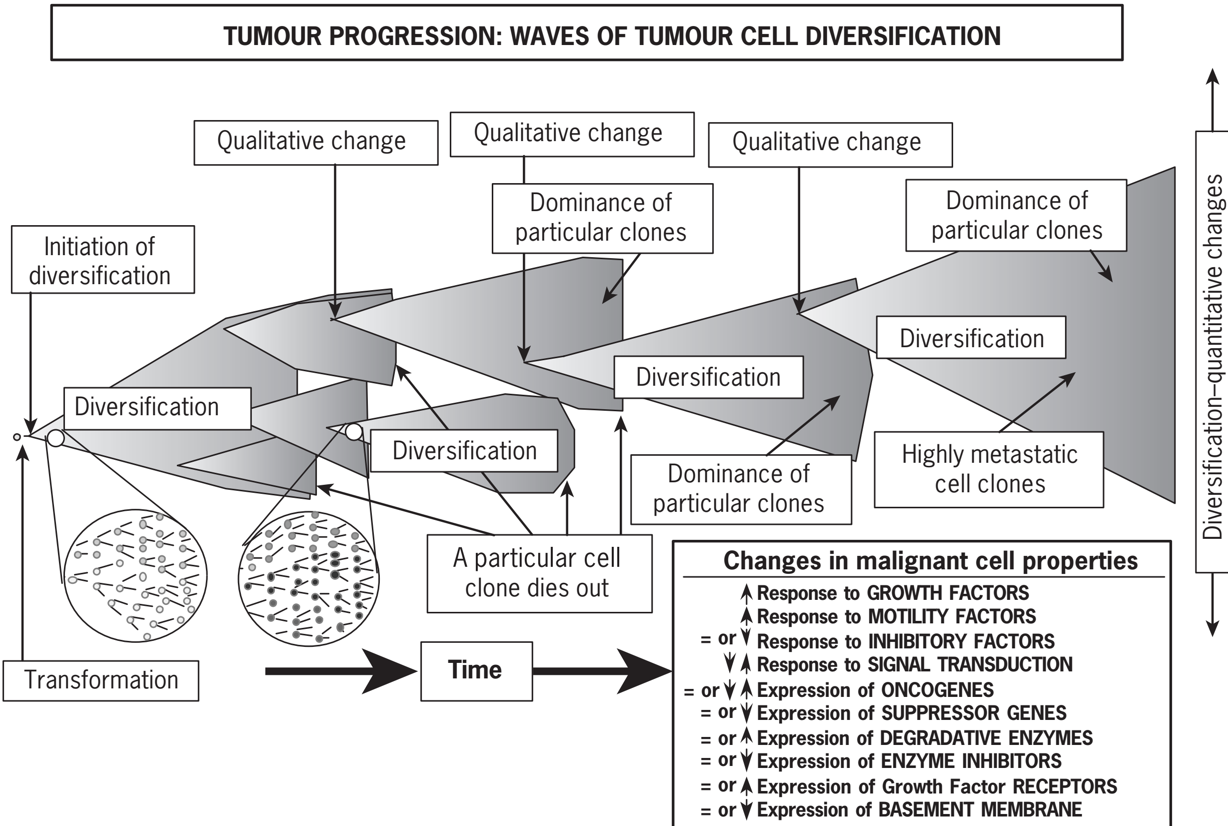


Figure 1 Example of how qualitative alterations in gene sequence and quantitative changes in gene expression could be related to tumour cell diversification and progression. In this diagram a single cell is transformed, proliferates and the progeny undergoes diversification due to quantitative changes in gene expression. As the tumour cells diversify, particular cell clones begin to dominate the cell population owing to growth advantages and host selection. At some point in time in one cell clone a qualitative change in a gene occurs that gives this clone an advantage over other clones in the population, and it proliferates and diversifies until clonal dominance again occurs. After several cycles of qualitative genetic changes, proliferation, and extensive quantitative changes in gene expression that drive diversification, a tumour cell subpopulation has progressed to a highly malignant state. (From Moustafa and Nicolson, 1998, *Oncology Research*, **9**, 505–525.)

abnormally, inappropriately or at improper concentrations, resulting in the circumvention of the normal cellular controls that regulate cellular proliferation and the state of differentiation (Klein and Klein, 1986; Nicolson, 1987). Although qualitative changes in oncogenes have been found in transformed cells, a more common finding is a change in oncogene expression due to chromosome translocations, gene amplifications and other changes, particularly in breast cancer (Yokota *et al*., 1986). Such single events by themselves are unlikely to be the underlying cause of neoplastic transformation, because further cellular changes are usually necessary.

Specific changes in oncogenes are often important in tumour progression. For example, oncogene amplification is an important mechanism (Gitelman *et al*., 1987). Although amplification of oncogenes has frequently been seen in various cancers, it is not universally seen in all tumours (Yokota *et al*., 1986). Oncogene amplification may be indicative of other, unrecognized genetic changes, or the amplification of oncogenes and other genes could contribute more directly to cancer progression (Nicolson, 1991; Moustafa and Nicolson, 1998). Since the expression of oncogenes can differ between primary breast cancers and their metastases, oncogene expression has been proposed to be important in breast tumour progression. However, examination of a variety of primary and secondary tumours reveals that oncogenes can be overexpressed, underexpressed or equally expressed in metastases compared with primary tumours (Yokota *et al*., 1986; Nicolson, 1987). Hence, the qualitative changes seen in oncogenes or the quantitative changes in their expression may contribute to progression, but they are unlikely to be the universal determinants.

Although oncogenes are important in tumour progression, the data are not convincing in support of a universal causative role for oncogenes in the progression of breast and other tumours to the metastatic state. In most studies metastases were compared with advanced primary tumours that may have already undergone all of the changes necessary to become metastatic (Yokota *et al*., 1986; Nicolson, 1987). Experimentally, the insertion of dominantly acting oncogenes into a suitable recipient cell can result in acquisition of the metastatic phenotype (Kerbel *et al*., 1987; Nicolson *et al*., 1990; Tuck *et al*., 1990). Often such experiments have been performed using aneuploid, unstable, easily spontaneously transformable animal cells as recipients, like unstable mouse fibroblast cell lines. In some untransformed cells conversion to the metastatic state only occurred when two different dominantly acting oncogenes were simultaneously inserted, an event rarely seen in spontaneous tumours. These rapid qualitative changes are unlike the slow, sequential changes that characterize spontaneous transformation and tumour progression to the metastatic state *in vivo*.

Some normal cells or even benign tumour cells are highly resistant to oncogene-mediated conversion to the metastatic phenotype (Kerbel *et al*., 1987; Nicolson *et al*., 1990; Tuck *et al*., 1990). Even within the same cell type there appears to be heterogeneity in the ability of dominantly acting activated oncogenes to cause metastatic conversion. For example, in rat mammary cells there was a considerable difference in the ability of an activated ras oncogene to convert benign cells to the metastatic phenotype (Muschel and Liotta, 1988; Nicolson *et al*., 1990). Moreover, in some systems the gene transfer techniques themselves may be as important as the transferred gene in causing metastatic conversion (Kerbel *et al*., 1987; Nicolson *et al*., 1990). (See also chapter on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*.) Often multiple gene copies are inserted, and the effects of their accompanying strong promoter/enhancer elements are not considered. It is usually assumed that oncogene constructs are randomly incorporated into the genome, but just the opposite is likely to be the case. In addition, non-random cytogenetic changes may occur concomitant with gene transfer (Muschel and McKenna, 1989). Hence it is difficult to conclude from experimental studies that the insertion of an oncogene is the only event required for conversion of benign cells to the invasive, metastatic state. In some cases oncogene-mediated conversion of a cell to the invasive, metastatic state may be dependent on the resulting concentration of an oncogene-encoded product (Muschel and Liotta, 1988; Greenberg *et al*., 1989); however, there was no obvious correlation between the expression level of an oncogene-encoded product and metastatic conversion in mammary cells (Nicolson *et al*., 1990). In addition to oncogene insertion and expression, other changes are probably necessary. Some of these changes may involve other oncogenes, suppressor genes, chromosomal structural alterations, and eventually cellular diversification.

Suppressor genes can affect mammary tumour progression and malignancy (Muschel and McKenna, 1989; Sobel, 1990). Metastasis suppressor genes were originally identified on the basis of cell fusion experiments that resulted in suppression of metastatic properties (Sidebottom and Clark, 1983). By examining differences in gene expression in nonmetastatic and metastatic cell lines, several candidate metastasis suppressor genes were identified and eventually isolated (Sobel, 1990). The best known of these in breast cancer is the *nm23* gene family. Steeg *et al*. (1988) cloned the *nm23* candidate metastasis suppressor gene, and low expression of this gene was associated with lymph node metastasis of breast cancers (Bevilacqua *et al*., 1989). The encoded product of the *nm23* gene is known. The predicted sequence of the nm23 protein was found to be identical with the *Drosophila* developmental gene *awd* product (Rosengard *et al*., 1989). The *awd* gene, in turn, was found to have a high degree of homology with nucleotide diphosphate kinase (Kimura *et al*., 1990), suggesting a possible role for the *nm23* gene product in microtubule assembly/disassembly, signal transduction and/or regulation of G proteins. An altered nm23 protein or changes in its expression could result in

modified signal transduction, gene expression and possibly progression in breast cancer.

HOST SURVEILLANCE MECHANISMS AND TUMOUR PROGRESSION

Other differentially expressed genes have been found in highly metastatic cells. For the most part, the precise roles of most of these differentially expressed genes or their gene products in tumour progression have yet to be determined. Some differentially expressed genes could be involved in tumour progression by acting on intracellular signalling in a manner that results in an increase in cell survival or an inhibition of apoptosis or programmed cell death. Gene products that act on host surveillance of tumour cells could also affect tumour progression. As tumour cells proliferate and diversify, they are under host surveillance pressures that continuously seek to inhibit or kill aberrant cells. Components that allow malignant cells to escape host surveillance mechanisms could be important in tumour progression by allowing more malignant cells to become dominant in a tumour cell population. For example, the cells that escape macrophage surveillance mechanisms might be expected to be more malignant by virtue of their ability to grow in the face of host antitumour responses. These less host-sensitive cells could then overgrow the more host-sensitive tumour cells and become the dominant cell population in the tumour. In breast cancer models the most malignant cell populations are also apparently those which are less susceptible to host surveillance systems, suggesting that *in vivo* selection occurs during the progression of mammary tumours to more malignant states (North and Nicolson, 1985). Moreover, in contrast to what has been seen in primary tumours, as metastases grow there are fewer host effector cells that infiltrate into the secondary tumour mass (Bugelski *et al*., 1987). This suggests that metastases may appear to be more resistant to host defences by virtue of the fact that fewer host effector cells can apparently penetrate into metastases.

CANCER PROGRESSION AND BREAST CANCER GENE EXPRESSION

Several differentially expressed genes associated with mammary tumour metastasis have been identified. Highly metastatic cells appear to over or underexpress a number of genes, including the overexpressed *mts1* gene (Ebralidze *et al*., 1989). The *mts1* gene has a high homology with calcium-binding proteins but is of unknown function. This gene has been found to be part of the *S100* gene family that encodes calcium-binding, cytoskeletal-binding proteins. Transfection of the *S100A4* gene into benign mammary cells resulted in these cells acquiring the metastatic phenotype (Davies *et al*., 1993). The *mts1* gene may be related to signal transduction systems, and its overexpression may lead to increased signalling of cell movement and growth.

Other differentially expressed genes have been found in mammary adenocarcinoma cells (Pencil *et al*., 1993). The most differentially expressed genes could, for the most part, be identified as known genes that had already been identified in the metastatic process, e.g. some encoded degradative enzymes, such as type IV collagenases, that are known to be differentially expressed in highly metastatic cells. Other genes encoded transcription or translation protein factors that could be involved in the metastatic process by increasing the ability of more malignant cells to proliferate under limiting growth conditions (Pencil *et al*., 1993). One of the overexpressed novel genes, *mta1*, was identified as overexpressed both in spontaneous rat mammary tumour metastases and in human breast cancer cells of high metastatic potential (Toh *et al*., 1994). This gene appears to function in signal transduction mechanisms but its exact role in maintenance of the malignant phenotype is not known. We have now identified and cloned the human *MTA1* gene, also a novel gene that appears to be involved in human breast epithelial cell motility and growth regulation. We sought to investigate the role of the *MTA1* gene by blocking its expression using an antisense oligonucleotide to the precise sequence at the transcription start site in the gene. In preliminary experiments, the antisense oligonucleotide but not a sense oligonucleotide blocked expression of the gene and inhibited human breast cancer cell motility and proliferation. The *mta1/MTA1* gene has as a part of its structure a Src homology or SH3-binding domain at its C-terminal region. This domain may function in Src and other related signal transduction pathways. The *mta1/MTA1* gene may act to modulate signal transduction pathways, and its overexpression in highly metastatic cells could result in an increased ability to circumvent host controls on cell growth and movement.

INTERCELLULAR COMMUNICATION AND TUMOUR CELL DIVERSIFICATION

The communication via intercellular junctions is an important cellular controlling mechanism in epithelial cells. These structures, in particular gap junctions, allow adjacent epithelial cells to be metabolically and electrically coupled, and this could be important in their normal maintenance of cellular proliferation and diversity. Using nonmalignant epithelial cells in culture we have found that cellular diversification and heterogeneity can be stimulated by oncogene transfer (Nicolson *et al*., 1992). This was demonstrated by transfecting relatively stable benign cell clones with the dominantly acting oncogene construct and observing the diversification of subclones derived from single transfected

cells. We found that stable, benign mammary epithelial cell clones that acquired an oncogene construct diversified rapidly concomitant with the cells acquiring the metastatic phenotype. In addition to increased diversity in metastatic properties, the transfected cells also showed increased diversity in the expression of a metastasis-associated cell surface mucin-like glycoprotein that may be involved in cell adhesion. In contrast, most of the cell clones that received the control gene construct remained relatively more stable and their subclones were for the most part nonmetastatic. The results suggest that rapid cellular diversification is an important property of highly malignant cells. A qualitative genetic change, such as sequence alteration of a dominantly acting oncogene, would be expected to occur at a low rate *in vivo*; however, the results demonstrated the relationships between qualitative gene changes, cellular phenotypic diversification and malignancy. This suggests that stimulation of cellular diversification could be an important step in tumour progression.

Environmental signals can also regulate cellular diversification and heterogeneity (see chapter on *Modelling Tumour–Tissue Interactions*). The dynamic regulation of electrical, ionic and metabolic coupling between epithelial cells is mediated by gap junctions. This form of cellular communication plays an important role in cell proliferation, differentiation, physiological responses and carcinogenesis (Dotto *et al*., 1989). Overexpression of oncogenes can disrupt gap junctional communication, and overexpression of tumour-suppressor genes is associated with an increase in gap junctional communication. When we examined the ability of a dominantly acting oncogene to cause diversification in the junctional communication properties of transfected benign mammary cells, we found that similar to their metastatic properties, their gap junctional communication was inhibited and more diverse among subclones obtained from single cell clones at various times during subculture (Nicolson *et al*., 1992). Intercellular junctional communication and its interference during tumour progression may be important in releasing cells from microenvironmental controls that regulate cellular diversification.

CANCER PROGRESSION AND CELLULAR DIVERSIFICATION

Highly progressed malignant cells exhibit rapid rates of diversification and phenotypic change, mainly due to quantitative differences in gene expression. This probably results in tumours that express a diversity of different immunological, biochemical, enzymological, structural and other cellular phenotypes. The cellular and extracellular mechanisms that control tumour and normal cellular stability and diversity have not been readily appreciated or extensively studied, hence it is difficult to ascribe precise molecular mechanisms for this process. Volpe (1988) proposed that certain genes are present in the genome that control cellular stability, and these may be altered during tumour progression, resulting in extensive tumour cell diversity, evolution and eventually acquisition of more malignant and metastatic cellular phenotypes. He termed these 'stability genes,' and such genes could be involved in karyokinesis and the repair, recombination and replication of DNA. These genes would be good candidates as genes controlling mechanisms that regulate cellular diversification. In addition to stability and other genes, a wide variety of epigenetic factors, such as those that control tissue and stromal organisation, could also control cellular diversification mechanisms (Nicolson, 1987; Moustafa and Nicolson, 1988; Miller and Heppner, 1990). The removal of cells from their usual microenvironments could also result in loss of tissue-specific controls. In this scheme cells are in a constant dynamic flux due to changes in their microenvironments.

There are apparently complex mechanisms that control cellular differentiation and the diversity of cells within tissues. Both normal cells and tumour cells are exposed to variations in the concentrations of various regulatory factors. Although normal cells may be more stable than tumour cells in their microenvironments, under certain conditions normal cells can undergo rapid change and diversification. In contrast, malignant cells can be made more stable in their cellular phenotypes. Thus, as malignant cells progress, they can reach states where they are no longer regulated by paracrine growth factors and inhibitors. Once this occurs, they can grow in a variety of microenvironments that would usually not be permissive for growth (Nicolson, 1993).

REFERENCES

Armstrong, P. B. (1984). *Invasion: Experimental and Clinical Implications*. 126–167 (Oxford University Press, Oxford).

Bevilacqua, G., *et al*. (1989). Association of low nm23 RNA levels in human primary infiltrating ductal breast carcinomas with lymph node involvement and other histopathological indicators of high metastatic potential. *Cancer Research*, **49**, 5185–5190.

Bissell, M. J. and Barcellos-Hoff, M. H. (1987). The influence of extracellular matrix on gene expression: is structure the message? *Journal of Cell Science*, **8**, Suppl., 327–343.

Bugelski, P. J., *et al*. (1987). The macrophage content of spontaneous metastases at different stages of their growth. *Cancer Research*, **47**, 4141–4145.

Davies, B. R., *et al*. (1993). Induction of the metastatic phenotype by transfection of a benign rat mammary epithelial cell line with the gene for p9Ka, a rat calcium-binding protein, but not with the oncogene EJ-ras-1. *Oncogene*, **8**, 999–1008.

Dotto, G. P., *et al*. (1989). Similar and synergistic inhibition of gap-junctional communication by ras transformation and tumor promoter treatment of mouse primary keratinocytes. *Oncogene*, **4**, 637–641.

Ebralidze, A. K., *et al*. (1989). Isolation and characterization of a gene specifically expressed in different metastatic cells and whose deduced gene product has a high degree of homology to a Ca^{2+}-binding protein family. *Genes and Development*, **3**, 1086–1093.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759–767.

Fidler, I. J., *et al*. (1990). Orthotopic implantation is essential for the selection, growth and metastasis of human cancer cells in nude mice. *Cancer and Metastasis Reviews*, **9**, 149–165.

Folkman, J. and Shing, Y. (1992). Angiogenesis. *Journal of Biological Chemistry*, **267**, 10931–10934.

Foulds, L. (1975). *Neoplastic Development* (Academic Press, New York).

Frei, E. (1982). Models and the clinical dilemma. In: Fidler, I. J. and White, R. J. (eds), *Design of Models for Testing Cancer Therapeutic Agents*. 248–259 (Van Nostrand, New York).

Frost, P. and Chernajovsky, Y. (1990). Transformation injury and the unicellular phenotype of malignant cells. *Cancer and Metastasis Reviews*, **9**, 93–98.

Gitelman, I., *et al*. (1987). DNA amplification and metastasis of the human melanoma cell line MeWo. *Cancer Research*, **47**, 3851–3855.

Greenberg, A. H., *et al*. (1989). Oncogenes and metastatic progression. *Invasion and Metastasis*, **9**, 360–378.

Heppner, G. H. (1989). Tumor cell societies. *Journal of the National Cancer Institute*, **81**, 648–649.

Kerbel, R. A., *et al*. (1987). Alteration of the tumorigenic and metastatic properties of neoplastic cells is associated with the process of calcium phosphate-mediated DNA transfection. *Proceedings of the National Academy of Sciences of the USA*, **84**, 1263–1267.

Kerbel, R. A., *et al*. (1988). Clonal dominance of primary tumors by metastatic cells: genetic analysis and biological implications. *Cancer Surveys*, **7**, 597–629.

Kimura, N., *et al*. (1990). Isolation and characterization of a cDNA clone encoding rat nucleoside diphosphate kinase. *Journal of Biological Chemistry*, **265**, 15744–15749.

Klein, G. and Klein, E. (1986). Conditional tumorigenicity of activated oncogenes. *Cancer Research*, **46**, 3211–3224.

Lynch, H. T., *et al*. (1998). Breast cancer genetics: heterogeneity, molecular genetics, syndrome diagnosis and genetic counseling. In: Bland, K. I. and Copeland, E. M., III (eds), *The Breast: Comprehensive Management of Benign and Malignant Diseases*. 370–394 (W. B. Saunders, Philadelphia).

Miller, F. R. and Heppner, G. H. (1990). Cellular interactions in metastasis. *Cancer and Metastasis Reviews*, **9**, 21–34.

Moustafa, A. F. and Nicolson, G. L. (1998). Breast cancer metastasis-associated genes: prognostic significance and therapeutic implications. *Oncology Research*, **9**, 505–525.

Muschel, R. and Liotta, L. A. (1988). Role of oncogenes in metastasis. *Carcinogenesis*, **9**, 705–710.

Muschel, R. J. and McKenna, W. G. (1989). Oncogenes and tumor progression. *Anticancer Research*, **9**, 1395–1406.

Nakajima, M., *et al*. (1987). Degradation of basement membrane type IV collagen and lung subendothelial matrix by rat mammary adenocarcinoma cell clones of differing metastatic potentials. *Cancer Research*, **47**, 4869–4876.

Neri, A. and Nicolson, G. L. (1981). Phenotypic drift of metastatic and cell surface properties of mammary adenocarcinoma cell clones during growth *in vitro*. *International Journal of Cancer*, **28**, 731–738.

Neri, A., *et al*. (1982). Development and biologic properties of malignant cell sublines and clones of a spontaneously metastasizing rat mammary adenocarcinoma. *Journal of the National Cancer Institute*, **68**, 507–517.

Nicolson, G. L. (1986). Oncogenes, genetic instability and the evolution of the metastatic phenotype. *Advances in Viral Oncology*, **6**, 143–167.

Nicolson, G. L. (1987). Tumor cell instability, diversification and progression to the metastatic phenotype: from oncogene to oncofetal expression. *Cancer Research*, **47**, 1473–1487.

Nicolson, G. L. (1989). Metastatic tumor cell interactions with endothelium, basement membrane and tissue. *Current Opinions in Cell Biology*, **1**, 1009–1019.

Nicolson, G. L. (1991). Gene expression and tumor progression to the metastatic phenotype. *Bioessays*, **13**, 337–342.

Nicolson, G. L. (1993). Pacarine/autocrine growth mechanisms in tumor metastasis. *Oncology Research*, **4**, 389–399.

Nicolson, G. L. and Poste, G. (1983). Tumor implantation and invasion at metastatic sites. *International Review of Experimental Pathology*, **25**, 77–181.

Nicolson, G. L., *et al*. (1990). Intercellular junctional communication, $p21^{rasEJ}$ expression, and spontaneous metastatic properties of rat mammary cells after transfection with c-H-rasEJ or neo genes. *Oncogene*, **5**, 747–753.

Nicolson, G. L., *et al*. (1992). Transfection of activated c-H-ras^{EJ}/pSV2neo or pSV2neo genes into rat mammary cells: rapid stimulation of clonal diversification in spontaneous metastatic and cell surface properties. *Oncogene*, **7**, 1127–1135.

North, S. M. and Nicolson, G. L. (1985). Effect of host immune status on the spontaneous metastasis of cloned cell lines of the 13762NF rat mammary adenocarcinoma. *British Journal of Cancer*, **52**, 747–755.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, **194**, 23–28.

Pearce, V., *et al*. (1984). Chromosome and DNA analysis of rat 13762NF mammary adenocarcinoma cell lines and clones of different metastatic potentials. *Clinical and Experimental Metastasis*, **2**, 271–286.

Pencil, S. D., *et al*. (1993). Candidate metastasis-associated genes of rat 13762NF mammary adenocarcinoma. *Breast Cancer Research and Treatment*, **25**, 165–174.

Poste, G. (1982). Experimental systems for analysis of the malignant phenotype. *Cancer and Metastasis Reviews*, **1**, 141–199.

Rosengard, A. M., *et al*. (1989). Reduced Nm23/Awd protein in tumor metastasis and aberrant *Drosophila* development. *Nature*, **342**, 177–180.

Rubin, H. (1990). The significance of biological heterogeneity. *Cancer and Metastasis Reviews*, **9**, 1–20.

Rudland, P. S., *et al*. (1998). Growth and differentiation of the normal mammary gland and its tumours. *Biochemical Society Symposia*, **63**, 1–20.

Sidebottom, E. and Clark, S. R. (1983). Cell fusion segregates progressive growth from metastasis. *British Journal of Cancer*, **47**, 399–406.

Sobel, M. E. (1990). Metastasis suppressor genes. *Journal of the National Cancer Institute*, **82**, 267–275.

Steck, P. A., *et al*. (1987a). Altered expression of glycosaminoglycans in metastatic 13762NF rat mammary adenocarcinoma. *Biochemistry*, **26**, 1020–1028.

Steck, P. A., *et al*. (1987b). Purification and partial characterization of a tumour metastasis-associated high Mr glycoprotein from rat 13762NF mammary adenocarcinoma cells. *Biochemical Journal*, **242**, 779–787.

Steeg, P. S., *et al*. (1988). Evidence for a novel gene associated with low tumour metastatic potential. *Journal of the National Cancer Institute*, **80**, 200–204.

Toh, Y., *et al*. (1994). A novel candidate metastasis-associated gene mta1 differentially expressed in highly metastatic mammary adenocarcinoma cell lines: cDNA cloning, expression and protein analyses. *Journal of Biological Chemistry*, **269**, 22958–22963.

Trosko, J. E., *et al*. (1993). Endogenous and exogenous modulation of gap junctional intercellular communication: Toxicological and pharmacological implications. *Life Sciences*, **53**, 1–19.

Thompson, E. W., *et al*. (1993). The invasive and metastatic properties of hormone-independent but hormone-responsive variants of MCF-7 human breast cancer cells. *Clinical and Experimental Metastasis*, **11**, 15–26.

Tuck, A. B., *et al*. (1990). Ras transfection and expression does not induce progression from tumorigenicity to metastatic ability in mouse LTA cells. *Clinical and Experimental Metastasis*, **8**, 417–431.

Viney, J. L. (1995). Transgenic and gene knockout mice in cancer research. *Cancer and Metastasis Reviews*, **14**, 77–90.

Volpe, J. P. G. (1988). Genetic instability of cancer. Why a metastatic tumor is unstable and a benign tumor is stable. *Cancer Genetics and Cytogenetics*, **34**, 125–134.

Yokota, J., *et al*. (1986). Alterations of myc, myb and rasHa protooncogenes in cancers are frequent and show clinical correlation. *Science*, **231**, 261–265.

FURTHER READING

Clarke, R. (1996). Animal models of breast cancer: their diversity and role in biomedical. *Breast Cancer Research and Treatment*, **39**, 1–6.

Edwards, P. A. (1999). The impact of developmental biology on cancer research: an overview. *Cancer and Metastasis Reviews*, **18**, 175–180.

Mattern, J., *et al*. (1988). Human tumor xenografts as model for drug testing. *Cancer and Metastasis Reviews*, **7**, 263–284.

Nicolson, G. L. (1981). The use of animal tumor models to study the metastatic process. In: Stroehlein, J. R. and Romsdahl, M. M. (eds), *Gastrointestinal Cancer*. 427–441 (Raven Press, New York).

Nicolson, G. L. (1994). The use of animal tumor models to study the role of the tumor microenvironment and paracrine and autocrine growth mechanisms in metastasis to specific sites. In: Fusenig, N. E. and Graf, H. (eds), *Cell Culture in Pharmaceutical Research*. 103–123 (Springer, Berlin).

Poste, G. (1982). Experimental systems for analysis of the malignant phenotype. *Cancer and Metastasis Reviews*, **1**, 141–199.

Rudland, P. S., *et al*. (1998). Growth and differentiation of the normal mammary gland and its tumours. *Biochemical Society Symposia*, **63**, 1–20.

Welch, D. R. (1997). Technical considerations for studying cancer metastasis *in vivo*. *Clinical and Experimental Metastasis*, **15**, 272–306.

Chapter 55

Basic Tissue Culture in Cancer Research

Philip Cavanaugh
Institute for Molecular Medicine, Huntington Beach, CA, USA

CONTENTS

- Biological Background
- Principles of Model Establishment
- Technical Considerations
- General Applications
- Biological Limitations
- Perspectives

BIOLOGICAL BACKGROUND

The maintenance of tissue or cell viability and function *in vitro* was first performed in the late 1800s. Initial attempts at cell culture used crude plasma extracts, chick embryos and ascites fluid as growth media. Most attempts at maintaining tissues *in vitro* were performed to use those tissues as a growth platform for viruses. The move to defined media supplemented with serum as a growth environment was not perfected until the late 1950s (Whitaker, 1972). The basic media used today were developed at around that time. The use of tissue culture to raise cells strictly for studies of the cells themselves then accelerated. It was only natural then for those interested in cancer research to take advantage of the system so that an unending supply of those cells could be grown and harvested to investigate multiple tumour cell properties.

PRINCIPLES OF MODEL ESTABLISHMENT

This chapter is actually devoted mostly to cell culture, which is a subset of tissue culture. For the purposes here, cell culture includes all studies were animal cells are maintained outside the donor in a viable state for long periods of time, encompassing many generations of cell life (Murrell, 1979). Tissue culture is so routinely used that it is not really thought of as a model, but an established dyed-in-the-wool technique. Cell culture is a valuable tool for cancer researchers as it enables investigators to tap a renewable source of stable tumour cells for various experiments. Tissue culture poses a number of drawbacks such as the inevitable *in vitro* alteration of certain cellular characteristics. However, it is an irreplaceable mechanism for the repeated acquisition of standardized cells for study, especially when one is concerned with exploration of human cancer cell properties. An orientation to the use of tissue culture is undertaken by examining the properties of the major tools required, the solutions required, the techniques needed to merge these two and finally special considerations needed to be taken into account when preparing cultured cancer cells for certain studies.

TECHNICAL CONSIDERATIONS

Major Equipment

The major equipment listed are items which one would want to consider obtaining if constructing a new tissue culture laboratory, or to find and acquaint one self with if one is suddenly found to be positioned in an existing facility.

Biological Safety Cabinet

The most vital piece of equipment in a tissue culture laboratory is the biological safety cabinet (BSC). These cabinets are constructed to provide an environment which is continuously flushed with sterile, particulate-free air; thus they provide the sterile environment where the bulk of tissue culture manipulations occur. These cabinets are equipped with high-efficiency particulate air (HEPA) filters which filter out any material $>0.3\ \mu m$ from an air stream. The type of cabinet usually used in tissue culture applications is a type II class A BSC, where the user is protected from biological agents arising from the material being manipulated and that material is protected from similar agents arising from outside the cabinet (Richmond and McKinney, 1995). Thus, air entering the work area and all air exhausting the cabinet pass through a HEPA filter. A diagram of a type II class A BSC is shown in **Figure 1a**. With these cabinets, workers are not protected from HEPA-passable airborne hazardous chemicals being used inside the cabinet. A type II class A cabinet can be ducted

to the outside to help prevent this, but work requiring the use of such hazardous materials is usually performed in a class II type B1 BSC, where all air drawn into the rear grille of the cabinet is vented to the outside through a hard duct (Richmond and McKinney, 1995) (**Figure 1b**). Type B cabinets also possess a greater air flow and direct a greater percentage of cabinet air to the rear. Users must aquaint themselves with the requirements and hazards of protocols and choose the proper cabinet. Other types of BSCs exist which provide additional protections, but these

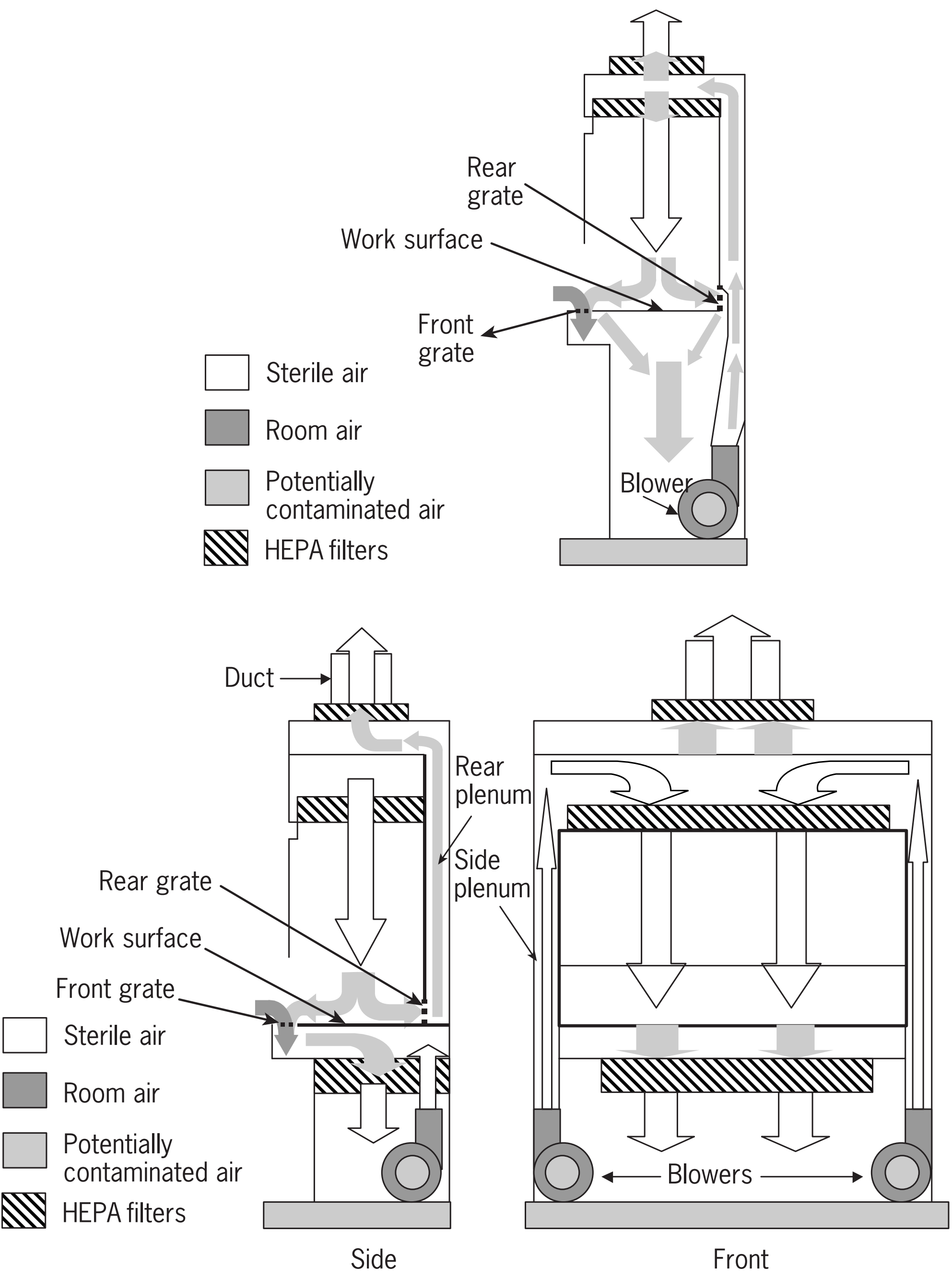


Figure 1 (a) Schematic of a class II, type A biological safety cabinet (BSC). Nonsterile air is drawn through the front and rear grates to the bottom via the blower motor. This air is directed up through the rear plenum where it is forced through one of two HEPA filters and either exits as sterile air through the top or returns into the interior through the bottom filter. (b) Schematic of a class II, type B1 biological safety cabinet. Nonsterile air is drawn through the front and rear grates. Air drawn in through the rear grate is pulled up through the rear plenum, through an HEPA filter and to the outside through a hard duct. Air drawn through the front grate passes through a HEPA filter, is blown up side plenums to the top, where it is passed through another HEPA filter and down on to the work surface.

are rarely found in ordinary tissue culture laboratories. Horizontal and vertical laminar flow clean benches also exist; however, these do not filter exhaust air and afford no user protection from the culture material; these are not BSCs and their use should be avoided in tissue culture operations (Richmond and McKinney, 1995).

Biological safety cabinets direct a flow of sterile air vertically straight downward from the inner top of the cabinet directly onto the working surface (Richmond and McKinney, 1995; Freshney, 2000) (**Figure 1**). At approximately 15 cm above the work surface, air flow is split and diverted horizontally to the front and rear grilles. The split point is typically half way from the rear of the cabinet (Richmond and McKinney, 1995). Since the flow of sterile air in a BSC is from the top down, it is imperative that any non-sterile items be kept away from the top of sterile items which are open, i.e., open bottles of media or open tissue cultures can only be approached from overhead by sterile items. Nonsterile entities such as hands and pipetters should always be positioned to the side of open flasks and bottles (**Figure 2**). Also, work should be performed a respectable distance (at least 10 cm) from the front of the cabinet so as to avoid chance room air contamination. Any object can divert the flow of sterile air, so the working surface of the cabinet should be occupied only with items that are necessary for a particular operation. Care should be taken not to place large objects in front of the rear grille, especially not to the rear of where one is working. The use of a flame in a BSC should be avoided. Air currents formed by the flame will disrupt the flow of sterile air and create unpredictable sterile air directions (Richmond and McKinney, 1995). Likewise, strong air movement outside the cabinet can interfere with uniform sterile air flow, so things such as fans should not be operated in the area of a BSC.

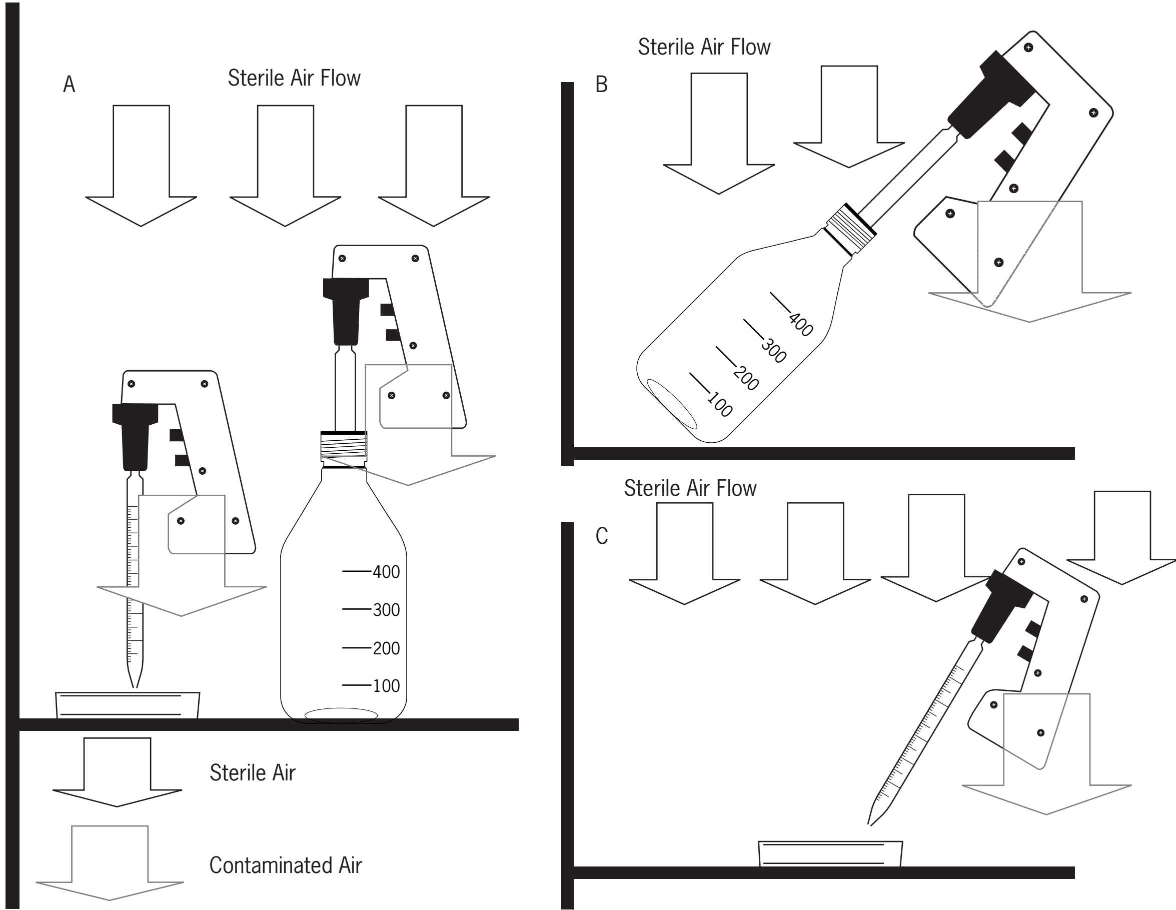


Figure 2 An illustration of a common error made when pipetting into/out of sterile containers or dishes in a tissue culture BSC. In **A**, many first time users will position a non-sterile pipettor (with associated non-sterile user hands) directly over open bottles of sterile media or tissue culture dishes, etc. The straight down vertical flow of sterile air will cause an array of microorganisms to be blown off these non-sterile items directly down into the sterile vessel. All approaches to sterile containers must be made at an angle (**B** and **C**), so that the vertical flow of contaminated air is off to the side of sterile open containers. A rule of thumb is that no sterile bottle can be pipetted into or out of when it is standing upright on its own. If you are not supporting said bottle with your hand while pipetting into/out of it, the approach to the inside will likely cause contaminants to fly in.

These cabinets usually feature side ports for vacuum and gas supply. Vacuum source in a BSC is highly recommended. Cabinet vacuum is required for many tasks, such as aspirating spent media from cultures and for filter sterilizing. A filter flask placed outside the cabinet and connected between the laboratory vacuum source and the hood's vacuum port is used to collect aspirated liquids. This flask usually contains bleach or some other strong disinfectant. An additional empty filter flask can be plumbed between the first flask and the vacuum source, to act as an emergency overflow reservoir. To protect fully the wanderers outside the cabinet, an in-line HEPA filter can be installed in the vacuum line between the vacuum supply and the overflow flask (Richmond and McKinney, 1995) (**Figure 3**).

Biological safety cabinets should be kept clean and all interior stainless-steel surfaces wiped down with 70% ethanol (Richmond and McKinney, 1995) and/or an additional disinfectant (e.g. 0.05–0.1% (w/v) cetylpyridinium chloride) on a regular basis. BSC maintenance and practice vary with the types of living things being handled. Performing routine tissue culture in a BSC and handling hazardous pathogens in a BSC are two separate issues. With the latter, all surfaces prior to the HEPA filters that are impinged upon with air that has buffeted off pathogenic material can be considered hazardous and should not be approached until the entire unit has been decontaminated with formaldehyde. Hence cleaning up under the grilles or spelunking underneath the work area is not a good idea. If one trusts the work materials to be safe, then the areas underneath the front and rear intake grates could be cleaned as this will diminish the flow of large particulates on to the filter. Since airborne particulates impinge upon the HEPA filters and do not really come off, and the filters eventually clog up. Therefore, all hoods should be monitored routinely by qualified technicians to verify adequate air flow and delivery of sterile air (Freshney, 2000).

Autoclave

A second vital piece of equipment is an autoclave (Whitaker, 1972; Paul, 1975; Freshney, 2000). Autoclaves are pressurized heat chambers which are able to kill all known living organisms via a combination of heat and pressure. Autoclaves utilize steam to achieve both of these ends, achieving internal pressures of 15–20 psi and temperatures of 110–140 °C. Autoclaves vary in their design and can be fed by an in-house steam system or can generate

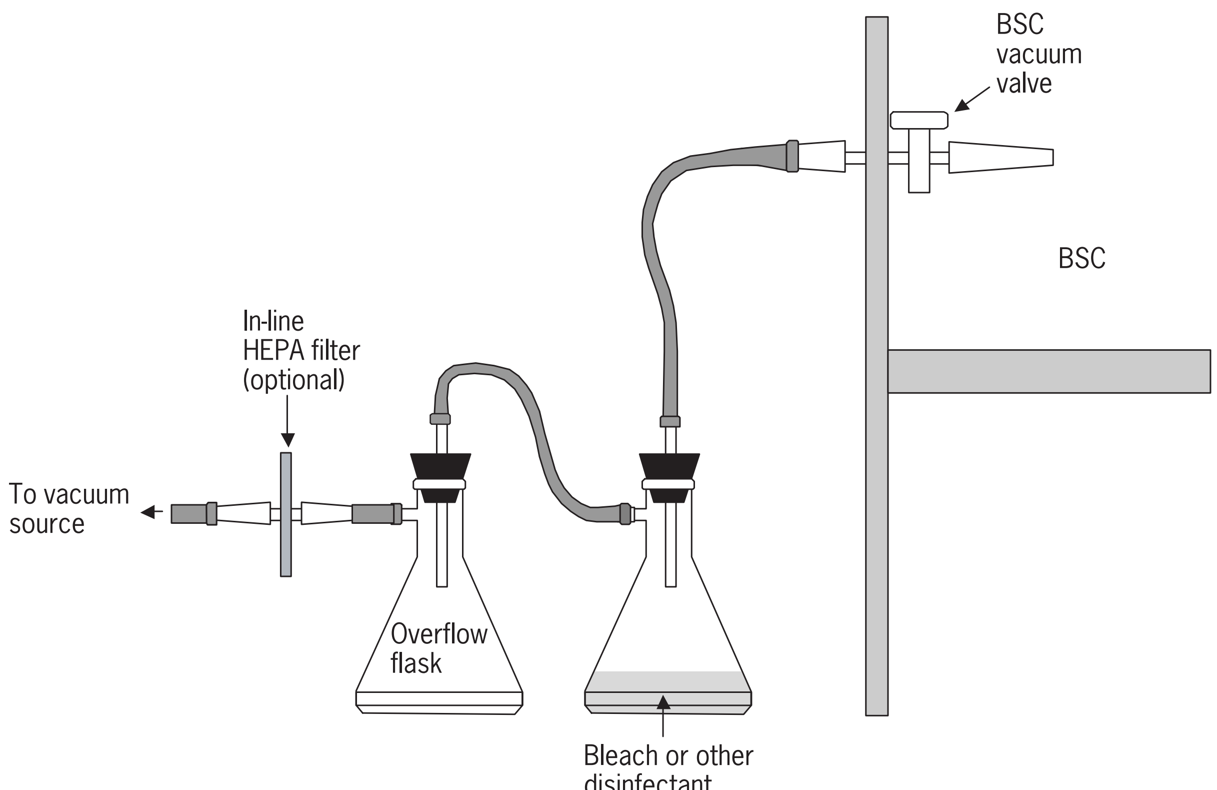


Figure 3 Schematic of a typical vacuum flask arrangement in connection with a BSC. The first flask contains some type of solution which can kill most everything. It serves as the primary reservoir of aspirated liquids. The second flask serves as an emergency overflow reservoir to protect the vacuum pump in graduate student-rich situations where flask number 1 is forgotten about. The in-line HEPA filter is to prevent pathogenic microorganisms from being distributed throughout the room via the vacuum pump or throughout the building via the building vacuum system.

steam from their own internal water supply. As such, manufacturers' instructions on operations specific for a given brand autoclave should be adhered to. The steam used must always originate from deionized water. Most autoclaves feature preset cycles for various items to be sterilized. Knowledge of the materials being autoclaved and the ability to push the proper button are all that is required to kill everything. A basic cycle consists of a rapid pressure build-up and a preset high-pressure incubation time, followed by a rapid depressurization. Typical additions to basic sterilization cycles include a drying step for wrapped or boxed materials, and a slow exhaust cycle for the sterilization of liquids. The latter allows for depressurization to occur slowly enough so that liquid samples do not boil and overflow their container. If a new autoclave is being considered, then examine models which pump sterile air through the chamber during the drying step. This shortens this cycle's time and ensures more uniform drying.

Paper, nylon and polypropylene wrappings and autoclave bags can be purchased which are specifically made for wrapping or enclosing items to be autoclaved. Items can also be autoclaved in stainless-steel, polypropylene or aluminium boxes, a number of which are particularly made for this purpose. Autoclave tape can be obtained which is provided with indicator stripes which turn from white to black when sterilization conditions and proper autoclaving have been achieved, hence this tape makes for a convenient sealer for the aforementioned wrapped or boxed items. The strategy is to wrap or enclose articles to be sterilized so that the steam in the autoclave has free access to all surfaces (Whitaker, 1972), but when removed, the wrappings or containers will be such that airborne contaminants will not be able to enter the sterilized inner sanctum. Nonsteam permeable autoclave containers cannot be totally sealed when put in the autoclave, since contents in the interior will not be exposed to the high pressure. Steam in the autoclave must be free to enter all vessels, so no container placed in the autoclave should be air-tight. Hence all rigid autoclave boxes or tray–cover combinations feature loose-fitting lids. Paper allows the passage of steam in the autoclave, but when dry, acts as a good microbial barrier, so items wrapped in paper can be totally sealed. Glass, stainless steel, aluminium foil, paper, rubber, polypropylene, nylon and polysulfone are all autoclavable. Other common laboratory materials such as polystyrene and polyethylene are not. Nalgene maintains a list of resistance properties (including autoclave resistance) of various plastics at http://nalgenelab.nalgenunc.com/tech.html. Any non-standard items to be autoclaved should be examined for the presence of materials which will undergo significant morphological changes upon undergoing an intra-autoclave experience.

Bottles act as their own autoclave container and should therefore always be autoclaved with the caps on loose, especially when containing liquids. Tight caps mean a sealed system and bottles will not only not be sterilized, but might implode. Some practitioners place a small amount of water in empty bottles or like vessels to be autoclaved, as this ensures adequate steam build up in the vessel interior (Whitaker, 1972). To ensure postautoclave sterility, one can place just-autoclaved bottles in a biological safety cabinet, wait for them to cool and then tighten the caps.

It seems that no matter how slow the depressurization, some liquids can be lost from their containers during autoclaving, hence these should always be placed in a tray when being autoclaved. The initial volume of liquids to be autoclaved should be marked on their bottle so that any loss can be noticed. Certain liquids such as microbiological media tend to foam up in the autoclave; these require more head room, i.e., their container should be twice the volume of the liquid. Many autoclave manufacturers recommend that liquid to be sterilized should occupy no more than two-thirds of the volume of the container.

Incubator

A third major item required for tissue culture is an incubator. Tissue culture incubators possess a number of features for successful culture. All should have a CO_2 feed, monitoring, and adjustment system. Ideally, incubators heat the interior through a water jacket, a water reservoir that surrounds the culture chamber. Water in the jacket is heated to 37 °C and acts as a heat reservoir and temperature stabilizer for the interior. The interior of any incubator should be all stainless steel with as many rounded surfaces as possible, for easy maintenance. Maintenance of a humidified atmosphere in the incubator is performed by the simple inclusion of a tray of deionized water at the bottom of the incubator chamber. Water should not be placed directly on the bottom of the incubator's chamber.

Inverted Microscope

A microscope is essential for monitoring of cell cultures and the performance of certain tasks such as cloning and manual cell counting. An inverted microscope is required to focus in on the bottom of cell culture dishes or flasks in such a way so that the physical same-space intersection of objective and culture vessel is avoided (Freshney, 2000).

Frozen Cell Storage Facilities

Liquid nitrogen is the traditional cooling agent for establishing an environment for frozen cells. A number of liquid nitrogen containers specifically designed for cell storage are commercially available. Any of those with narrow-neck tops are far more efficient at conserving liquid nitrogen than those with flush-diameter tops. Electric freezers which can achieve an internal temperature of −135 to −150 °C are available. These are suitable for cell storage and many laboratories will utilize these in place of liquid nitrogen freezers.

Warm Room

A room or large walk-in chamber kept near 37 °C is necessary for large-scale tissue culture operations. Large-volume roller bottle or spinner tissue cultures for cell or conditioned media harvesting usually cannot be kept in incubators. These facilities do not require humidification or CO_2 addition, as large-scale cultures are usually self-contained (Freshney, 2000).

Cell Counter

Cell counting can always be accomplished on a microscope via the use of a haemocytometer slide. However, this process is tedious and cell enumeration of a large number of samples in this way is awkward. Automated cell counters such as those made by Coulter (Luton, UK) (Paul, 1975; Freshney, 2000) aid in the accurate counting of cell suspensions, provided that the cells are relatively well dispersed. Manual inspection of cells to be counted is always recommended as automated counters do not discern between living and dead cells.

Minor Equipment

A 37 °C water bath is useful for prewarming media, thawing cells, etc. Another instrument convenient to the immediate tissue culture area is a low-speed centrifuge capable of around 3000 *g* and of holding 15- and 50-mL centrifuge tubes. This is used for spinning down cells after various washings, dissociations, etc.

Some type of small-volume hand-operated pipettors are required. Many laboratories use pipettors operating in only two volume ranges: those capable of delivering 20–200 μL and those capable of delivering 200–1000 μL. On occasion, a unit handling 0.5–10 μL may be necessary. An eight- or 12-channel multichannel pipettor will be required if culture in 96-well plates is to be performed. Also required is a pump-operated pipetting device for aspirating and delivering liquid from conventional larger-volume glass or disposable plastic pipettes.

Sterile equipment kept on hand for tissue culture laboratories is fairly similar and includes large and small pipette tips; 1 (or 2), 5 and 10-mL pipettes (with cotton plugs); Pasteur pipettes (with and without cotton plugs); various sized bottles; cell freezing vials; 15- and 50-mL centrifuge tubes; 0.22-μm filters with holders, or disposable 0.22-μm filter units; 0.22-μm syringe filters; various sized syringes; centrifuge micro filter units; etc. All of these can be purchased as presterilized diposables. However, one can also make use of reusable items that are sterilized in house. This offers certain advantages in long-term cost, ease of use and ability to monitor sterility and resterilize when necessary. All pipette tips for common pipettors are made of polypropylene and can be packaged into covered polypropylene racks and sterilized in any autoclave. The most practical glass pipettes for tissue culture use are the short variety. Tissue culture varieties of these accept cotton plugs and are usually sterilized in stainless-steel canisters. Likewise, glass Pasteur (transfer) pipettes can be placed into and sterilized in steel canisters. These can also be plugged with cotton, if needed. Bottles commonly used for tissue culture are the 45-mm wide-mouth variety manufactured by a number of companies. These offer a secure, well-sealed cap, and a drip-proof wide mouth for ease of filling, dispensing and washing. The most affordable reusable filter sterilizing units available are NALGENE® Brand units from Nalge Nunc International (Rochester, NY) and are manufactured from polysulfone. These are available as stand-alone or bottle-top units (**Figure 4**). They are typically fitted with cellulose acetate or nylon 0.22-μm filters and can be repeatedly autoclaved. Disposable filters cost $4–5 per unit, so the $60 reusable Nalgene filter pays for itself after 15–20 rounds of use.

Culture dishes, flasks and plates in virtually all laboratories are disposable polystyrene types. Polystyrene is not autoclavable so the use of these is a one-shot deal. Polystyrene is naturally hydrophobic and untreated vessels are only suitable for the growth of nonadherent cells. Polystyrene is treated under vacuum in the presence of an oxygen-rich plasma and an electrical discharge to create charged hydroxyl, carbonyl and carboxyl functional groups on the plastic (Vitesse-Dadey, 1999). These dishes are now said to be ‘tissue culture treated’ and will permit the adhesion and spreading of adherence-dependent cells. Many different types of treated polystyrene tissue culture flasks and dishes are available. A nomenclature which has developed around flasks is to precede their surface cell growing area in square centimetres with a T, so a 75-cm^2 growth area flask is referred to as a T-75.

The use of Petri-type tissue-culture treated dishes is more awkward than that of screw-top flasks; however, they are more economical, exchange gases more easily and are utilized by many laboratories successfully. Greiner (Bel-Air, MD, USA) sell these dishes and certain flasks and multi-well tissue culture plates at prices significantly less than other suppliers.

This concludes the equipment section and has mentioned most of the basic tools needed for tissue culture. Particular laboratories will possess a number of other items not mentioned here, depending on their own field and requirement for ‘special needs’. We now proceed to the liquid part of tissue culture laboratory requirements.

Solutions

Water

High-purity water is essential for the final rinsing of glassware and the making of tissue culture solutions.

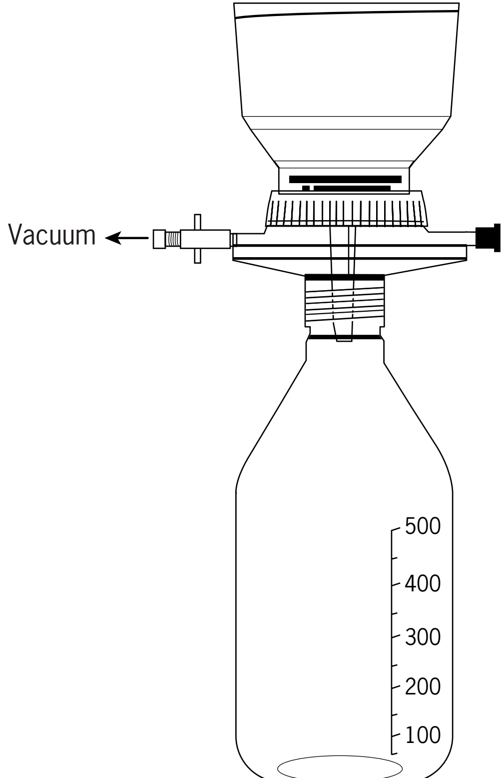


Figure 4 Schematic of the Nalgene 47-mm reusable polysulfone bottle-top filter. The model shown is fitted on to a commonly used 45-mm neck diameter tissue culture bottle. Solutions to be filtered are placed in the upper reservoir and are filtered by vacuum directly into the bottle. These filters are supplied with gas-tight lids (not shown) with fittings for tubing so that continuous filtration of large volumes can be performed. These are usually equipped by the user with nylon 0.22-μm filters and are autoclaved in paper autoclave bags. With the bottle-top design, both sides of the filter are exposed to equal amounts of pressure in the autoclave, with the result being that filter tears are uncommon. Filtration directly into sterile bottles eases handling and reduces the chances of contamination.

Most laboratories have access to water purified by deionization or reverse osmosis. This is adequate for glassware rinsing. Typically, this water is then further purified using commercially available systems which use additional ion-exchange and charcoal filtration (Freshney, 2000). More rarely, distillation is used. Names for high-purity water vary but usually reflect the name of the unit used to do the job. So names such as Nano-pure, Milli-pure, Milli-Q or doubly distilled (DD) water are added to the vocabulary of various laboratories. An article by Brush (1998) is a practical Internet-based reference which summarizes various water-purification systems.

Liquid Sterilization

For the most part, all solutions used in tissue culture are cold filter sterilized. Filter sterilization is achieved by passing the liquid through a sterile membrane which possesses pores of 0.22 μm or smaller. The vessel to hold the sterile material is first sterilized (autoclaved) as is the filter and filter holder assembly. Many facilities use disposable preassembled sterile filter units for this. Disposable 250-mL filter units cost $4–$6 each. Permanent 250-mL autoclavable polysulfone filter holders can be acquired for $60 and the 47-mm diameter 0.22-μm membranes to fit them are around $0.60 each. In the long run, significant savings can be made in purchasing permanent units, if one is not afraid of autoclaving.

A high-purity water system is essential in preparing tissue culture solutions as water from these systems is not only free of contaminants, but also has often been passed through a filter during its production. If a large volume of a particular solution is to be filtered, then an advantageous preliminary (but not necessary) step is first to pass the prepared solution through a glass-fibre and/or 1.0-μm filter on the bench top, prior to entering the hood for final sterile filtering. Certain disposable filter units are equipped with a large-pore prefilter which does the same thing. Sterilization of liquids in an autoclave is performed for certain tissue culture applications. Liquids require some head room when being autoclaved: the container should be 1.2–1.3 times the volume of the solution being sterilized. Still, some water can be lost when liquids are autoclaved, and therefore the use of a graduated container helps, with lost water replaced up to the original mark with sterile water. Water itself may be autoclaved to act as a sterile diluent for sterile 10× concentrated stock solutions such as 10× PBS or 10× liquid media, etc. Certain types of media are autoclavable, but the use of these seems to be rather uncommon.

Salt Solutions

Reproduction of media and salt solution ingredient tables with a listing of all components is not done here. To find such tables and view common media ingredients, do not go to the library. The best references for media component listing are the catalogues of companies (e.g. Invitrogen (Life Technologies Gibco), Carlsbad, CA; Sigma Chemical, St. Louis, MO, USA) which sell them. The Sigma cell culture catalogue is a convenient reference which offers a number of informative tables and methods. The Gibco catalogue contains a helpful reference section which is available on-line at http://www.lifetech.com (Tech Online).

Balanced or isotonic salt solutions are used as a base for media formulation and as a base for other reagents such as trypsin–EDTA solution, or used on their own for various purposes such as cell washing (Kuchler, 1977; Freshney, 2000).

Dulbecco's phosphate-buffered saline (PBS, DPBS or D-PBS) is the simplest formula. It is buffered with sodium and potassium phosphate and can be used in a plain air environment. It contains sodium, potassium, calcium and magnesium as anions and chloride and phosphate as cations. Glucose- and pyruvate-containing versions are available. With Earle's balanced salt solution, the phosphate concentration is lowered 10-fold and bicarbonate is added as the buffering agent. Sulfate is added to the cation list. This is used primarily as a media base or for incubator use since it contains 26 mmol L^{-1} sodium bicarbonate as a buffering agent and requires a 5% CO_2 environment for pH stabilization. Earle's solution always contains glucose as an energy source. Hank's balanced salt solution (HBSS) is similar to Earle's solution but contains less bicarbonate, resulting in an air-utilizable solution. Many laboratories supplement Hank's balanced salt solution with 20–25 mmol L^{-1} HEPES (*N*-2-hydroxyethylpiperazine-*N′*-2-ethanesulfonic acid; final pH 7.4–7.5), to provide additional buffering capacity. All of these salt solutions can be obtained calcium and magnesium free. The calcium-, magnesium-free version of DPBS is sometimes referred to as Dulbecco's phosphate buffered saline A (PBSA). These calcium-, magnesium-free versions are ideal as a base for reagents which require a lack of divalent cations such as trypsin–EDTA solution (for cell dissociation).

Media

Common Media

It seems that the simplest medium is Basal Medium Eagle (BME). Eagle further modified BME, including adding larger amounts of amino acids, to make Minimum Essential Medium Eagle, also known as Minimum Essential Medium (MEM). Dulbecco modified this by adding iron, fourfold more amino acids and vitamins to produce Dulbecco modified MEM (DME). Some DME formulations feature higher glucose concentrations also.

Tissue culture media are complex mixtures of various compounds found to be essential by various investigators for various cell types to survive and proliferate. A number of different types of media exist with various sets of ingredients and uses. Most suppliers classify tissue culture media ingredients into four divisions: salts, amino acids, vitamins and 'other' components. Kuchler (1977) divides media types into four groups. In order from basic to exotic, the ingredient list rises from 28 in group I to 50–60 in group IV. A major difference in the jump from group I to II is the inclusion of nonessential amino acids (Kuchler, 1977). Whitaker (1972) divides media components into nine categories: inorganic salts, amino acids, vitamins, glucose, nonglucose carbohydrate sources, coenzymes, reducing agents, nucleic acid derivatives and lipid sources. The Whitaker system gives one a better understanding of media components and allows for obvious discrimination of the differences between simple and complex media. By and large, the addition of more exotic items was performed in an effort to develop media that would support cell growth in the absence of serum. Those categories and example contents are as follows (Whitaker, 1972).

Inorganic Salts

The basic 'core' salts employed in media are those found in the aforementioned balanced salt solutions. Additional salts are added for various purposes. A major difference in additional salt composition among media is the addition of iron. Many media (e.g. BME, MEM) lack an iron salt whereas others (e.g. DME, Ham F12, Medium 199) are iron supplemented. Other minor metal salts such as zinc, selenium and copper are found in certain media (Ham F12) and others yet are supplemented with rarer salts such as molybdenum and vanadium (MCDB 131).

Amino Acids

Certain media contain only the essential amino acids (BME, MEM, DME) whereas others contain all amino acids (Ham F12, IMDM, NCTC 109). Amino acid concentrations can vary and tend to increase from lower levels in basic media to much higher levels in specialized media such as Waymouth and NCTC 109 (Whitaker, 1972). Close attention should be paid to glutamine concentrations in media as it has a half-life of about 3 weeks at 4 °C and 1 week at 37 °C. Many researchers obtain media glutamine free and add it from a 200 mmol L^{-1} 100× concentrate just prior to use.

Vitamins

A core group of vitamins is present in all media, typical of that found in BME. Additional vitamins such as *p*-aminobenzoic acid and vitamins A, D, E and K are present in others, such as NCTC 109.

Glucose

Glucose is found in nearly all media. Most media contain a minimum of 1000 mg L^{-1}. This rises to 3000 mg L^{-1} in media such as McCoy 5A, to 4500 mg L^{-1} in IMDM and certain formulations of DME and to 5000 mg L^{-1} in Waymouth medium 752/1. Exceptions to this exist, e.g. L-15 medium contains no glucose and substitutes it with galactose.

Carbohydrate Sources Other than Glucose

These are found in specialized media and include compounds such as sodium pyruvate, galactose, ribose and glucoronolactone.

Coenzymes

This category includes compounds that can act as enzyme cofactors and includes things components such as ATP, FAD and coenzyme A, etc., all found in NCTC 109.

Reducing Agents

Ascorbic acid is placed here, although in many tables it is listed as a vitamin. Other agents include glutathione and L-cysteine.

Nucleic Acid Derivatives

This does not include ribose but includes compounds such as hypoxanthine, xanthine, adenine, guanine, cytosine, uracil, etc. Medium 199 is an example of a medium with additional nucleic acid derivatives.

Lipid Sources

These are just plain fats, including cholesterol (medium 199), lipoic acid (Ham F-12) and Tween 80 (medium 199, NCTC 109).

The choice of medium for a certain cell type should be indicated by the supplier and adhered to. Changes in medium for any reason should be accompanied by rigorous examination for any changes in cell behaviour. On occasion, we have had need to switch cells to an iron-deficient medium and this dictated a change in basic medium type, which was kept constant for tests and controls to eliminate media-created changes.

Tissue culture media is usually buffered using a sodium bicarbonate/gaseous CO_2 combination. The buffering system is based on physiological ones and is only effective when an excess of gaseous CO_2 is present. In a $HA \leftrightarrow H^+ + A^-$ acid/base system, at a pH where A^-/HA ratio is high, the presence of an exogenous limitless amount of HA will result in an efficient buffering system, as the excess HA maintains H^+ and A^- at a given ratio, however, the maximum concentration of HA in solution must be limited. A gaseous source of HA can do this: act as a limitless supply, but still exist in a way where the maximum concentration existing in the aqueous phase can be restricted. Most media contain sodium bicarbonate or direct its addition upon preparation. Sodium bicarbonate readily dissociates in aqueous solution forming Na^+ and HCO_{3^-} (carbonate). Carbonate will complex with H^+ to form carbonic acid: H_2CO_3, an acid with a pK_a of 6.1. At a pH of 7.4, carbonate will function as a weak base, as this pH is distant from the pKa, and only about 5% of the added HCO_{3^-} will exist as H_2CO_3. Gaseous CO_2, readily goes into aqueous solution and reacts with water to form H_2CO_3, the concentration of which can be adjusted by adjusting the partial pressure of CO_2 gas in the surrounding atmosphere. At a pH of 7.4, if the concentration of carbonate is 25 meq/L, then the concentration of H_2CO_3 is 1.25 meq/L. If the partial pressure of gaseous CO_2 is set so that H_2CO_3 is maintained at 1.25 meq/L maximum, then the system will buffer at pH 7.4, since H_2CO_3 lost in any way will be replenished by H_2CO_3 formed from gaseous CO_2. Thus, the excess gaseous CO_2 maintains H_2CO_3 and HCO_{3^-} at a given ratio (White *et al.*, 1978). This results in a buffering system which imitates physiological buffering systems to great extent. This system also neutralizes the propensity of CO_2 generated by cellular metabolism to form carbonic acid and lower media pH. To maintain this system, most tissue culture incubators are CO_2 incubators, with a gaseous CO_2 feed line delivering CO_2 from a pressurized tank to the interior, an internal CO_2 monitor, and an automatic gas feed system which will continuously maintain the internal CO_2 monitor, and an automatic gas feed system which will continuously maintain the internal CO_2 at 5–10% (or whatever is set). Confluent cultures will produce other acids such as lactic acid and will utilize carbonate for carboxylation reactions. Therefore, the buffering system will eventually fail in crowded cultures unless media is frequently changed.

Additional buffering capacity is provided in many tissue culture situations by the addition of HEPES to the media. This is a non-toxic zwitterionic buffering agent with a pK_a of 7.3, hence it buffers well at the usual tissue culture pH of 7.4. These media still contain carbonate, for buffering and for synthesis processes. As mentioned, most media contain or direct the addition of sodium bicarbonate as a buffering agent in atmospheres supplemented with 5% CO_2. However, certain media are formulated to buffer adequately in air only and contain no bicarbonate. An example is L-15 (Leibovitz 15) medium. Cells growing in this medium would require a CO_2-free incubator. Other media, such as DMEM, may direct the addition of sodium bicarbonate, which is in keeping with an atmosphere supplemented with 10% CO_2. Care should be taken to maintain awareness with regard to the type of medium and atmosphere required for a particular cell type.

Serum

A medium alone will not usually support the growth of most mammalian cells. The medium is supplemented with serum at 5–20% (v/v), in order to support cell growth. In many cases, excess serum can be toxic, and it is rarely used at >20% (v/v). Serum provides a multitude of growth factors, nutrient transport factors, hormones, lipids and other components required to maintain cells. The most common serum used to supplement media in tissue culture is fetal bovine serum (FBS). This is available from many suppliers and is usually tested for tissue culture use and for the presence of many viral and other microbial agents. FBS from most suppliers is supplied sterile and is thoroughly tested for tissue culture use. The only precautionary note to keep in mind is to have on hand enough serum from the same supplier of the same lot number to last through an entire experiment. Other sera commonly used are calf serum and horse serum. These are also usually supplied sterile and again the only concern is to maintain serum consistency throughout an experiment.

When defined growing conditions are required, or for harvesting of secreted cell products, many investigators utilize serum-free media. Serum-free media are available

in two basic different types of preparations. The first are those containing protein supplements, examples being AIM-V, CHO-A-SFM and endothelial-SFM media (Invitrogen/(Life Technologies/Gibco)). These have a restricted list of cells of which they are advertised to support the growth of. Proteins commonly found in these types of media are albumin, transferrin and insulin. The second type of premade serum-free media are the protein-free variety, examples being PFHM-II and CHO III PFM (Invitrogen/(Life Technologies/Gibco)) and serum-free and protein-free hybridoma medium (Sigma). These are even less universally applicable than the protein-containing type. Serum substitutes are available which can be used to custom augment one's own base medium to achieve serum-free cell growth. The simplest of these is ITS, a supplement containing insulin, transferrin and selenium. More complex mixtures such as controlled-process serum replacements (CPSR, Sigma) are available for custom formulations using one's own medium. The Invitrogen/(Life Technologies/Gibco) catalogue reference guide contains sections on serum-free media and applications and on the adaptation of cells to serum-free conditions.

Antibiotics/Antimycotics

Many researchers avoid the use of these as much as possible. In these cases, their use is avoided in routine culture with established and tested cell lines. Other laboratories place antibiotics in media as a matter of course. In this case, the antibiotics of choice are penicillin and streptomycin. Their concentrations as preventative bacterial inhibitory agents are 100 000 units L^{-1} and 100 mg L^{-1}, respectively. A combination of these two is available commercially as a 100× stock solution. The antifungal/antiyeast agent most commonly used is amphotericin B. It is also available as a 100× stock solution and the final concentration used in media is 2.5 mg L^{-1}. All three of the above are also available as a combination antibiotic/antimycotic solution which is supplied as a 100× preparation. Penicillin and streptomycin are readily soluble in solutions such as PBS or HBSS and in-house sterile filtered preparations of these can easily be produced. Amphotericin B is difficult to solubilize and requires a detergent such as sodium deoxycholate to achieve this. For this reason, amphotericin B is best purchased ready to use as the detergent complex, which is known as fungizone. Other antibiotics are also useful in tissue culture. Gentamicin and kanamycin can be used and both exhibit a broader range of activity than do penicillin and streptomycin, including activity against mycoplasma. The Sigma tissue culture catalogue contains a usefeul table on the actions, activity and recommended concentrations for many tissue culture antibiotics. Antibiotics/antimycotics are usually used when primary cultures are established and are often placed in media after extensive out-of-hood manipulation of cells has occurred, such as with FACS sorting. Antibiotics can sometimes be used to resolve contaminated cultures, in which case the concentrations used can be significantly higher than those listed above, e.g. the cytotoxic concentration of streptomycin is >20 000 mg L^{-1} (Paul, 1975). We have found kanamycin to be particularly nontoxic and have used concentrations as high as 5 mg mL^{-1} to eliminate contamination. The Invitrogen/(Life Technologies/Gibco) catalogue reference guide contains a section on the decontamination of cultures using antibiotics. Elimination of fungal or yeast contamination with amphotericin B is more problematic as this agent is useful as a preventative but is not nearly as effective as a cure. Novel methods to decontaminate yeast-ridden cultures have been described with one (Behrens and Paronetto, 1984) using added macrophages to clear up this type of contamination.

Cell Removal Solutions

Cells are usually removed from tissue culture plates or flasks using 0.25% trypsin and 0.53 mmol L^{-1} EDTA, made in a basic salt solution. Trypsin–EDTA solution can be purchased ready to use or as a 10× solution containing 2.5% trypsin and 5.3 mmol L^{-1} EDTA. The latter is diluted 1:10 in sterile Ca^{2+}-, Mg^{2+}-free PBS or Ca^{2+}-, Mg^{2+}-free HBSS to produce a working solution. Either trypsin or EDTA alone can be used to remove certain cells from surfaces. We have found that Ca^{2+}-, Mg^{2+}-free HBSS containing 0.5–1.0 mmol L^{-1} EDTA can remove many cell types from culture plates, and with certain cells, exposure to 4 °C enhances the effect. Other enzymes such as collagenase or dipase are used on occasion. For simple subculturing and passaging of cells, trypsin–EDTA is used in most cases. The Invitrogen/(Life Technologies/Gibco) catalogue reference guide contains a section on various methods and reagents for the dissociation of cells from culture vessels.

Cell Freezing Solutions

Freezing solutions contain a cryoprotectant which lowers the solution's freezing point and protects cell membranes from freeze-related injury. Cells suspended in these are cooled slowly, and when the freezing point is reached, ice crystals initially form outside the cells. This causes the osmolarity of the extracellular solution to rise and water to diffuse from the cells, resulting in stabilization of the cells in a supercooled dehydrated state. Anyhow, the intracellular crystallization of water is prevented. Glycerol and dimethyl sulfoxide (DMSO) are the usually used cryoprotectants for freezing-point depression. The most prevalent formulation uses DMSO. These solutions are usually made up of the cell's usual growth medium with the addition of the cryopreservative. Thus, a typical freezing solution is 70–80% (v/v) medium, 10–20% (v/v) FBS and 10% (v/v) DMSO. Serum- and/or DMSO-free

solutions are available for special needs such as the cryopreservation of cells grown in serum-free media. Freezing solutions are available premade and presterilized from many suppliers.

Cells

Various cell lines are usually obtained from colleagues or collaborators. Certain companies, such as Clonetics (Walkerton, MD, USA), PromoCell (Heidelberg, Germany) and Cytotech (Copenhagen, Denmark) sell a number of cell lines, mostly normal. The most plentiful source of low-cost (although their costs have risen dramatically from 1995 to 2000) normal and tumour cell lines from various species is the American Type Culture Collection (ATCC) (Rockville, MD, USA) (www.atcc.org). Cells from this source are rigorously checked for contamination and are verified to be of the species stated. Also, the history and growth conditions of the lines are well documented. Any cell lines obtained from a private source should be checked for mycoplasma contamination and, if possible, tested for species verification and the presence of viral agents. Some of this can be accomplished by sending out samples to for-fee providers of these services, e.g. the ATCC offers a mycoplasma testing service. Cross-cellular contamination of cultures has been found frequently, with HeLa cell contamination of numerous cell lines having been reported (Stulberg *et al*., 1976; Ogura *et al*., 1997). Specific instances of other types of cross-species contamination have been noted (Muir and Gunz, 1977; Price *et al*., 1998). If not detected prior to the start of an experiment, this will play havoc with intended studies. Methods for monitoring such contamination are cumbersome, but worthwhile, with fluorescence *in situ* hybridization being a current suggested protocol (Multani and Pathak, 1999), with the indication that the earlier recommendation of the use traditional cytogenetic methods (Pathak and Hsu, 1985) may not be foolproof. Certain other multiparameter comparisons can be performed (Gignac *et al*., 1993), depending on the cell type and available assay techniques. Other preliminary methods can be devised depending on the species being used and the expertise available. For instance, we have found that most antihuman transferrin receptor antibodies will not react with rat (Cavanaugh *et al*., 1999) or mouse transferrin receptors. Simple treatment of rat, mouse or other nonhuman cultured cells with a fluorescent-tagged antibody of this nature followed by fluorescent microscopy will provide evidence of human cell contamination.

Mycoplasma Contamination

Mycoplasma are cell wall-deficient intracellular-dwelling parasitic bacteria which can infect a cultured cell line without causing any apparent morphological changes. Mycoplasma consist of only a plasma membrane, ribosomes and a genome of approximately 580 kb. Mycoplasma can exist during their life cycle as single cells smaller than 0.2 μm in diameter (Raab, 1999). A number of detection methods for mycoplasma are available, with the most recent being polymerase chain reaction-based detection of mycoplasma DNA in culture media (Tang *et al*., 2000). A number of commercially available services exist to which one can send samples for paid mycoplasma testing (Raab, 1999). Mycoplasma contamination can be controlled with antibiotics; however, the proper antibiotic must be chosen (Drexler *et al*., 1994; Freshney, 2000). The Sigma tissue culture catalogue antibiotic table indicates which of these are useful against mycoplasma. Kanamycin and gentamycin are often used. Cell cultures should be tested for mycoplasma on a routine basis, and any new cultures arriving in a particular laboratory should be tested. Trust no-one.

Techniques

Basic Culture

Initiating Cultures from Donated or Purchased Starting Cultures or Frozen Cell Vials

Cell Culture as a Start. Volumes are based on starting with and using a 10-cm diameter dish. In the BSC, aspirate the medium from the source cell culture, using a sterile transfer pipette. Rinse the cells once with 5–10 mL of Ca^{2+}-, Mg^{2+}-free HBSS and aspirate. Add 5–10 mL of sterile trypsin–EDTA and incubate in a tissue culture incubator until the cells detach. Pipette the cells into a sterile centrifuge tube. Centrifuge at 1000 *g* for 5 min, aspirate out the spent trypsin–EDTA solution and suspend the cell pellet in 2–5 mL of sterile growth medium (media plus serum). Add an aliquot of the suspension to waiting sterile fresh tissue culture dishes or flasks containing sterile fresh growth medium (e.g. 10 mL of medium per 10-cm dish). Place the dishes in a humidified 37 °C, 5% CO_2 incubator.

Frozen Cells as a Start. Thaw the cell vial rapidly in a 37 °C water bath. In a BSC, using a sterile pipette, dilute the thawed cell suspension in at least 10 volumes of prewarmed sterile growth medium and add to a fresh sterile culture dish or flask. To reduce the concentration of DMSO, the vial can be centrifuged once thawed, the freezing solution aspirated off and the cell pellet resuspended in growth medium; however, this is not usually necessary. Place the dish in a humidified 37 °C, 5% CO_2 incubator.

Subculturing

Attached Cells

Subculturing is the act of removing cells from a given dish or flask and replating them into additional vessels. This is

synonymous with passaging. This is done to maintain the cells in a viable state and/or to expand the population into additional dishes or flasks for experimental purposes. Each time a cell line is subcultured, the cells are said to have been passaged. Thus, the passage number indicated on a cell culture or vial represents the number of times that they have gone through trypsinization and reapplication to another dish or flask. The method is essentially the same as described above for initiating cultures from cultures.

Suspension Cultures

Many hybridoma cell lines and also leukaemia and lymphoma lines will grow unattached to a dish or other substrate. These cells will not require tissue culture-treated plasticware and can be grown in un-treated Petri dishes. For subculturing, trypsin is not required and an aliquot of cells from an old overgrown culture is pipetted sterilely into a sterile new dish replete with sterile fresh growth medium. Since subculturing is merely a dilution of cells into a new dish, the concept of passaging is hazy here and many investigators do not regard these cells as being passaged in the conventional sense. On occasion, suspension culture cells will be encountered with no passage number.

Cell Counting and Viability

Counting relatively monodispersed cells with an automated cell counter such as a Coulter Counter is the most convenient and accurate method for cell density determination. For this, we usually pipette 100 μL of a cell suspension into 10 mL of Coulter isoton, and count 100 μL of that diluted suspension in the counter. So, the result provided by the instrument is cells per 100 μL of a 1:101 diluted preparation; this value is multiplied by 1010 to provide cells per millilitre in the original suspension. Most cell counters can be programmed to correct for dilution and will present data in terms of cells per millilitre directly. One should make sure that the correction factor fits in with your current state of dilution.

A Coulter Counter has to be set up for particular cell counting operations. For most tumour cells, a 100-μm aperture tube is used and a lower threshold which eliminates counts due to anything with a diameter of less than 4–5 μm is set. A counting instrument such as a Coulter Counter detects particle numbers and sizes. It cannot discriminate between viable and nonviable cells and cannot discriminate like-sized debris from cells. For examination of a cell suspension for viability, debris and degree of cell aggregation, one must visually examine the cells directly under a microscope. Examination, counting and viability are all done at once by adding one volume of 0.4% trypan blue (in PBS) into one volume of a dilute cell suspension. An aliquot of the suspension is placed into a cover-slipped haemacytometer slide and examined under a microscope. Nonviable cells will take up the dye and appear blue. A haemacytometer counting area is divided into nine major squares which are 1 mm per side (**Figure 5**). With an official haemacytometer cover-slip in place, the depth of the counting area is 0.1 mm. All of the cells counted in one of the squares represents cells per 0.1 mm^3. Most individuals count the live and dead cells in four corner squares to acquire cells per 0.4 mm^3. This is multiplied by 2500 to produce cells per millilitre in the dye-treated suspension. The divisions inside the major squares are used as guides; the finer divisions in the side and centre squares are not usually needed for normal tumour cell examination and counting.

Cell Freezing

The process of cell freezing has been made much easier with the development of propan-2-ol-filled freezing chambers, the interior of which will cool from room temperature to −70 °C at a rate of 1 °C min^{-1} in a −70 °C freezer. Cells to be frozen are first removed from a culture dish using sterile trypsin–EDTA as described above, the removed cells are centrifuged and the spent trypsin–EDTA solution is aspirated off the pellet sterilely. The cell pellet is suspended in sterile freezing solution at 37 °C. Some investigators resuspend at no more than 1×10^6 cells ml^{-1}; however, many practitioners use higher densities with no apparent ill effects. The cell suspension is aliquoted into sterile screw-capped freezer vials and placed in a room-temperature freezing chamber. The chamber is placed in a −70 °C freezer overnight. The cell vials are then placed into a liquid nitrogen-cooled permanent storage facility.

Large-scale or High-density Cultures

A number of methods exist to increase cell culture surface area without using billions of dishes (Meisenholder, 1999). A common technique is to use roller bottles. This requires a roller-bottle rolling apparatus, which could be considered as major equipment. The roller is placed in a 37 °C chamber or warm room. Large cell culture equipment of this nature is often not located in a CO_2 environment, and therefore non-CO_2 buffering systems are used, or the bottle interiors are flushed with sterile 5% CO_2 prior to placement in the incubator. Roller bottles with interior pleats are available, which offer an additional increase in surface area. Roller-bottle cultures are often used when cell secreted products are collected (Wakabayashi *et al*., 1995), as the medium to cell surface culture area can be reduced significantly. Another method for establishing large surface area cultures in a small place is to use microcarrier beads (Meisenholder, 1999). A number of different beads with various surface functionalities are available. Cells are subcultured on to or 'plated' on to the beads similarly to how they are plated on to dishes in subculturing. Procedures for the different beads available vary, so the manufacturer's recommendations should be followed for rehydration, sterilization and plating techniques. These

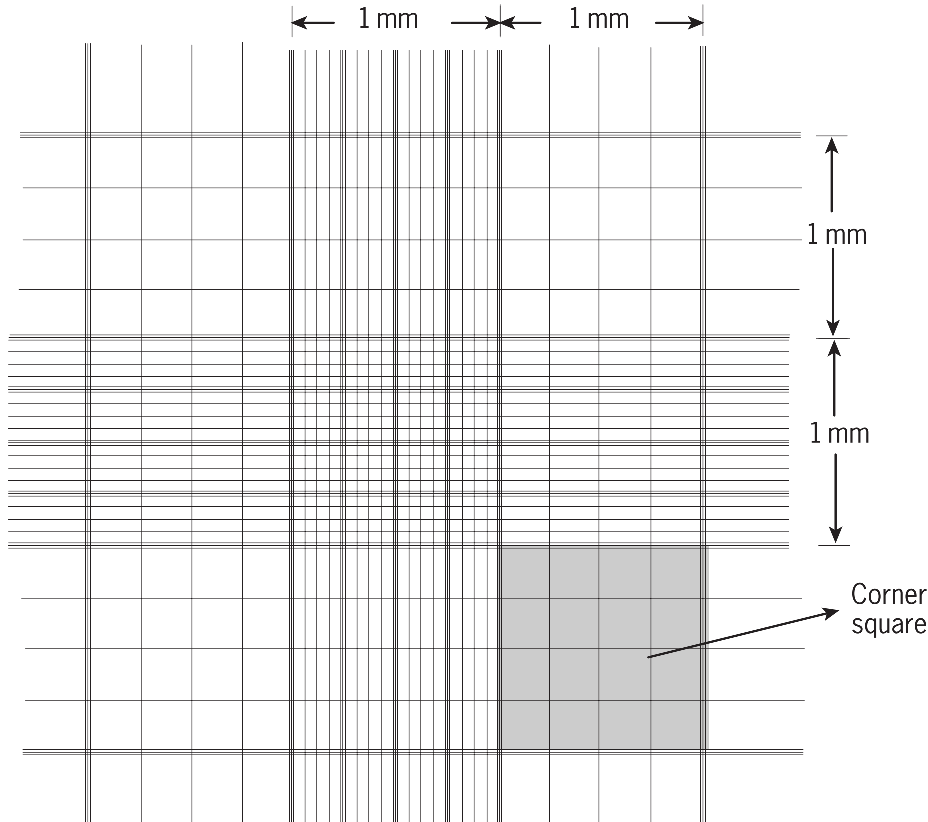


Figure 5 Schematic of a haemocytometer slide. Each major square is 1 mm per side. The depth throughout the entire counting area when the slide is equipped with a haemocytometer cover slip is 0.1 mm. Hence the total cells in one major square represents the number per 0.1 mm^3 in the suspension that was applied to the slide. Usually, corner squares are used since they are the least subdivided. The centre, side, top and bottom squares can be used to count additional 0.1-mm^3 volumes.

cultures are stirred at a slow rate using specialized vessels and stirrers. The stirrers used rotate at extremely low speeds and can often be set to rest at certain time intervals. Once again, the placement of a culture of this nature in a CO_2-supplemented environment is difficult, so these are usually buffered using HEPES or some other agent.

Primary Cultures

For this, cells are dissociated from viable tissue, usually a surgical explant, and are established as growing populations in normal tissue culture environment (Murrell, 1979).Tumour cells will frequently plate out and grow in tissue culture from surgical explants. Tumour cells often readily dissociate from a tumour mass and their transformed nature often results in successful tissue culture growth. For the establishment of tumour cell culture from animals, a section of tumour surgically removed from an anaesthetized host is suspended in the medium of choice. Mechanical manipulation of the tissue with forceps and scissors will often release enough cells to initiate a culture. If necessary, tissue-disrupting enzyme preparations such as collagenase/dispase can be used to dissociate cells. In either case, cells are initially washed 2–4 times by centrifugation and resuspension in fresh media, and then an aliquot is placed in growth medium in a fresh tissue culture dish or plate and placed in the incubator. If contamination with normal cells is seen, then the tumour cells can be further isolated by cloning. In this case, the initial disrupting media, dissociation medium and growth media all usually contain antibiotics and antimycotics.

Cloning

Cloning essentially involves establishing a culture from one cell. One can dilute a monodispersed cell suspension to exceedingly low cell densities (e.g. 10 cells mL^{-1}) and plate this on to a 96-well tissue culture plate at 100 μL per well. Individual cell colonies that arise in given wells can be assumed to have started from one or two cells. These are harvested, recultured and recloned to ensure single-cell origin status. Dilute cell suspensions can also be plated on to 10-cm tissue culture dishes and colonies arising can be harvested using cloning rings. These are ~3–5-mm diameter open-ended glass cylinders which are coated on one rim with sterile silicone grease. Medium is aspirated from

the dish and a ring is placed grease-side down over the colony of choice. Trypsin–EDTA solution is added inside and removed cells are aspirated out and recultured. A good substitute for cloning rings is sterile 3–5-mm Whatman No 1 filter-paper discs that are soaked in trypsin–EDTA. A disc of this nature is placed over the colony of choice and released cells are 'entrapped' by the paper. After the trypsin had done its work, the disc is removed and placed with its captive cells into a fresh dish with fresh growing medium. Many cells do not grow well at low density. Minimizing the increase in culture surface area and media volume when subculturing harvested colonies can help, in addition to the addition of pyruvate to the medium. For hybridomas, commercially available hybridoma cloning factors or supplements can be added to enhance low-density cell survivability.

GENERAL APPLICATIONS

Tissue culture-raised tumour cells have their use in a number of *in vitro* and *in vivo* assays. This chapter is not intended to discuss particular assays, but to warn new investigators of tissue culture manipulations that may have to be performed to make assays work.

Intact cultured cancer cells are used for *in vivo* tumorigenesis and metastasis assays, *in vitro* invasion assays, migration assays, transfection studies, drug sensitivity assays, FACs analysis, adhesion assays, etc. Lysates made from tissue-cultured cancer cells are used for enzyme assays, Western blot analysis, DNA analysis, RNA analysis, signal transduction assays, etc. Any manipulation of the culture system must be done so as not to interfere with the assay system and to ensure that all cell lines being analysed and compared are treated identically.

Usually, fellows, post-docs and graduate students are situated in a laboratory and have to work on established cancer types and their corresponding cell lines as dictated by the laboratory's director and tasks to be accomplished for certain grants. In the rare case that one is presented with a wide-open prospective research situation, and a cell culture-dependent study has not yet been committed to a certain tumour line or species, a decision as to what to use as models must be made. The first of these perhaps is to decide whether or not to work with human or animal lines. If one is not going to extend studies into the *in vivo* arena, then the more human disease state-relevant human lines would probably be the ideal choice. Raising human and raising animal tumour lines in culture are almost identical processes and no major differences in difficulties in propagating either are likely. The handling of human tumour lines does not require any special barrier devices as they are not recognized to pose a threat to researchers manipulating them. If cells are eventually to be injected into host animals and assessed for tumorigenicity, metastatic capability, drug sensitivity, etc., then another set of decisions must be made. Human lines are usually tested in this regard in nude mice, an immunocompromised host which requires special housing and the donning of protective paper pyjamas by the workers. The protective gear needed to be worn to work with nude mice does not make any kind of fashion statement, result in a physically awkward work situation. Animal lines present the advantage of a more ideal *in vivo* testing system, where a tumour line's *in vivo* behaviour can be assessed in fully immunocompetent syngeneic hosts. These latter studies do not require special housing, present a more amiable animal/tumour cell/reagent manipulation system, do not require the wearing of special suits and are much cheaper. Other considerations must be kept in mind, e.g. if one is contemplating the use of certain reagents in an upcoming experiment, then the choice of species of cell line to be used may be limited by the reagent itself. For example, if analysis of certain cell antigens via Western blotting is to be performed, then the antibody to be used must be verified to recognize the antigen in the species of cell line contemplated.

Many studies require the comparison of tumour cells to corresponding normal cells. The culture of certain normal cells can require significantly different conditions than tumour cells. Many normal cells will only go through a limited number of cell divisions prior to a final exit from known reality, whereas their transformed counterparts will keep on ticking. The acquisition and maintenance of normal cells has been made easier by the commercial availability of many of these. Companies such as Clonetics, PromoCell and Cytotech supply normal cell types although these are in many cases restricted to human varieties. Clonetics will supply detailed instructions and media specific for the various normal cells offered. The downside of these sources is that they tend to be rather expensive.

An example of a system where great differences in handling occurs is with normal mammary epithelial cells. These require low concentrations of calcium for optimal cell division (McGrath and Soule, 1984), depending on the reference (Emerman and Wilkinson, 1990). Certain immortalized normal cell types such as normal rat kidney cells will grow readily in culture and handle similarly to tumour cells. Whether or not these types of lines constitute precise normal cell behaviour is not always accepted. Many normal endothelial cell lines are available or have been isolated (Belloni *et al*., 1992), and these often appear to culture and subculture as easily as many tumour cell lines. When dealing with normal cells, three things should be kept in mind (1) unusual media requirements such as low Ca^{2+} for normal mammary epithelial cells; (2) limited number of divisions in culture–experiments may have to be designed so that any normal cell types do not exceed their probable life span; (3) certain handling characteristics such as subculturing at a high density may have to be performed.

BIOLOGICAL LIMITATIONS

Procedures and reagents listed above were developed for day-to-day tissue culture cell maintenance. However, not all routine practices can be immediateley transferred to the experimental situation. All reagents and procedures must be examined for their ability to affect one's own field of study.

Trypsin-induced Artifacts

Cell removal from plates must be done so as not to interfere with subsequent tumour cell action. Trypsinization of tumour cells can inhibit a number of processes. For instance, trypsin exposure can inhibit the ability of tumour cells to aggregate platelets (Gasic *et al*., 1977), so if an assay of this nature is planned, then EDTA removal would be a preferred alternative. Trypsin may cleave cell surface antigens, e.g. the transferrin receptor is known to have a trypsin-sensitive site (Larrick and Creswell, 1979). Analysis of removed cultured cells for cell surface transferrin receptor levels would dictate an avoidance of the use of trypsin.

Media and Serum Effects

As mentioned earlier, FBS used in an experiment should be a constant. The lot number of FBS used should remain the same for the length of the experiment and should be the same used across all cells being compared. Likewise, if two different cell types are being compared in a given assay, the effect on the assay of different sera, different sera concentrations and different media types used must be taken into account. If possible, all of these should be the same. For example, if two different cell types were being analysed for internalized transferrin content, then one might want to be wary of the medium used and consider whether one cell type was grown in a medium containing iron salts whereas the other was not.

FBS is very useful for growing cells, but might present problems when it comes to analysing those cells for a particular activity. The binding of residual serum proteins on to tumour cell surfaces might effect *in vivo* assays, e.g. remaining bovine serum proteins might instigate tumour cell rejection. FBS proteins have been known to be retained by cultured cells for prolonged periods (Johnson *et al*., 1990), and cell-bound FBS proteins have been observed to induce immune responses in cell-injected animals (Mizushima and Cohen, 1985). One might want to substitute the serum of the animal host species into cultures for 1–2 days prior to harvesting those cells for a metastasis assay. On occasion, investigators have avoided FBS altogether, to circumvent possible artifacts (Emerman *et al*., 1990). Bovine serum might influence cell behaviour in ways not necessarily reflective of that shown *in vivo*. If one is analysing human cultured cancer cells for any activity, the replacement of FBS with human serum should be considered. A move to serum-free conditions for any postculture assay (McKenzie *et al*., 1990) could preclude FBS induced artifacts. Culture of cells in same-species serum for particular assays may be the most appropriate growth conditions (Emerman *et al*., 1987).

It is difficult to monitor all media additives and assess the possible contribution of all to artifactual effects. The best practice is to keep abreast of recent developments in a particular area and to be aware of any reports of negative effects of particular media components. A good example is the discovery that phenol red, a common media ingredient used as a pH indicator and which imparts its red colour to media, was found to possess oestrogen-like qualities (Berthois *et al*., 1986; Rajendran *et al*., 1987). This was determined to produce undesired background effects in certain breast cancer studies. This provided greater incentive for the marketing of phenol-red free media by many companies, and the avoidance of the use of phenol red by those engaged in those types of experiments.

Passage Number

A major consideration in the use of tissue cultured cells in cancer research is passage number. A number of different studies have indicated that certain cultured cancer cells will show a drift in certain behaviours as passage number increases. Welch *et al*. (1984) found that the drug resistance and metastatic capability of certain rat mammary adenocarcinoma cells will change with increasing passage number. Investigators will usually freeze down large numbers of early passage cells or cells at a state known to be stable, and limit their scope of experiments to cells which have undergone minimal passaging from that state.

Monodispersed Cells

The ability to obtain monodispersed culture cells can be problematic. Isolated single cell suspensions or monodispersed cells are the best for certain assays such as tail-vein injection (experimental) metastasis assays, and for certain other techniques such as FACs analysis or sorting. Many tumour cells will grow as sheets and will be released from plates as sheets after removal with trypsin–EDTA or EDTA alone. Passage of suspensions through sterile nylon meshes of various sizes can aid in reducing multicellular aggregates. One method is to trypsinize stock cultures and vigorously to pipette up and down to disaggregate cells as much as possible. This suspension is then passed through a sterile 45-μm nylon mesh and all cells replated at approximately 25% of the starting density. The next day, these are removed from plates with EDTA only, passed through a 45-μm filter and used in assays.

Logarithmic-phase Growth

When rapidly growing cells exist in culture with plentiful nutrients and growing space, their numbers will increase exponentially. Confluency is the degree of coverage of the growing vessel that cells occupy. If cells are covering far less than the available growing area, then exponential growth can usually be assumed. Cells in this state are said to be in logarithmic phase growth (Paul, 1975). When two or more cell lines are to be compared in a given assay; e.g. for metastatic potential, both should be in logarithmic phase growth as this will dictate that the cells are highly viable and that both have approximately the same percentage of cells in each of the cell cycle phases. Many researchers do not perform complicated growth curve assays to ascertain logarithmic phase growth and ensure that a given culture is in this state. Instead, care is taken to only harvest cells for an experiment when the dish or flask on which they are growing is 40–60% confluent. The cells should therefore exhibit an identical size and DNA content range, etc. When dealing with cells growing in suspension, it is best to be familiar with the maximum density (in cells per millilitre) that a particular culture will normally reach, and to harvest these for a procedure when the cell count is no more than 60% of maximum.

Cell Cycle Synchronization

A number of agents are available which will synchronize cells at a particular stage of the cell cycle without inducing toxicity (Spadari *et al*., 1985). Synchronizing cells is one method of eliminating cell cycle size and other cell cycle variations in a population. Tumour cells have been shown to display varied behaviours dependent on cell cycle (Iwasaki *et al*., 1995). This might be considered important when performing experiments such as tail vein injection metastasis assays, as cell size will change with cell cycle along with the propensity of a certain percentage of the cells to divide and in effect increase the injected cell number. Synchronizing cells for procedures is not usually performed but is a method that can be considered to render cells more constant (Merrill, 1998).

PERSPECTIVES

Cell culture is the foundation upon which many cancer research operations rest. Voyaging back to the tissue culture facility to obtain the cells needed for a particular experiment is a routine trip in many laboratories. Learning the basics of the properties of the equipment, the solutions, the cells and procedures involved is perhaps the most valuable background information that a researcher can have, whether involved in the actual performance of the tasks or engaged in a supervisory role. This chapter was written to acquaint those entirely inexperienced with tissue culture with the terms and materials used. The pitfalls associated with the use of routine tissue culture reagents in certain experiments were pointed out in an effort to convince new researchers to monitor continuously all procedures for reagents and practices which will adversely affect their own particular experiments.

REFERENCES

Behrens, U. J. and Paronetto, F. (1984). Elimination of the yeast *Candida parapsilosis* from lymphoid cells and monolayer cells in culture. *In Vitro*, **20**, 391–395.

Belloni, P. N., *et al*. (1992). Organ-derived microvessel endothelial cells exhibit differential responsiveness to thrombin and other growth factors. *Microvascular Research*, **43**, 20–45.

Berthois, Y., *et al*. (1986). Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. *Proceedings of the National Academy of Sciences of the USA*, **83**, 2496–2500.

Brush, M. (1998). Water, water, everywhere: a profile of water purification systems. *The Scientist*, **12**, 18.

Cavanaugh, P. G., *et al*. (1999). Transferrin receptor overexpression enhances transferrin responsiveness and the metastatic growth of a rat mammary adenocarcinoma line. *Breast Cancer Research and Treatment*, **56**, 203–217.

Drexler, H. G., *et al*. (1994). Treatment of mycoplasma contamination in a large panel of cell cultures. *In Vitro Cell Development and Biology*, **30**, 344–347.

Emerman, J. T. and Wilkinson, D. A. (1990). Routine culturing of normal, dysplastic and malignant human mammary epithelial cells from small tissue samples. *In Vitro Cell Development and Biology*, **26**, 1186–1194.

Emerman, J. T., *et al*. (1987). Effects of defined medium, fetal bovine serum and human serum on growth and chemosensitivities of human breast cancer cells in primary culture: inference for *in vitro* assays. *In Vitro Cell Development and Biology*, **23**, 134–140.

Emerman, J. T., *et al*. (1990). *In vitro* sensitivity testing of human breast cancer cells to hormones and chemotherapeutic agents. *Cancer Chemotherapy and Pharmacology*, **26**, 245–249.

Freshney, R. I. (2000). *Culture of Animal Cells: A Manual of Basic Technique* (Alan R. Liss, New York).

Gasic, G. J., *et al*. (1977). Platelet aggregating material in mouse tumor cells. Removal and regeneration. *Laboratory Investigation*, **36**, 413–419.

Gignac, S. M., *et al*. (1993). Multiparameter approach in the identification of cross-contaminated leukemia cell lines. *Leukemia and Lymphoma*, **10**, 359–368.

Iwasaki, T., *et al*. (1995). Cell-cycle-dependent invasion *in vitro* by rat ascites hepatoma cells. *International Journal of Cancer*, **63**, 282–287.

Johnson, M. C., *et al*. (1990). Persistence of fetal bovine serum proteins in human keratinocytes. *Journal of Burn Care and Rehabilitation*, **11**, 504–509.

Kuchler, R. J. (1977). *Biochemical Methods in Cell Culture and Virology* (Dowden, Hutchinson and Ross, Stroudsburg, PA).

Larrick, J. W. and Cresswell, P. (1979). Transferrin receptors on human B and T lymphoblastoid cell lines. *Biochimica Biophysica Acta*, **583**, 483–490.

McGrath, C. M. and Soule, H. D. (1984). Calcium regulation of normal human mammary epithelial cell growth in culture. *In Vitro*, **20**, 652–662.

McKenzie, R. C., *et al*. (1990). Fetal bovine serum contains an inhibitor of interleukin-1. *Journal of Immunological Methods*, **133**, 99–105.

Meisenholder, G. (1999). Postmodern culture: maximizing cell culture output at every level. *The Scientist*, **13**, 21.

Merrill, G. F. (1998). Cell synchronization. *Methods in Cell Biology*, **57**, 229–249.

Mizushima, Y. and Cohen, E. P. (1985). A study on an artifact introduced by fetal bovine serum-supplemented medium. *Hokkaido Igaku Zasshi*, **60**, 321–326.

Muir, P. D. and Gunz, F. W. (1977). Contamination of human melanoma cell lines by mouse L cells. *Pathology*, **9**, 301–309.

Multani, A. S. and Pathak, S. (1999). Conventional cytogenetics alone is not sufficient for identifying interspecies cell line contamination. *Anticancer*, **19**, 1753–1764.

Murrell, L. R. (1979). Vertebrate cell culture: an overview. In: Maramorosch, K. and Hirumi, H. (eds), *Practical Tissue Culture Applications*. 9–23 (Academic Press, New York).

Ogura, H., *et al*. (1997). Detection of HeLa cell contamination-presence of human papillomavirus 18 DNA as HeLa marker in JTC-3, OG and OE cell lines. *Japanese Journal of Medical Sciences: Biology*, **50**, 161–167.

Ota, T., *et al*. (1996). Inhibition of metastasis by a dialysable factor in fetal bovine serum in B16 melanoma cells. *Cancer Letters*, **110**, 201–205.

Pathak, S. and Hsu, T. C. (1985). Cytogenetic identification of interspecies cell-line contamination: procedures for non-cytogeneticists. *Cytobios*, **43**, 101–104.

Paul, J. (1975). *Cell and Tissue Culture* (Churchill Livingstone, Edinburgh).

Price, J. E., *et al*. (1998). Distinctive karyotypes and growth patterns in nude mice reveal cross-contamination in an established human cancer cell line. *Oncology Reports*, **5**, 261–266.

Raab, L. S. (1999). Cultural revolution: mycoplasma testing kits and services. *The Scientist*, **13**, 21.

Rajendran, K. G., *et al*. (1987). Estrogenic effect of phenol red in MCF-7 cells is achieved through activation of estrogen receptor by interacting with a site distinct from the steroid binding site. *Biochemical and Biophysical Research Communications*, **142**, 724–731.

Richmond, J. Y. and McKinney, R. W. (eds) (1995). *Primary Containment for Biohazards: Selection, Installation and Use of Biological Safety Cabinets* (US Department of Health and Human Services, Washington, DC).

Spadari, S., *et al*. (1985). Control of cell division by aphidicolin without adverse effects upon resting cells. *Arzneimittelforschung*, **35**, 1108–1116.

Schmitt, K., *et al*. (1988). A safe and efficient method for elimination of cell culture mycoplasmas using ciprofloxacin. *Journal of Immunological Methods*, **109**, 17–25.

Stulberg, C. S., *et al*. (1976). Identification of cells in culture. *American Journal of Hematology*, **1**, 237–242.

Tang, J., *et al*. (2000). A polymerase chain reaction based method for detecting Mycoplasma/Acholeplasma contaminants in cell culture. *Journal of Microbiological Methods*, **39**, 121–126.

Vettese-Dadey, M. (1999). One to grow on a profile of tissue culture labware. *The Scientist*, **13**, 20.

Wakabayashi, H., *et al*. (1995). Purification and identification of mouse lung microvessel cell-derived chemoattractant for lung-metastasizing murine RAW117 large-cell lymphoma cells: identification as mouse monocyte chemotactic protein-1. *Cancer Research*, **55**, 4458–4464.

Welch, D. R., *et al*. (1984). Multiple phenotypic divergence of mammary adenocarcinoma cell clones. I. *In vitro* and *in vivo* properties. *Clinical and Experimental Metastasis*, **2**, 333–355.

Whitaker, A. M. (1972). *Tissue and Cell Culture* (Williams and Wilkins, Baltimore).

White, A., Handler, P., Smith, E. L., Hill, R. L. and Lehman, I. R. (1978). *Principles of Biochemistry* (McGraw-Hill, New York).

FURTHER READING

Freshney, R. I. (2000). *Cultue of Animal Cells: a Manual of Basic Technique* (Alan R. Liss, New York).

Chapter 56

Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes

Anthony Wynshaw-Boris
University of California, San Diego School of Medicine, La Jolla, CA, USA

CONTENTS

- Introduction
- Principles of Model Establishment and Technical Considerations: Production of Genetically Altered Mice
- General Applications: Use of Transgenic and Knockout Mice to Study the Function of Cancer-causing Genes
- Interpretation of Experimental Results
- Biological Limitations
- Perspectives

INTRODUCTION

As discussed elsewhere in this book, oncogenes were first discovered by studying oncogenic viruses and by transfection studies of tumour DNA into cell lines. Tumour suppressors were identified based on loss of heterozygosity in human tumour studies. However, the understanding of the mechanism of action of oncogenes and tumour suppressors *in vivo* has been made possible by the use of genetically altered organisms, particularly the mouse. The technology for mammalian germ-line genetic manipulation was developed in the mouse, which has become the workhorse model organism for genetic studies. Like much of modern biology, genetic manipulation of the mouse via transgenic or knockout techniques has revolutionized the study of genes that cause cancer. The development of this technology has allowed investigators to examine the biological activity of genes important for the formation of cancer.

The basic experimental design differs for the study of oncogenes and tumour-suppressor genes. For oncogenes, transgenic mice are typically produced that overexpress oncogenes in specific tissues under the guidance of tissue-specific promoters. The expectation is that mice will develop tumours in the tissue where the oncogene is overexpressed. To study tumour suppressors, however, it is necessary to inactivate the gene to study its function. Therefore, knockout mice are produced with inactivation of tumour-suppressor function. If the tumour suppressor is required for embryonic development or viability of the organism, then special conditional knockouts can be produced that allow for gene inactivation in specific tissues at specific developmental times or adulthood.

In this chapter, the methodology for genetic manipulation of the mouse are outlined, then examples of the application of these techniques for the study of oncogenes and tumour suppressors are presented.

PRINCIPLES OF MODEL ESTABLISHMENT AND TECHNICAL CONSIDERATIONS: PRODUCTION OF GENETICALLY ALTERED MICE

The ability to manipulate the mouse genetically was made possible by those who developed techniques for the culture of mouse embryos and embryonic stem cells. Scientists developed preimplantation embryo culture conditions in the 1950s and 1960s, and included Wesley Whitten at the Australian National University and the Jackson Laboratory and John Biggers and Ralph Brinster at the University of Pennsylvania. In addition, uterine and oviduct transfer techniques were developed at the same time by Anne McLaren at Cambridge University. These techniques made possible the study and development of embryos outside of the mammalian uterus. This led to the development of teratocarcinoma cell lines that could be injected into these early embryos and contribute to a variety of somatic tissues, and then to embryonic stem cells that could contribute to somatic tissues and the germ line. Embryonic stem cells were derived from blastocysts independently by Gail Martin at UCSF and Martin Evans at Cambridge University, and soon were demonstrated to contribute to the germ line after injection into blastocysts.

These embryological techniques opened the door for the targeted genetic manipulation of the mouse genome.

Genetically altered mice can be produced in one of two ways: direct pronuclear injection of cloned DNA into the male pronucleus of a fertilized mouse egg, or targeted introduction of transgenes by homologous recombination using embryonic stem (ES) cells. Cloned genes were injected into fertilized eggs in the early 1980s by Ralph Brinster of the University of Pennsylvania, Tom Wagner at Ohio University, Erwin Wagner and Beatrice Mintz of the University of Pennsylvania, Jon Gordon and Frank Ruddle at Yale Univeristy and Frank Costantini and Elizabeth Lacy from Columbia University, and were found to pass stabily through the germ line.

A number of investigators studied the requirements for homologous recombination in cultured cells during the 1980s, including Mario Capecchi at the University of Utah, Oliver Smithies at the University of Wisconsin and then University of North Carolina and Raju Kucherlapati at Princeton. These principles were applied successfully to the modification of ES cells and then mice were made with germline transmission of modified alleles by the groups of Capecchi and Smithies, as well as groups from the laboratories of Elizabeth Robertson of Columbia University, Rudolph Jaenisch of Heidelberg University and MIT and Allan Bradley at Baylor University.

Detailed protocols and brief descriptions of the history of the development of transgenic manipulation and gene targeting techniques are available in several excellent laboratory manuals (see Further Reading). These techniques have allowed for the stable transfer of normal, altered or chimaeric genes into the mouse germ line, thus providing powerful tools with which to study mechanisms underlying development, function, tumorigenesis and gene expression within a physiological context.

Production of Transgenic Mice by Pronuclear Injection

Transgenic mice are produced by microinjection of DNA directly into pronuclei of fertilized mouse eggs, and results in the random integration of foreign DNA into the mouse genome. To express a particular protein in a specific tissue, the injected DNA is built by standard molecular biological cloning techniques in the form of an expressed eukaryotic gene (**Figure 1**). Constructs for injection consist of a promoter and polyadenylation site to provide signals for initiating and terminating transcription, respectively, and it has been found with extensive experience that an intron in the construct leads to much higher levels of expression of the transgene. These artificial genes used for injection are termed ‘transgenes,’ and mice made from the insertion of transgenes are referred to as ‘transgenic mice.’ The gene of interest must be expressed in the appropriate cell type within the mouse, and at an appropriate time in development to have a desired effect. Therefore, the choice of an appropriate promoter is critical to insure appropriate patterns of expression of the transgene. It often takes several modifications of endogenous promoters to identify sequences that will give appropriate levels of tissue specific transcription of transgenes in animals. Fortunately, such modified promoters have been identified that reproducibly give restricted expression patterns in transgenic mice, and can be used to express oncogenic sequences.

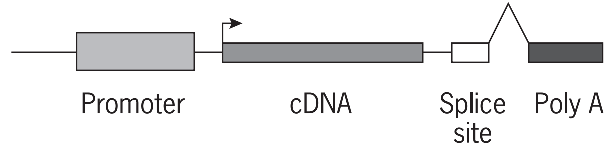


Figure 1 Structure of a transgene or transgenic construct. The transgene consists of a promoter, the cDNA to be expressed, a splice site and a polyA addition site.

After fertilization, the male and female pronuclei do not fuse immediately. The male pronucleus is large and, using micromanipulators to stabilize the egg, 100–200 copies of the purified DNA can be injected into the male pronucleus (**Figure 2**). Most of the injected eggs do not integrate the injected transgene. However, some fraction of injected eggs will integrate the injected DNA into a random chromosomal locus. The transgenic DNA usually integrates at the one cell stage and therefore foreign DNA will be present in every cell of the ‘transgenic’ animal. The manipulated embryo is implanted into the oviduct of a pseudopregnant female where the embryo will continue to develop until delivered at term. Offspring are analysed by molecular techniques such Southern blotting or polymerase chain reaction (PCR), and transgenic animals that have integrated the gene into their DNA are referred to as founders. These founders usually transmit the transgene as a heritable trait and hence a transgenic line is established. Under ideal conditions, the efficiency for generating transgenic founder mice containing 5–50 kb fragments is between 10 and 30%. In most cases, the transgene integrates at a single site randomly in the genome as 1–50 copies arranged in a head-to-tail orientation (tandem repeats). The site of integration can have profound effects on the expression of the transgene, so different founder lines may display vastly different levels of expression. In addition, enhancer elements at the site of integration can sometimes direct transgene expression to unexpected tissues in unpredictable spatial and temporal patterns of expression. Founders must therefore be screened for expression of the transgene RNA or protein, and the best founder line(s) for a particular experiment can be selected from the available founder lines. Once an appropriately expressing founder line has been identified, the expression pattern from integrated transgenes generally remains stable over many generations.

Recently, it has been possible to produce transgenic mice by pronuclear injection of large DNA fragments

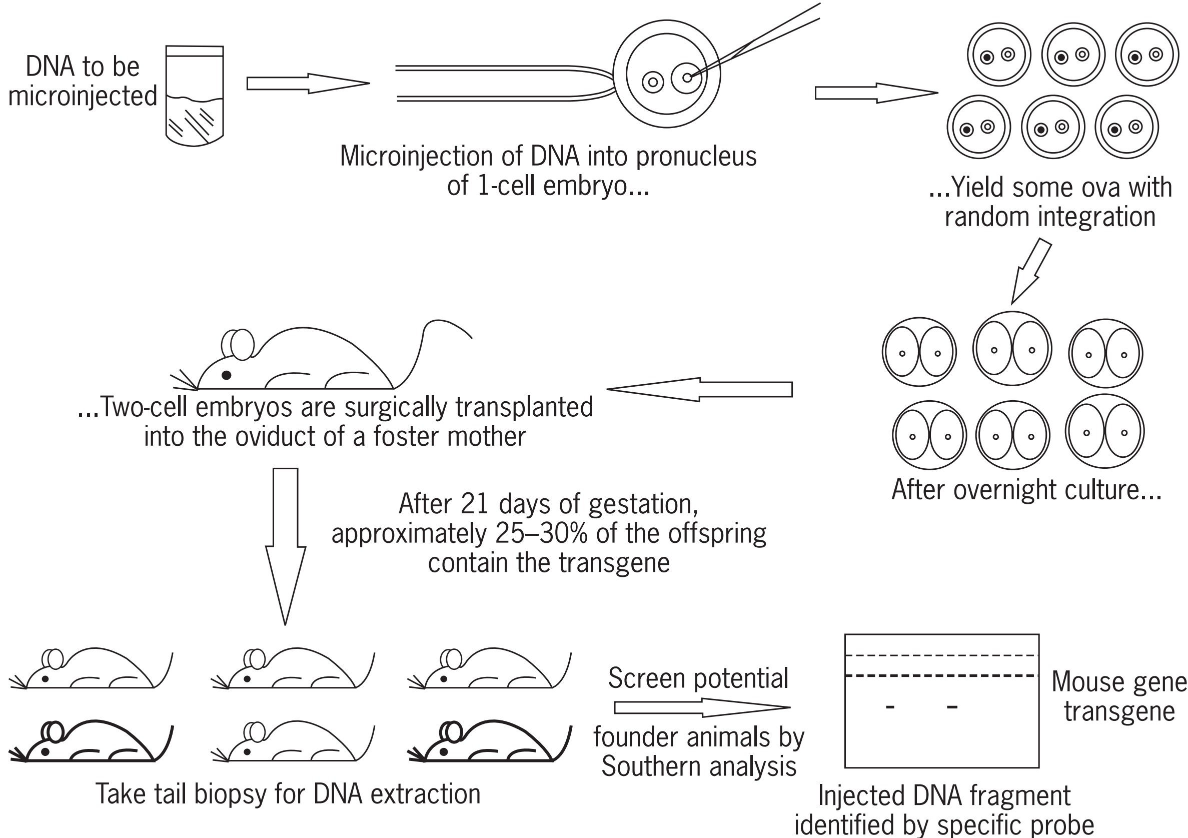


Figure 2 Generation of transgenic mice.

derived from yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs). Even large genes will be included on most YACs or BACs with important regulatory regions. Expression from these large DNA fragments appears to be more uniform and predictable in transgenic mice than expression from traditional transgenic constructs. This is probably the consequence of at least two factors: it is more likely that all regulatory sequences important for reproducible expression are contained in these large DNA fragments; and this added DNA sequence insulates the transgene from influences from the integration site. YACs can accommodate larger inserts of DNA (several hundred kilobases) compared with BACs (100–150 kb). However, YACs are more unstable during manipulation and transgenesis than BACs, and several excellent mouse BAC genomic libraries are available. Therefore, BACs have become the vector of choice for producing transgenic mice from large DNA fragments.

Gene Targeting via Homologous Recombination in ES Cells

In contrast to the random integration of DNA that occurs due to nonhomologous recombination during pronuclear injection of DNA, DNA can also integrate in a site-specific manner by homologous recombination. Foreign DNA introduced into mammalian cells generally integrates into the host genome via nonhomologous recombination, with a frequency of homologous to nonhomologous recombination of 1 in 1000. However, techniques were developed for the growth and transfection of ES cells in culture and selection for those cells that have undergone homologous recombination, while maintaining the ES cells in an undifferentiated (pluripotent) state. This technology has allowed for the extensive manipulation of the mouse genome that is occurring in many laboratories throughout the world. Meticulous culture procedures are required to maintain ES cells in culture so that they retain a normal karyotype and do not differentiate, to insure that the genetically manipulated cells can contribute to the germ line after injection into the host blastocyst.

ES cells are grown in culture and manipulated DNA of interest is transferred by electroporation (**Figure 3**). The manipulated DNA contains a marker that allows for the selection of cells that have stabily incorporated the DNA, whether it is integrated via homologous or nonhomologous recombination. Following electroporation, cells are subcloned so that DNA can be extracted. Genomic DNA from the clones is screened using standard molecular biology procedures to identify clones that

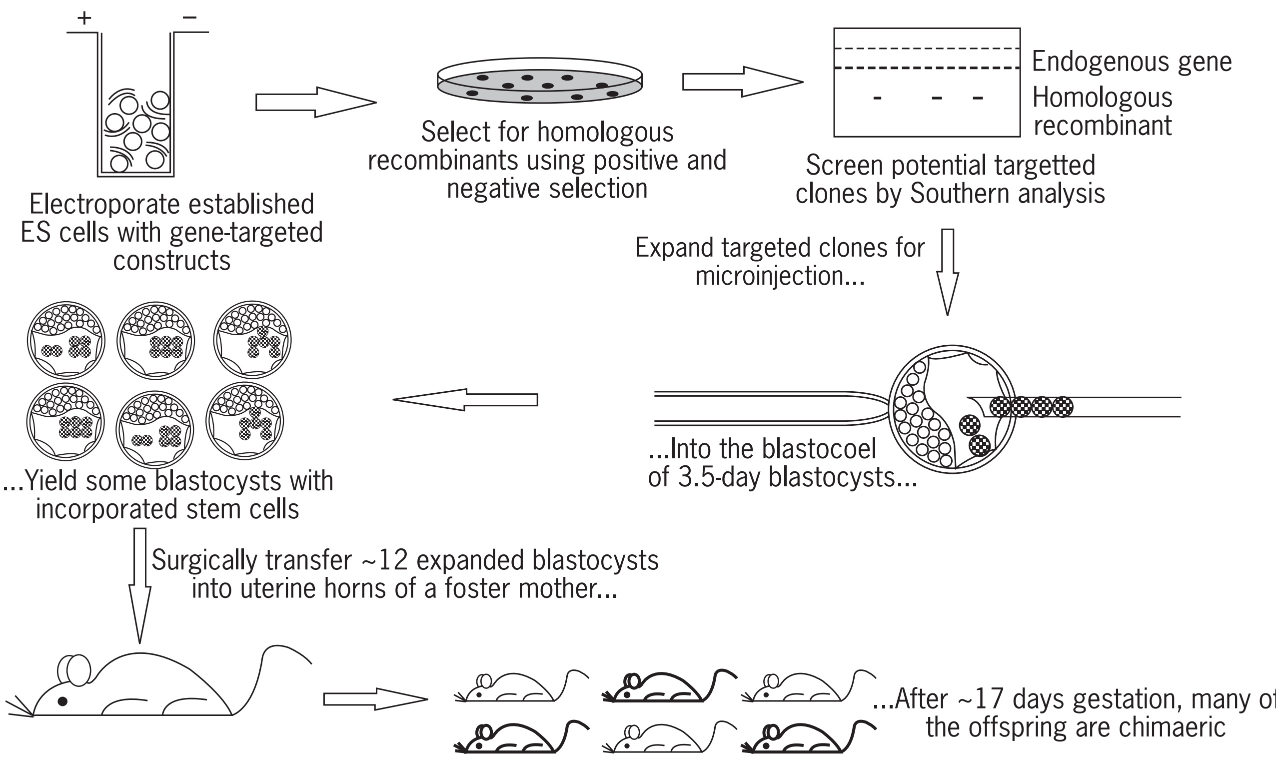


Figure 3 Construction of transgenic mice using embryonic stem cells.

have undergone homologous recombination at one allele. Once ES cell clones have been identified they are injected into blastocysts (host blastocysts) obtained by flushing the uterus of a 3.5-day pregnant female. The injected blastocysts are then transplanted into the uterus of a pseudopregnant female where the embryo continues to develop. Offspring are screened for the presence of the altered allele. Offspring are referred to as chimaeras, as some of the animal is derived from the host blastocyst and some from the injected ES cells that contain the altered allele or transgene. If the manipulated ES cells contribute to the germ line, then the chimaera can be bred and some of its offspring will be heterozygous for the altered allele. This is referred to as germ-line transmission. Germ-line transmission is essential for establishing a knockout line.

Several factors have been defined which facilitate the growth of ES cells, enhance the frequency of homologous recombination and simplify screening procedures. For example, use of isogenic DNA (DNA from the identical strain of mice as the ES cells), use of linearized DNA and the use of two selectable markers improve the ratio of homologous to nonhomologous clones surviving selection to the range of 1 in 15–300. In general, rates of homologous to nonhomologous recombination using optimized systems are 1 in 25 to 1 in 100.

Using homologous recombination in ES cells, it is possible to make a variety of modifications to specific genomic loci. Most commonly, gene targeting has been used to completely inactivate a specific gene (**Figure 4**). A selectable marker (usually neomycin phosphotransferase, or *neo*, expressed from a promoter active in ES cells) can be placed into a convenient restriction enzyme site of an exon, introducing stop codons into all three reading frames of the gene. This will completely inactivate the gene as long a stable truncated protein is not produced, or splicing does not occur around the exon containing the inserted *neo* gene. Alternatively, *neo* can be used to replace exons containing critical protein domains. The goal of these types of disruptions is to create complete loss-of-function or null alleles.

It is possible to make more subtle changes to genes using gene targeting. Various strategies have been used to introduce point mutations at specific residues of several genes. For example, a point mutation can be introduced in an exon of a gene, closely linked to the selectable marker. Clones that have undergone homologous recombination can be screened for the presence of the point mutation by restriction mapping or PCR. Conditional knockouts are designed to inactivate a gene in specific tissues at specific times of development or adulthood. Currently, several laboratories are refining this technology, involving Cre/*loxP*-mediated recombination and tetracycline-mediated induction and repression of gene expression to allow such subtle genetic alterations to be done routinely. Finally, it is possible to create large deletions in the mouse germ line using Cre/*loxP* technology. *LoxP* sites are introduced into the genome to surround a region to be

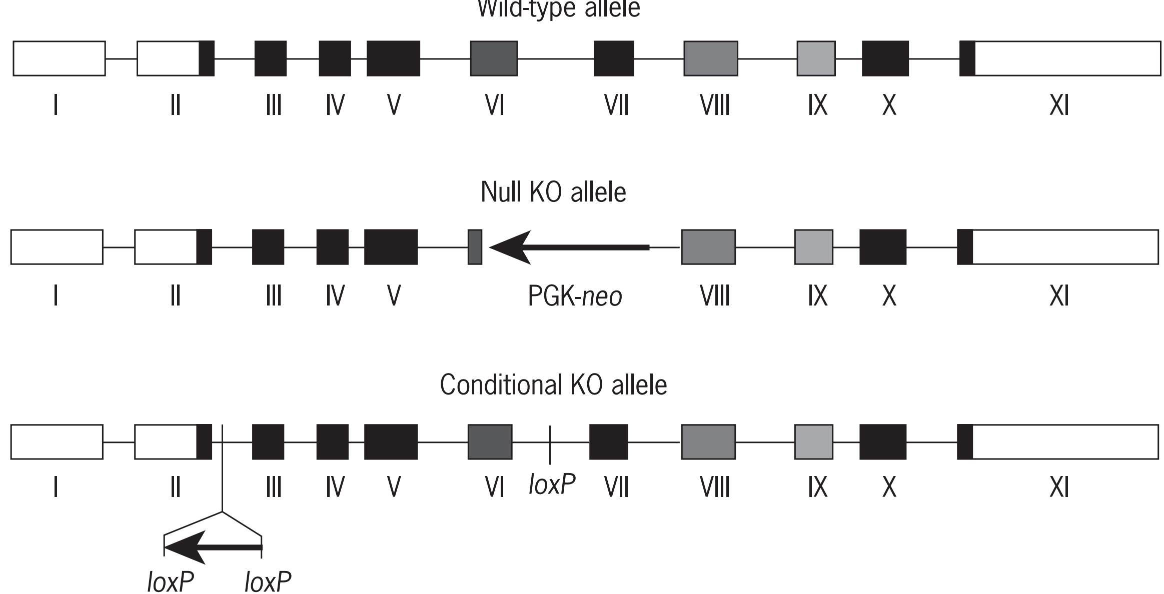


Figure 4 The general form of the most common types of gene targeting constructs. A wild-type gene with 11 exons (boxes, where clear boxes are 5′ and 3′ untranslated sequences and the filled boxes are coding sequences) is shown at the top. For the knockout construct (Null KO allele, middle), part of exon 6 and all of exon 7 are replaced by a gene that can be selected in culture, usually the neomycin transferase gene driven by a broad strong promoter (PGK-*neo*). For the conditional KO allele (bottom), the PCK-*neo* gene (arrow) surrounded by *loxP* recombinase sites are placed in intron 2, and a third *loxP* site is placed in intron 6. The inclusion of the *neo* gene in an intron often results in a hypomorphic allele, so this gene is surrounded by *loxP* sites for removal.

deleted. Cre-mediation deletion of the entire region can be accomplished in cell lines or in animals. Using these approaches, deletions of more than 1 Mb have been made *in vivo*.

The use of specific strains for the derivation of ES cells which have dominant coat colour, in combination with the isolation of blastocysts from strains with a recessive coat colour allow for the simplified screening procedure and identification of offspring in which the ES cells have contributed to the mouse. The most commonly used mouse strain for derivation of ES cells is one of the 129 strains. This strain carries the agouti coat colour locus that is dominant over the recessively inherited black coat colour locus of the C57BL/6J mouse strain from which the host blastocysts are obtained. Offspring in which the 129 ES cells have contributed to the embryo will have agouti and black fur due to the difference in origin of the cells, agouti from the ES cells or black from the cells of the host blastocyst. This is referred to as coat colour chimaerism. If ES cells have also contributed to the germ line, then mating of the chimaera to a female with a recessive coat colour results in offspring with the coat colour of the ES cell-derived strain. ES cell-derived offspring can then be screened for the presence of the mutated allele. Once animals heterozygous for the recombined allele have been identified, they can be interbred to generate animals homozygous for the mutated allele. In this way, null mutations or ‘knockouts’ of a specific gene can be obtained.

GENERAL APPLICATIONS: USE OF TRANSGENIC AND KNOCKOUT MICE TO STUDY THE FUNCTION OF CANCER-CAUSING GENES

In this section, examples will be presented that illustrate the value of genetic manipulation in the mouse to study the role of oncogenes and tumour-suppressor genes.

Study of Oncogenes *In Vivo* in Transgenic Mice

Mouse mammary tumour virus (MMTV) is a retrovirus that, like other retroviruses, integrates randomly into DNA. As a consequence of integration, it induces mammary tumours in a high proportion of infected animals. It was discovered that MMTV harbours regulatory sequences in its long terminal repeat (LTR) that allow for high levels of expression in the breast. Random integration of this LTR near the site of an oncogene that can promote breast cancer results in mammary tumorigenesis. The LTR of MMTV has been isolated from the integrated provirus and

characterized in transgenic mice. It has been found to be highly efficient in directing transgene expression to the mammary ductal epithelium.

Several investigators have taken advantage of these properties of the MMTV promoter to make transgenic mice that overexpress oncogenes in the breast. The first of these were performed in the laboratory of Philip Leder at Harvard (Stewart *et al*., 1984). A number of transgenic constructs were produced that fused the normal mouse c-*myc* gene with various lengths of its own promoter to the MMTV promoter (**Figure 5**). c-Myc is a basic helix–loop–helix (bHLH) transcription factor. It heterodimerizes with another bHLH transcription factor Max to activate transcription at E-box sites in promoters. It was one of the first identified oncogenes. In these first studies, two female founders of thirteen transgenic lines developed spontaneous mammary adenocarcinomas during an early pregnancy. The tumours and the breast tissue of these founder animals expressed RNA transcripts of the fusion gene. In addition, all female progeny of these two founder lines that inherited the MMTV/c-*myc* transgene also developed mammary adenocarcinomas during their second or third pregnancies. These mice were the first to describe the utility of the MMTV promoter in directing expression in the breast of transgenic mice, and were the first genetically manipulated mice to be susceptible to cancer. These mice were also the first mice to be patented and have been termed the Oncomouse.

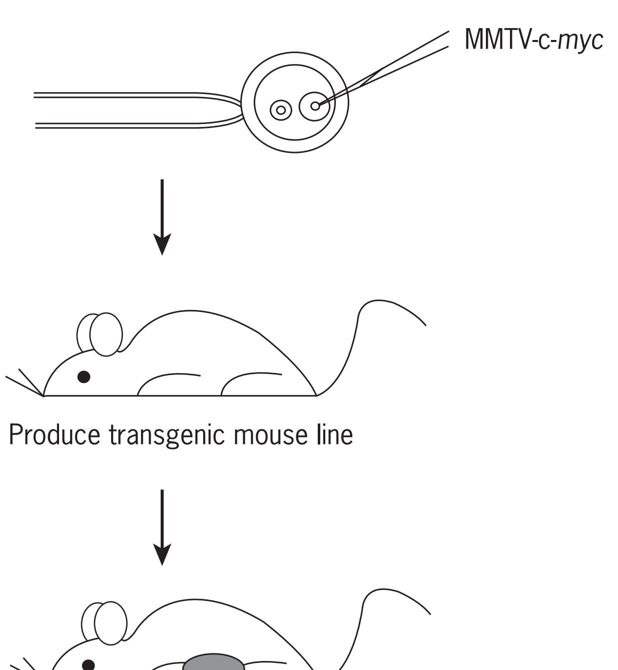


Figure 5 The production of MMTV-c-myc transgenic mice that get breast cancer.

Inactivation of Tumour Suppressors by Gene Targeting

In response to cellular DNA damage, cell cycle checkpoints are activated to arrest progression so that DNA can be repaired, and, if damage is extensive, the cell undergoes apoptosis. p53 plays a central, crucial role in regulating the distinct pathways that control these responses after DNA damage by diverse agents such as ultraviolet light, ionizing radiation (IR) and chemical carcinogens. p53 is maintained at a low level within the normal, undamaged cell. When a DNA damage signal is transmitted to p53 via upstream molecules that sense and respond to particular types of DNA damage, p53 is induced and regulates distinct downstream cell cycle checkpoint and apoptotic pathways. The importance of p53 in DNA damage responses is highlighted by the fact that mutation of p53 is the most common genetic abnormality in human cancer. Humans with Li–Fraumeni syndrome carry heterozygous p53 mutations and are predisposed to a wide variety of tumours.

p53 is a transcription factor that confers its effects by DNA binding and transcriptional regulation (Ko and Prives, 1990; Levine, 1997). p53 activates transcription of several genes by binding as a tetramer to p53-specific promoter elements, recruiting general transcription factors and activators to promoters of these genes. At least six genes have been demonstrated to be direct transcriptional targets of p53: *p21/WAF1/Cip1*, *mdm-2*, *GADD45*, *cyclin G*, *Bax* and *IGF-BP3*. Several of these genes have potential roles in cell cycle control and apoptosis, particularly *p21/WAF1/Cip1* (*p21*) and *Bax*. p53-dependent G1 arrest is mediated at least in part by transcriptional activation of p21, since p21 binds to cyclin/CDK complexes, and at high concentration p21 inhibits their activity. Bax forms a heterodimer with Bcl-2 and other family members, and these two proteins have opposing effects on apoptosis: Bax enhances apoptosis, whereas Bcl-2 promotes cell survival.

p53 activity can be increased by induction of protein levels or by activation of an inactive form of p53. DNA damage-induced increases in levels of p53 protein occur post-transcriptionally from decreased turnover or increased translation of p53. p53 can also be converted from inactive to active forms *in vitro* by a variety of agents, including antibodies to specific regions of p53, redox conditions, single stranded DNA and phosphorylation, suggesting that p53 activation may occur *in vivo*. Altered phosphorylation is a particularly attractive mechanism for rapid regulation of p53 activity. There are several phosphorylation sites at the N- and C-terminal regions of p53, and p53 can be multiply phosphorylated. These regions of the protein may be required for growth inhibition, suggesting a potential role for phosphorylated forms of p53 in cell cycle responses.

Because of the central role of p53 in choreographing the response to DNA damage, animals deficient for p53

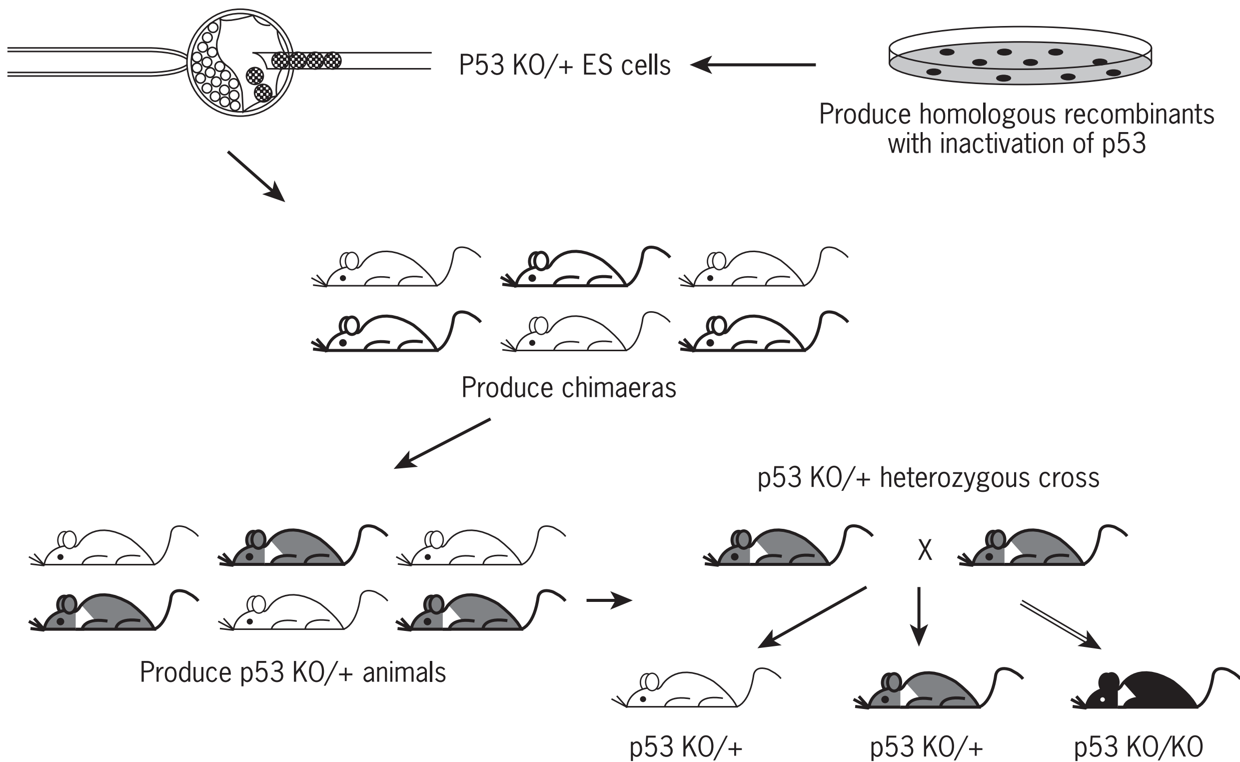


Figure 6 The production of p53-deficient mice by gene targeting.

were produced in two different laboratories (Donehower *et al*., 1992; Jacks *et al*., 1994). These animals were produced by gene targeting in the mouse (**Figure 6**). In both laboratories, it was found that p53 homozygous null mice were viable and developed normally, but displayed the spontaneous development of a variety of cancers by 4–6 months of age. Heterozygous animals also have an increased cancer risk, although with a longer latency to tumour formation. The distribution of tumour types in p53 heterozygous animals differs from that in homozygous mutants. In most cases, tumorigenesis in heterozygous animals is accompanied by loss of the wild-type *p53* allele. (See also chapter on *Regulation of the Cell Cycle*.)

Conditional Gene Knockouts for Tissue-specific Gene Inactivation

Breast cancer is the most common female cancer and the second leading cause of mortality in women. Approximately one in nine females are affected during their lifetime. About 5% of these women display an inherited form of breast cancer caused by loss of function mutations in the *BRCA1* gene. In fact, 90% of women with familial breast and ovarian cancers and about 50% of women with familial breast cancer alone display heterozygous germ-line mutations in *BRCA1*. Consequently, as for p53, many groups attempted to make a mouse model for this important human disease. However, it was found that homozygous *BRCA1*-null mice died during early embryonic development, and heterozygous mice did not develop tumours.

To overcome the early lethality, the *Brca1* gene was inactivated specifically in mammary epithelial cells of female mice using the Cre–*loxP* recombination system (**Figure 7**) (Xu *et al*., 1999). Mice were generated that carry a *Brca1* null allele, a *Brca1* conditional allele where exon 11 was flanked with *loxP* recombinase sites, and a *Cre* transgene under the control of the MMTV-LTR or the whey acidic protein (*WAP*) gene promoter, another mammary epithelial promoter. Cre-mediated excision of *Brca1* exon 11 occurred in the majority of mammary epithelial cells and was accompanied by a sharp reduction of *Brca1* transcripts. The reduction of BRCA1 impaired ductal outgrowth during pregnancy, and was frequently accompanied by apoptosis. Specifically, the ducts did not fully penetrate the fat pad. These results demonstrate that BRCA1 controls ductal elongation and branching morphogenesis during mammary gland development.

Most importantly, though, mammary tumour formation occurred after a long (10–12 months) latency. These tumours displayed genomic instability that was characterized by aneuploidy, chromosomal rearrangements and alterations of p53 transcription. Mating of these mice with the p53 null mice demonstrated that a loss of p53 accelerated the formation of mammary tumours. This example demonstrates the versatility of options available to study

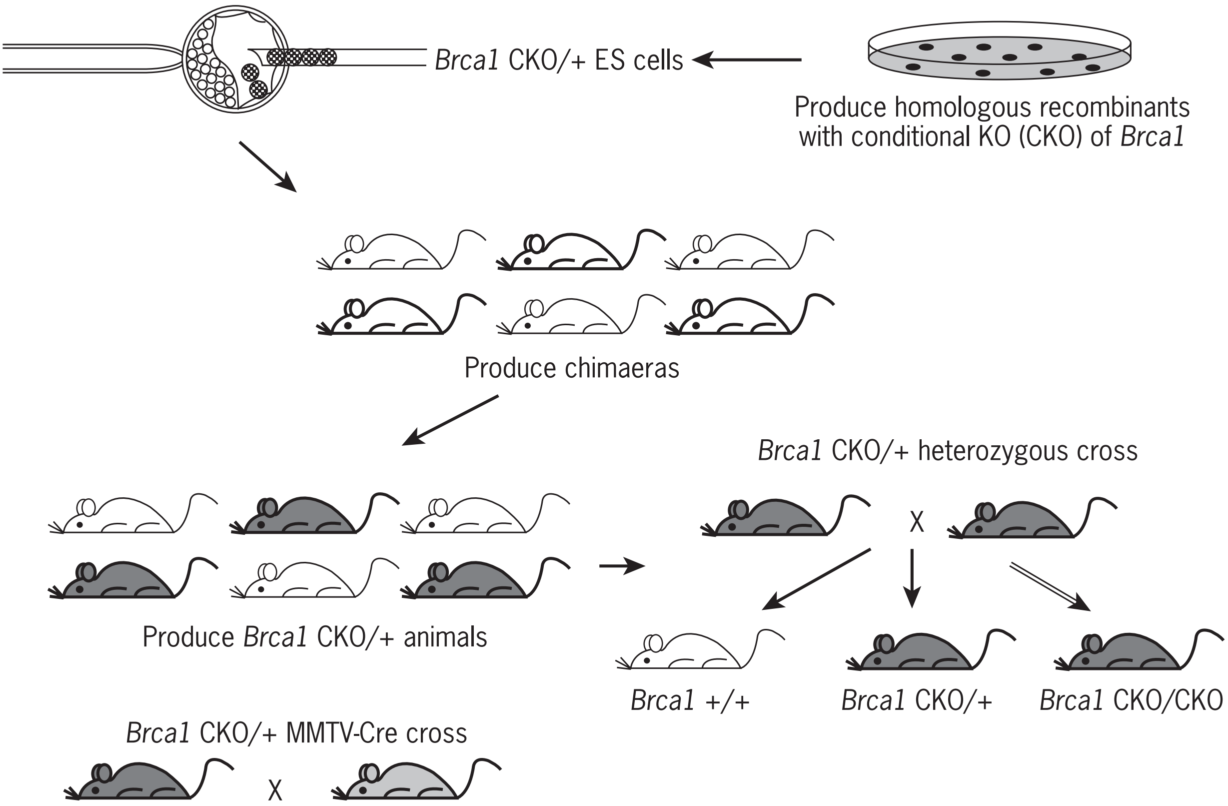


Figure 7 The production of *Brca1* conditional KO mice by gene targeting.

the role of cancer-causing genes *in vivo* in the mouse, even when the gene is essential for normal development.

Construction of Mice Bearing Deletions of Chromosome 11

Although single gene mutations often predispose to the development of cancer, tumorigenesis is ultimately caused by genomic instability. Among the changes that are often observed in tumours are regions of chromosomal deletion and duplication. These changes arise only in the tumour (i.e. they are not germ-line mutations). In addition, similar regions are often deleted or duplicated in the same type of cancer, suggesting that these regions contain genes important for certain types of cancer. The identification of the causative gene in these regions is often difficult.

Over the past few years, Allan Bradley's laboratory has been developing the ability to make deletions and/or duplications across an entire chromosome using ES cell technology (**Figure 8**), and approach that they have termed 'chromosome engineering' (Ramirez-Solis *et al*., 1995). Their first study focused on a small portion of chromosome 11. They constructed deficiencies, duplications and inversions in ES cells with sizes ranging from 1 Mb to 22 centimorgans (cM). Two deficiencies and three duplications were made into mice. Mice with a 1-Mb duplication developed corneal hyperplasia and thymic tumours, whereas two different 3–4-cM deficiencies were embryonically lethal in heterozygous mice. A duplication

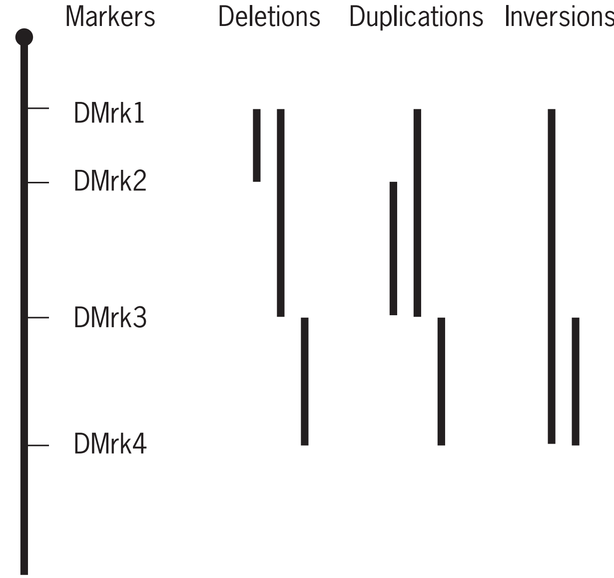


Figure 8 Deletions produced in mice from chromosome engineering. (Adapted from Liu *et al.*, 1998.) *DMrk1*, *DMrk2*, *DMrk3* and *DMrk4* refer to hypothetical marker sites.

corresponding to one of these two deficiencies was able to rescue its lethality, demonstrating that haploinsufficiency was in fact the cause of the lethality. Thus, they were able to produce mice that had defined deletions and duplications, and one of the deletions resulted in cancer.

INTERPRETATION OF EXPERIMENTAL RESULTS

Transgenic mice produced to express a dominant oncogene in a particular tissue are examined at regular intervals (usually weekly) for the presence of tumours in that tissue. Using the example of MMTV-driven oncogenes, breast tumours can be palpated along the lateral body wall of mice. Mice have five pairs of mammary glands, and tumours can arise in any of the glands. There are some unique properties to the MMTV promoter. Pregnancy stimulates expression from this promoter, so in practice, female mice bearing the MMTV-oncogene transgene are continuously mated to generate the highest possible levels of transgene expression. MMTV is often also expressed in the salivary glands and Harderian glands. Harderian glands are specialized tear glands located in the orbit. These tissues can also be the site of tumours in MMTV-oncogene transgenic mice.

Litter mate or age-matched controls are used to compare latency to tumour detection. Wild-type mice display spontaneous tumours only in rare cases and at relatively old age. Therefore, reproducible tumorigenesis in the target tissue is strong evidence for an oncogenic effect of the transgene.

For tumour suppressors in which the gene is inactivated in all tissues, any tissue could be the site of a tumour. The commonly used inbred mouse strains are somewhat susceptible to lymphomas, so lymphomagenesis is a likely outcome of broad inactivation of tumour-suppressor genes. As an example of this, p53 mutant mice mutant display lymphomas at a high frequency. As described above, litter mate or age-matched controls are used to compare latency to tumour detection, and reproducible tumorigenesis in any tissue is strong evidence for tumour-suppressor effect.

Latency to tumorigenesis or to tumour-free survival in a population of animals is normally represented as a Kaplan–Meier plot. Tumorigenesis is represented in a population by determining the age at which all animals in that population succomb to tumours (**Figure 9**). The proportion of animals surviving at each age is plotted over the entire life span of animals in a population. As mentioned above, wild-type animals rarely display spontaneous tumours, so the surviving fraction is usually 100% at all ages. In contrast, a population of tumour-prone animals will display a progressive decline in the numbers of animals. The age at which 50% of animals survive is used as a simple comparative measure between two populations.

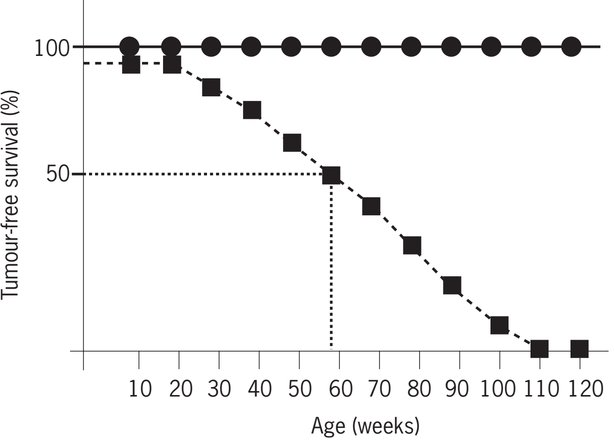


Figure 9 Kaplan–Meier survival analysis of wild-type (circles) and mutant cancer-prone (squares) mice. Note that 50% of the mutant mice have died by 58 weeks.

Genetic interactions can be determined in mice by mating two or more tumour-prone mice, and performing genetic epistasis analysis on the double mutants. If there is a genetic interaction between two different mutants, this provides strong evidence that the two mutant proteins cooperate in tumorigenesis. For example, suppose a putative oncogene 'A' overexpressed in the breast gives no detectable tumours, and that a different oncogene 'B' overexpressed in the breast gives tumours in 50% of animals by 1 year of age. Suppose now that double 'A + B' mutant mice with both transgenes overexpressed in the breast display aggressive breast tumours in 50% of animals by 2 months of age. These data suggest that these two oncogenes interact to promote tumours in the breast.

The hallmark of cancer is genomic instability, and it can be measured using several approaches (Ried *et al*., 1997). The major cytogenetic approaches are multicolour FISH techniques such as spectral karyotyping (SKY) (Schröck *et al*., 1996) and comparative genomic hybridization (CGH). SKY is a method which allows the simultaneous, unambiguous identification of each chromosome using individual chromosome painting probes, and is primarily used to determine whether translocations are present in a tumour sample (**Figure 10; see colour plate section**). SKY allows for the rapid identification of any chromosome, and is performed using flow sorted individual chromosome libraries. Using SKY and mouse chromosomal painting probes, all mouse chromosomes can be analysed and identified simultaneously in metaphase spreads and each displayed using a different colour. Metaphase chromosomes are prepared from a tumour using standard procedures. The mouse chromosome specific painting probes are generated from flow sorted mouse chromosomes differentiated by size. DNA labelling of each chromosome is performed in the presence of different combinations of five fluorochromes, such that

after excitation, each chromosomal probe will emit a slightly different wavelength of light. *In situ* hybridization is performed by combining the labelled DNA pools with an excess of a *Cot-1* fraction of mouse DNA to reduce background from repeat sequences. Multicolour hybridization is visualized using a combination of Fourier transform spectroscopy, CCD imaging and optical microscopy. The spectrum of the emitted light is recorded by the CCD camera and measured by the spectrophotometer at each pixel of the image to assign a specific spectrum to each chromosome. The measurement of specific emission spectra then allows the simultaneous display of all mouse chromosomes, each with a different colour, as shown in **Figure 10; see colour plate section**. Note how easy it is to determine the presence of multiple translocations in a metaphase spread. The location of translocation breakpoints can be used to identify genes deregulated as a result of the genomic rearrangement. Often the location will suggest candidate genes that can be tested using conventional FISH probes. Alternatively, the breakpoint can be cloned by conventional molecular biological techniques.

CGH is used primarily to identify regions of chromosomal gains and losses in a tumour (**Figure 11; see colour plate section**). DNA is prepared from a tumour and a nontumour control tissue or cell line. It is important that this control DNA is uniformly diploid, since it will be used as a standard to compare gene dosage in tumours. The control DNA is labelled with one fluorochrome (green, for example) and the tumour DNA is labelled with another fluorochrome (red, for instance). These two probes are mixed in equimolar amounts, and used as a hybridization probe against a normal metaphase chromosome spread. If a chromosomal region is diploid in the tumour sample, there will be equal hybridization of the control and tumour DNA, and the chromosomal region will appear yellow. In contrast, if there are regions of amplification or deletion in the tumour, then the chromosomal region will appear red for amplification and green for deleted regions. These CGH metaphase spreads can be captured and analysed by commercially available software dedicated to CGH analysis, and can even estimate relative levels of amplification and loss of chromosomal regions. The location of these genomic rearrangements can point to candidate genes in that region of the chromosome for further analysis.

Microsatellite instability is an additional type of genomic instability seen in mismatch repair proteins of the MLH and MSH families. Microsatellites are small repeat sequences present in the genome that can change sizes in the absence of fully functional mismatch repair. The length of these microsatellites can be determined by sensitive PCR assays, and this assay can be used to provide evidence for the involvement of mismatch repair pathways in tumorigenesis of a particular mutant.

It is clearly of great interest to understand the genes and ultimately the genetic pathways that are disrupted during the genesis of a particular tumour. It has recently become possible to analyse global gene expression by the use of a variety of PCR-based techniques to assay broadly for changes in gene expression. Subtracted libraries are used to clone mRNAs that show differences between the reference and test samples. Serial analysis of gene expression (SAGE) and differential display methods employ PCR techniques to measure differences between two different populations of mRNA, and to clone these differences. Differential display techniques use oligonucleotides synthesized with all possible sequence combinations to assess broadly gene expression between two populations of mRNA. Each of these methods depends upon large differences between reference and test samples. Recently, it has become possible to sample the expression pattern of thousands of genes simultaneously using high-density cDNA arrays (Duggan *et al*., 1999). The advantage of this method is that each spot of a microarray represents a known gene. As all genes of a genome are discovered, it is possible to develop arrays containing every gene in a genome to assess quantitatively mRNA expression of every gene simultaneously.

There are several available options to carry out reproducible, high-throughput analysis of the expression of thousands of genes simultaneously using microarrays. It is possible to purchase premade filters or arrays that contain hundreds to thousands of cDNAs. These filters are hybridized against radioactive probes made from a single source, so two filters must be purchased for any comparative study, and they must be read on a phosphorimager. Premade filters have little flexibility, in terms of adding genes as the human and mouse genome sequencing efforts progress, but have the advantage of being simple to use in any laboratory that performs Southern or Northern blot hybridization analysis. In addition, it is thought that these filters will be useful for detecting only high- and medium-abundance mRNAs. Another option is to use high-density microarrays of cDNAs or of oligonucleotides that represent parts of different cDNAs made by companies such as Incyte or Affymetrix. At the moment, these commercially available arrays are expensive, and the supply of chips is limiting.

Microarray technology has been developed in academic laboratories and industry as a cost-effective and reliable means to analyse comparatively genome-wide patterns of mRNA expression (**Figure 12; see colour plate section**). Ultimately, the goal is to develop arrays containing every gene in a genome to assess quantitatively mRNA expression of every gene simultaneously. At this time, every gene is known only for three eukaryotic organisms: *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster*, where the genomes have been completely sequenced. A working draft of the human genome has now been published (February 15, 2001, *Nature*, **409**; February 16, 2001; *Science*, **291**), and a working draft of the mouse genome is predicted for 2005. For mammalian organisms at present, UniGene sets for the human (15 000, cDNAs, or approximately 15% of the

human genome) and mouse (8000 cDNAs, or approximately 8% of the murine genome) are commercially available. cDNA-based microarrays are made by robotically printing cDNA inserts on to a glass slide and subsequently hybridizing these slides to two differentially labelled fluorescent probes. The probes are pools of cDNAs generated after isolating mRNA from cells or tissues in two or more different states, so that one can compare expression patterns of all genes under different conditions. The resulting fluorescent intensities are assayed with a fluorescent reader, and the ratio of expression in the two states is obtained following image processing. The crucial elements to this technology are an arrayer, a representative cDNA set, a microarray reader, image analysis software and database management protocols. With these elements, it is possible to compare global RNA expression patterns in two or more different RNA pools and to define differences in expression between pools. Several academic and commercial groups have developed robotic arrayers to place thousands of cDNA clones on defined spots of an array, and these arrayers are often matched with chip readers to capture the data after hybridization. In addition, many groups are working to develop the bioinformatics capabilities needed to handle this very large volume of data.

BIOLOGICAL LIMITATIONS

In many ways, the mouse is the perfect experimental system for the creation of models of human genetic diseases such as cancer. The genome size and number of genes are similar between human and mouse, as are patterns of development. It is relatively cheap to house mice, in comparison with other mammals. Mice have relatively short gestation periods, a brief time of maturation to sexual maturity and large litter sizes. Most importantly, it is the best mammalian species for genetic manipulation, and the only one where it is currently possible to perform germline knockouts in embryonic stem cells.

Although the mouse is good as an experimental system, it is not perfect for modelling human diseases. Often a knockout or transgenic mouse has a phenotype that is different from the human genetic disease. There are many explanations for why this might occur, and a few examples are provided here.

1. There may be redundant pathways in the mouse not present in the human. There are many examples of this in the literature. Mice have redundant pathways not present in the human for Tay–Sachs disease, Lowe's syndrome and galactokinase deficiency, to name a few examples.
2. The pattern of overexpression of the transgene is not similar to that of the human disease gene. The inherent variability of expression of transgene expression in mice often leads to unpredictable and variable phenotypes. Regulatory elements may not be present in the transgene, or the site of integration may modify the pattern of expression.
3. The developmental pathway that is disrupted in the human disease is either not present or is different in the mouse. In general, the mouse is related closely enough to humans that this should not pose a major problem. However, there must be genes in the human that are involved in unique processes such as complex brain folding and development which may not have a similarly important role in the mouse. If so, it may not be possible to use the mouse to model certain genetic diseases.
4. The life span of the mouse is too short to develop cancer. An advantage to using the mouse as a model is that it develops rapidly. However, if a certain type of cancer requires years to develop, a mouse may not live long enough to develop tumours.
5. The human mutation is not faithfully reproduced in the mouse. Transgenic constructs are usually made with promoters that express cDNAs in a desired spatial and temporal pattern. Knockout constructs are usually made by insertion of a *neo* gene into the gene that is being targeted for positive selection, with or without the deletion of coding sequences. These strategies, although extremely valuable, may not faithfully reproduce the mutation that is seen in a human disease.
6. Inherent tumour spectrum of inbred mouse strains. Many inbred mouse strains used for transgenic experiments seem particularly susceptible to certain types of tumours. For example, lymphomas are commonly seen in mice with inactivation of tumour suppressors such as *p53* and *ATM*. Fortunately, many inbred strains are available that can be used to alter the tumour spectrum displayed in a particular transgenic mouse.
7. Developmental role for the cancer-causing gene. Although a somatic mutation in a particular gene can lead to cancer, the complete inactivation of the same gene in the germ line could result in severe developmental defects. As described above, it is possible to inactivate genes in somatic tissues using conditional knockouts of such genes in conjunction with mice that express recombinase in the appropriate spatial and temporal pattern. This type of approach was illustrated above for *Brca1*.

Notwithstanding these limitations, the mouse remains an excellent mammalian genetic organism to model human cancer, and will likely be the model of choice for the next decade and longer.

PERSPECTIVES

The development of transgenic and gene targeting techniques for genetic modification of the mouse has resulted

in tremendous insights into the genetic processes that cause cancer. The continued use of these techniques to analyse *in vivo* the consequences of these genetic changes will allow for even greater insights into the genetic causes of cancer. With the imminent completion of the human and mouse genomic sequences, we will soon have at our disposal a catalogue of all of the genes in these mammalian organisms. The importance of any of these genes to tumorigenesis can be tested directly in the mouse using the techiques described in this chapter. Although much has been learned, much remains to be unravelled using these approaches.

REFERENCES

Donehower, L. A., *et al*. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, **356**, 215–221.

Duggan, D. J., *et al*. (1999). Expression profiling using cDNA microarrays. *Nature Genetics*, **21**, 10–14.

Jacks, T., *et al*. (1994). Tumour spectrum analysis in p53-mutant mice. *Current Biology*, **4**, 1–7.

Ko, L. J. and Prives, C. (1996). p53: puzzle and paradigm, *Genes and Development*, **10**, 1054–1072.

Levine, A. J. (1997). p53: the cellular gatekeeper for growth and division. *Cell*, **88**, 323–331.

Liu, P., *et al*. (1998). Embryonic lethality and tumorigenesis caused by segmental aneuploidy on mouse chromosome 11. *Genetics*, **150**, 1155–1168.

Ramirez-Solis, R., *et al.* (1995). Chromosome engineering in mice. *Nature*, **378**, 720–724.

Ried, T., *et al.* (1997). Tumour cytogenetics revisited: comparative genomic hybridization and spectral karyotyping. *Journal of Molecular Medicine*, **75**, 801–814.

Schrock, E., *et al.* (1996). Multicolour spectral karyotyping of human chromosomes. *Science*, **273**, 494–497.

Stewart, T. A., *et al.* (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell*, **38**, 627–637.

Wynshaw-Boris, A. (1996). Model mice and human disease. *Nature Genetics*, **13**, 259–260.

Xu, X., *et al.* (1999). Conditional knockout of *Brca1* in the mammary epithelial cells results in blunted ductal morphogenesis and tumor formation. *Nature Genetics*, **22**, 37–43.

FURTHER READING

Donehower, L. A., *et al*. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, **356**, 215–221.

Hogan, B., *et al*. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Jacks, T., *et al*. (1994). Tumor spectrum analysis in p53-mutant mice. *Current Biology*, **4**, 1–7.

Joyner, A. L. (1993). *Gene Targetting: A Practical Approach.* (Oxford University Press, Oxford).

Ko, L. J. and Prives, C. (1996). p53: puzzle and paradigm. *Genes and Development*, **10**, 1054–1072.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323–331.

Ramirez-Solis, R., *et al*. (1995). Chromosome engineering in mice. *Nature*, **378**, 720–724.

Ried, T., *et al*. (1997). Tumour cytogenetics revisited: comparative genomic hybridization and spectral karyotyping. *Journal of Molecular Medicine*, **75**, 801–814.

Schröck, E., *et al*. (1996). Multicolour spectral karyotyping of human chromosomes. *Science*, **273**, 494–497.

Stewart, T. A., *et al*. (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell*, **38**, 627–637.

Wassarman, P. M. and DePamphilis, M. L. (1993). *Guide to Techniques in Mouse Development. Methods in Enzymology*, Vol. 225. (Academic Press, San Diego.)

Xu, X., *et al*. (1999). Conditional knockout of *Brca1* in the mammary epithelial cells results in blunted ductal morphogenesis and tumor formation. *Nature Genetics*, **22**, 37–43.

WEB PAGES

The Jackson Laboratories: http://www.jax.org/.

Induced Mutant Resource: hhttp://www.jax.org/resources/documents/imr/.

Oak Ridge National Laboratories Mutant Mouse Database: http://www.bio.ornl.gov/htmouse/.

Trans-NIH Mouse Initiative: http://www.nih.gov/science/models/mouse/.

The Whole Mouse Catalogue: http://www.rodentia.com/wmc/.

Chapter 57

Gene Knockouts in Cancer Research

Taha Merghoub and Pier Paolo Pandolfi
Memorial Sloan-Kettering Cancer Center, New York, NY, USA

CONTENTS

- Biological Background: Mouse Modelling of Cancer: A Needed Tool?
- Principles of Model Establishment
- General Applications and Interpretation of Experimental Results
- Biological Limitations and Technical Considerations
- Perspectives

BIOLOGICAL BACKGROUND: MOUSE MODELLING OF CANCER: A NEEDED TOOL?

'Mouse modelling' is intrinsically and conceptually an inappropriate definition in that it implies that the ultimate goal of these experiments is to reproduce in the mouse a human disease. On the contrary, the data accumulated in the past two decades have shown that these experiments go beyond pure mimicry of the disease itself, and have indeed enabled us to gain tremendous insight and dissect the molecular mechanisms underlying oncogenesis and the pathogenesis of any cancer.

Numerous laboratories have used the transgenic approach as well as gene targeting in order to create mouse models for the study of cancer pathogenesis. The analysis of these mutants has led to a better understanding of the normal function of genes involved in tumorigenesis, the relationship between the histopathological progression of cancer and acquisition of specific genetics lesions as well as many other aspects of disease progression such as the biology of metastasis, which have to be studied in the context of the whole animal.

Transgenic mice harbouring activated forms of proto-oncogenes or dominant negative forms of tumour-suppressor genes have been generated in order to test the causal relationship between their expression and tumour development as reviewed in the chapter *Transgenic Technology in the Study of Oncogenes and Tumour Suppressor Genes*. (See also the chapter on *Human Tumours in Animal Hosts*.) Mouse strains where putative tumour-suppressor genes have been inactivated (or 'knocked out') have been useful to assess the consequences of the lack of their function in tumorigenesis (reviewed in this chapter). Animal models have also determined the importance of specific gene mutations in tumorigenesis. These studies have provided an in-depth insight into cancer-related molecular pathways and the relationship between normal developmental processes and tumorigenesis. Interbreeding of transgenic and knockout mutants has allowed the introduction of multiple cancer-associated mutations in a single animal and, in turn, the discovery of cooperating mutations/pathways. Mouse models, unlike *in vitro* studies, allow one to integrate the analysis at the organismal and cellular levels with the overall physiological response of the animal to the disease (e.g. immune response).

Understanding genetic pathways still remains a prerequisite for the development of molecular and pharmacological therapies in order to treat and prevent cancer effectively. Thus, these models will represent an invaluable instrument not only to study the pathogenesis of human cancer, but also to exploit novel therapeutic approaches.

PRINCIPLES OF MODEL ESTABLISHMENT

The Knockout Approach in the Mouse

The possibility to manipulate and to culture pluripotent mouse embryonic stem (ES) cells has allowed for the generation of mice with a specific alteration in a chosen target gene. ES cells are in fact capable of contributing to the formation of all organs including the germ line when introduced into the developing blastocysts (3.5 dpc embryo). To engineer knockout mice, the first step is therefore to generate a targeting vector, which is transfected into ES cells (**Figure 1**). This vector contains sequences that allow the introduction of a specific disrupting mutation into a defined locus by homologous recombination (**Figure 1**). The vector is next transfected into embryonic stem cells. Recombination between the endogenous locus and the vector in the regions of homology occurs spontaneously since ES cells are competent for recombination. However, recombination occurs at very low frequency, thus rendering necessary a selection step to identify the cells

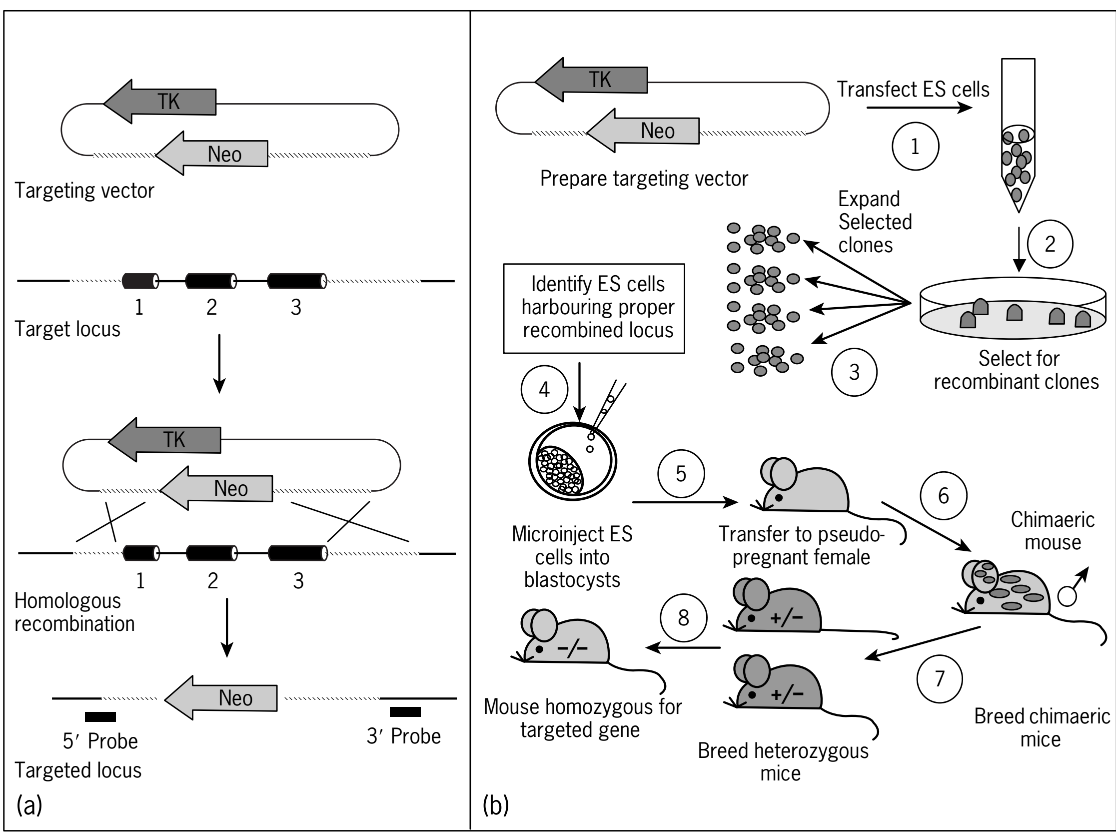


Figure 1 Schematic representation of the strategy for the generation of knockout mice. (a) targeting strategy. The targeting vector contains two regions homologous to the desired chromosomal integration site (▨). Homologous recombination between sequences of the targeting vector and the target locus will result in the disruption of the target gene by the replacement of the region containing exons 1, 2 and 3 (as an example) with the positive selectable marker: the *neo* (neomycin phosphotransferase) resistance gene cassette. The HSV*tk* (herpes simplex thymidine kinase) gene, which is used as a negative selectable marker and renders the cells sensitive to Gancyclovir, is lost in the case when proper recombination has occurred. Thus, recombined clones become resistant to both G418 and Gancyclovir. Southern blot analysis on individual clones with appropriate probes external to the targeting vector will identify the recombined locus in view of the fact that the presence of the *Neo* cassette will change the configuration of the targeted locus. The hatched lines represent the regions of homology to the target locus; the thin line represents the bacterial plasmid. The black cylinders identify the exons of the target gene. The two large crosses represent the crossover points. (b) Generation of knockout mice. (1) The replacement vector is linearized outside the region of homology and transfected into ES cells. (2) Transfected ES cells are then plated on feeder layer cells and subjected to positive selection (G418 in the case of the *neo* gene) and in some cases to negative selection (Gancyclovir in the case of HSV*tk* gene). This process gives rise to ES cells colonies after few days of selection (7–10 days). (3) The different ES clones are picked and expanded in order to have enough cells to prepare DNA. Genotyping is subsequently performed either by Southern blot or PCR in order to identify recombined clones. (4) Recombined clones are then independently microinjected in the blastocoelic cavity of blastocysts. (5) Injected blastocysts are transferred into pseudopregnant females. (6) Pups are born with various degrees of chimaerism. The chimaerism is scored by simple colour coat inspection. The ES cells are typically derived from 129*sv* mouse strain with agouti coat. The recipient blastocysts are generally derived from C57B/6 strain with black coat. Thus, a high degree of chimaerism will be shown by a highly agouti coat in chimaeric mice. (7) The chimaeric mice are crossed with wild-type mice to obtain heterozygous mutants in case the mutation has been transmitted to the germ line. (8) Finally, heterozygous mutants are intercrossed to generate null homozygous mice.

where the recombination took place. ES cells are therefore subjected to both positive and negative selection for markers, which are contained in the targeting vector (**Figure 1**).

Mutant ES cells that harbour a mutation on one allele (heterozygous or hemizygous for the mutation) are next microinjected into mouse blastocysts to generate mice constituted in part by cells of the recipient blastocyst

and the mutant ES cells. These organisms are therefore referred to as chimaeric mice (**Figure 1**). When ES cells contribute to the germ line of a chimaera it is possible to identify mice derived entirely from the mutant ES cells and therefore heterozygous/hemizygous for the mutation among its offspring. Subsequent interbreeding of heterozygotes yields animals homozygous for the desired mutation, which can be studied for the phenotypic manifestations of gene inactivation (e.g. tumorigenesis if the gene inactivated is a putative tumour-suppressor gene). This approach can be utilized not only to disrupt genes, but also to introduce specific mutations in the coding of a gene (e.g. an activating or dominant negative mutation in an oncogene or a tumour-suppressor gene, respectively). Gene targeting in mouse ES cells is, without a doubt, the most elegant and selective tool for the manipulation of the vertebrate genome ever devised in order to study *in vivo* the function of a gene and how it relates to human cancer pathogenesis. However, this approach also has limitations described in the following paragraphs that rendered necessary further implementations of this powerful technique.

The Cre–*loxP* System for *In Vivo* Tissue-specific Gene Targeting

As mentioned before, the conventional 'knockout' approach has been extremely useful in studying the function of numerous genes involved in cancer pathogenesis; however, in many instances the mutant mice die prenatally because the missing gene exerts a critical role in embryogenesis. In order to circumvent this limitation, a conditional targetting strategy utilizing the Cre–*loxP* system has been developed (**Figure 2**). The Cre–*loxP* system represents an ideal instrument for the manipulation of chromosomal organisation and the control of gene expression *in vivo*, in a tissue-specific manner, and at a specific time in development. Cre, the recombinase of bacteriophage P1, is able to mediate the specific excision of DNA flanked by *loxP* recognition sites from the genome of mammalian cells in culture and transgenic mice. The experimental procedure to generate mouse models by using this system consists of three distinct phases: (1) generate a mutant mouse, where the gene of interest is flanked by *loxP* sites; (2) generate transgenic animals in which the *Cre* gene is expressed in a tissue-specific manner; (3) intercross the mutant mice generated in step 1 with the ones generated in step 2. As a result, Cre will excise the intervening sequence, removing the coding region of the gene of interest and resulting in inactivation of the gene only in the tissue or cells in which the *Cre* gene is expressed. The consequences of the tissue specific gene disruption can now be analysed, thus circumventing embryonic lethality. A similar system termed Flp–frt has been also widely utilized. In this case, Flp is the recombinase and FLP recognition targets (frt) are the equivalent of the *loxP* sites.

The 'Knock-in' Approach

The 'Knock-in' approach refers to the generation of mice in which one particular gene (or sequence) is replaced by another gene or aberrant gene by homologous recombination (**Figure 3**). ES cells containing the targeted gene are generated following the same procedure as in the knockout approach. These cells are subsequently utilized to generate mice expressing the 'knocked-in' gene. This strategy is generally utilized as an alternative approach to the generation of transgenic mice which, although time consuming, offers several advantages. Compared with the transgenic approach, the main advantage is the possibility of expressing the gene of interest under the control of an endogenous promoter, thus providing expression levels similar to that observed in normal physiological situations (e.g. a cancer-associated fusion gene can be targeted into the locus of the parental gene, thus closely mimicking what happens at the somatic level in a cancerous cell upon a chromosomal translocation). An additional advantage of this methodology is that the copy number of the 'knocked-in' gene is a known parameter to start with, and can be increased simply by modifying the targeting vector. Finally, this approach can also be used to knock in a reporter gene such as GFP (green fluorescent protein) or Lac-Z, under the control of regulatory elements of the gene of interest, in order to define more precisely its expression pattern during embryogenesis and in adulthood (e.g. the reporter gene can be placed in frame with the first coding amino acids of the target protein).

GENERAL APPLICATIONS AND INTERPRETATION OF EXPERIMENTAL RESULTS

We will provide here representative examples of how knockout technology has been utilized to prove the critical tumour-suppressive role of genes implicated in human cancer pathogenesis. We will also assess the usefulness of some of these knockout mutants as models for the corresponding human cancer syndrome.

p53-deficient Mice (Li–Fraumeni Syndrome)

Mutations in the *p53* tumour-suppressor gene occur in about 40% of all human tumours and represent one of the most frequent genetic alterations observed in human cancer. Most *p53* alterations are missense mutations resulting in the expression of a functionally altered protein. Discovered

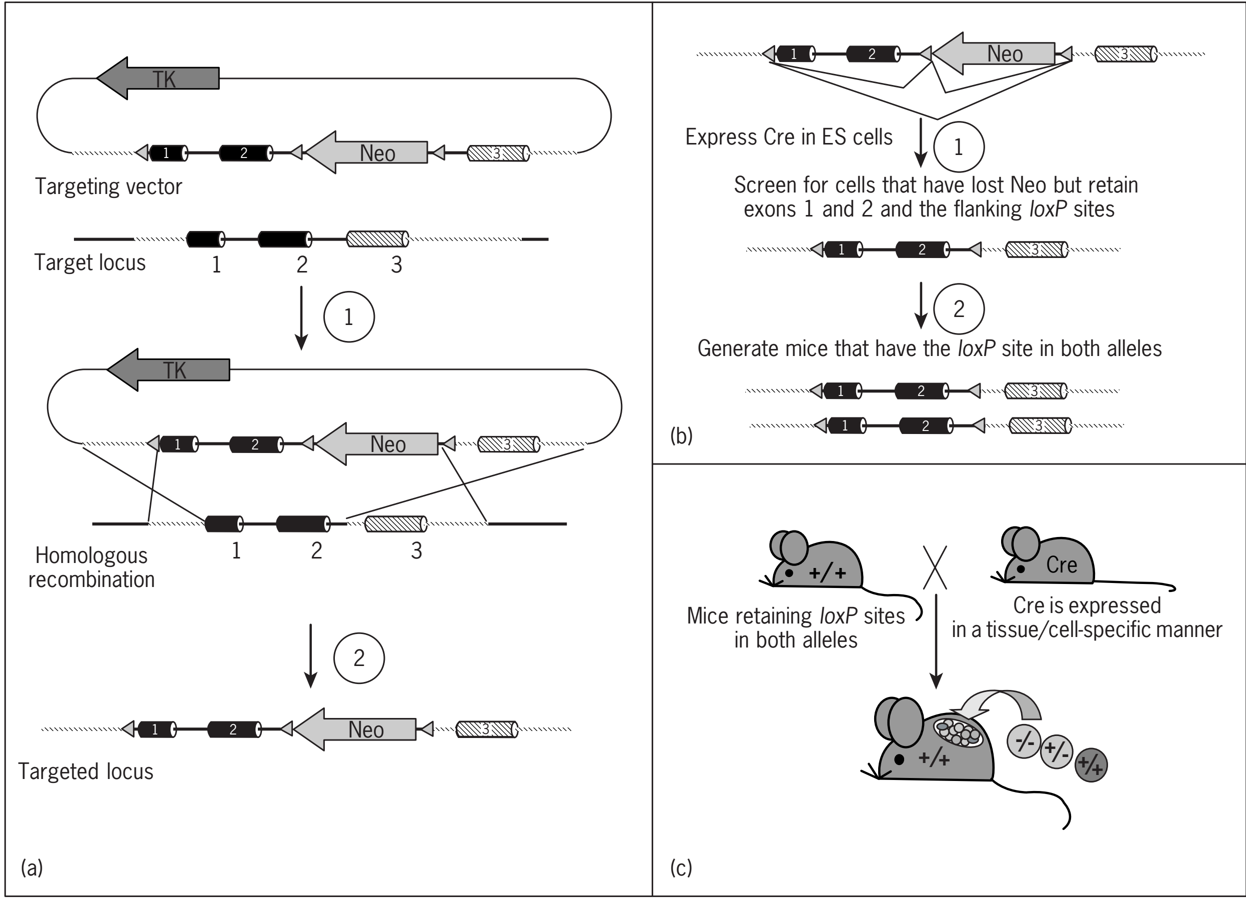


Figure 2 Schematic representation of the strategy for *in vivo* tissue specific Cre–*loxP* mediated recombination. (a) Targeting strategy. Following the same strategy as described in **Figure 1**, *loxP* sites and *neo* cassette will be inserted within intronic regions of a gene. (b) Cre-mediated recombination *in vitro*. In a first step, Cre recombinase is expressed in ES cells (this can be achieved by retroviral expression or by transient transfection). Cre will mediate the specific excision of DNA sequence flanked by *loxP* sites. *LoxP* sites are approximately 20 base pairs in length and do not interfere with the proper transcription of the locus. Consequently, this step results in the generation of ES cells with three distinct genotypes: (i) allele lacking *neo*, exons 1 and 2; (ii) allele lacking exons 1 and 2; (iii) allele only lacking *neo*. The ES cells lacking *neo* (iii) are expanded, and chimaeric mice are generated as described in **Figure 1**. Chimaeras and heterozygous mutants are intercrossed in order to generate mice homozygous for exon 1 and 2 flanked by two *loxP* sites. (c) Generation of tissue specific knockout mice. Transgenic mice expressing Cre in a tissue-specific manner are intercrossed with homozygous mice from step (b) In the resulting double mutants, Cre will excise exons 1 and 2, leading to the inactivation of the gene of interest only in tissues and cells where Cre is expressed. The hatched lines and hatched cylinders represent the vector homology to the target locus. Black cylinders indicate exons; the thin black line represents the bacterial plasmid. The triangle represents the *loxP* sites. The neo (neomycin phosphotransferase) is the positive selectable marker. The *TK* (herpes simplex thymidine kinase gene) is the negative selectable marker. The large crosses represent the crossover points.

in 1979 as a cellular protein complexed with a tumour virus protein, *p53* was initially considered an oncogene, because mutant p53 had an oncogenic activity in transformation assays. However, subsequent reports have shown that the *p53* gene encodes for a tumour-suppressor protein. Wild-type p53 has been referred to as the ‘guardian’ of the genome, as p53 responds to DNA damage or cell cycle checkpoint failure, either by arresting cells in the G1 phase of the cell cycle for damage repair or through the initiation of an apoptotic pathway to eliminate the damaged cell entirely. Whether arrest or apoptosis is induced, the result is that the cell with DNA damage is prevented from replicating damaged DNA templates, thus avoiding potential oncogenic lesions. The role of p53 in cell cycle arrest is mainly due to direct regulation of gene transcription. p53 can positively regulate the transcription of genes such as $p21^{\mathrm{WAF1/CIP1}}$, a G1 cyclin-dependent kinase inhibitor, *mdm*-2, *GADD*45, *IGF-BP3* and *Bax*. p53 is also involved in the negative regulation of certain genes that are implicated in positive growth regulation such as PCNA

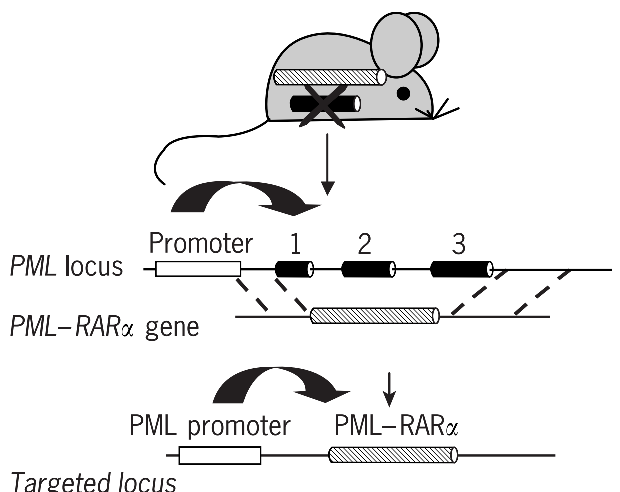


Figure 3 Schematic and representative example of a 'knock-in' experiment. In this case, a gene such as the PML gene of APL can be replaced by an oncogenic aberrant fusion gene such as *PML–RARα*. As a result, the expression of *PML–RARα* will be under the control of the promoter/regulatory regions of the PML gene as in human APL blasts. The targeting vector can be designed in a way that the first few exons of the *PML* gene are precisely replaced by the coding region of the *PML–RARα*.

and Bcl-2. p53 is critical for the maintenance of genomic stability. Several assays have shown that cells missing p53 display frequent genetic aberrations of various types: aberrant ploidy, gene amplification, increased recombination and centrosomal dysregulation. Mutation of p53 may result in the loss of wild-type function and also in the generation of dominant negative and gain of function mutants. p53 heterozygous mutations have been found in Li–Fraumeni patients which are susceptible to develop tumour of various histological origins. (See also chapter on *Regulation of the Cell Cycle*.)

Owing to its importance for cell cycle control and cancer prevention, the generation of p53 knockout mice was essential. Several groups have developed p53 null mice. These mice have a normal development, but, as expected, 75% of the p53 null mice develop tumours of various histotypes by 6 months of age and all die by 10 months of age. p53 null mice developed mainly thymic T cell lymphomas and soft tissue sarcomas and testicular teratomas and osteosarcomas. $p53^{+/-}$ mice are also prone to develop cancer. Tumours are observed starting from the ninth month of age and by 24 months more than 95% of the animals die. $p53^{+/-}$ mice are more prone to develop osteosarcomas and soft tissue sarcomas than null mice, which more commonly develop lymphomas.

The $p53^{+/-}$ mouse shares similarities with Li–Fraumeni patients. The tumour types found in mice are also observed in Li–Fraumeni patients. However, the overlap is not complete since, for instance, brain and breast tumours observed in Li–Fraumeni patients are not observed in the mouse model.

Of interest, Southern blot and sequencing analysis of the wild-type allele in tumours arising from $p53^{+/-}$ mice shows that it is structurally intact in more than 50% of tumours. Furthermore, the analysis of p53 function in these tumours shows that the p53 protein retains normal transcriptional activity and apoptosis induction. The inactivation of both copies of a tumour-suppressor gene is considered the prerequisite for tumour formation as postulated in the 'two-hit' model of Knudson. In the case of p53, however, the reduction of gene dosage to heterozygosity seems to be sufficient for tumorigenesis.

It is also worth noting that onset of tumorigenesis may vary depending on the mice strain. For instance, the $p53^{-/-}$ and $p53^{+/-}$ mice generated in the 129/*sv* background develop tumours sooner than mutants in a C57BL/6 background. Moreover their tumour spectrum is different. In the 129/*sv* background the mice develop mainly testicular teratomas and lymphomas, whereas the C57BL/6 mice develop mainly lymphomas. These differences demonstrate the importance of tumour-modifier genes in cancer susceptibility by p53 inactivation and make the mouse a unique tool for gene discovery, also in view of the fact that the mouse genome project will soon be completed.

PTEN-deficient Mice (Cowden Syndrome)

The *PTEN* (phosphatase and tensin homologue deleted from chromosome 10 gene), also known as *MMAC1* (mutated in multiple advanced cancers) or TEP1 (TGFβ-regulated and epithelial cell-enriched phosphatase) has been recently identified as a tumour-suppressor gene. Somatic mutation or deletions of this gene have been found in a large fraction of tumours, including glioblastomas, endometrial and prostate cancer, placing *PTEN* among the most commonly mutated genes in human cancers. Germline heterozygous *PTEN* mutations are associated with Cowden disease. This syndrome is characterized by developmental defects as well as susceptibility to breast, thyroid and brain tumours. *PTEN* is also mutated in two related familial syndromes, Lhermitte–Duclos disease and Bannayan–Zonana syndrome, which show similar predisposition to tumours but distinct developmental defects.

PTEN is the first example of a tumour-suppressor gene that encodes a phosphatase. PTEN contains a tyrosine phosphatase domain with features resembling those of the dual-specificity phosphatases, which are able to dephosphorylate both tyrosine and serine/threonine residues. Despite its homology with the dual-specificity phosphatses, PTEN is an inefficient protein phosphatase *in vitro*. However, phosphatidylinositol 3,4,5-triphosphate has proved to be the main PTEN substrate. Recent studies suggest that the focal adhesion kinase and Shc may be

regulated by PTEN. However, the tumour-suppressive function of PTEN is dependent solely on its lipid phosphatase activity.

Several groups have generated PTEN-deficient mice. Null embryos displayed an early embryonic lethal phenotype between days 6.5 and 9.5 d.p.c. (possibly depending on the genetic background), owing to aberrant differentiation and patterning of the embryonic germ layers. Analysis of the embryos shows poorly organized ectodermal and mesodermal layers and over growth of the cephalic and caudal region due to aberrant patterning.

The role of PTEN in tumorigenesis was confirmed by the fact that $PTEN^{+/-}$ mice develop a broad range of tumours including thyroid, prostate, small and large intestine, endometrial tumours and phaeochromocytomas. Furthermore, it has recently been shown that, the concomitant inactivation of one *Pten* allele and one or both $p27^{KIP1}$ alleles dramatically accelerates spontaneous neoplastic transformation and incidence of tumours of various histological origins. Cell proliferation, but not cell survival, is markedly increased in $Pten^{+/-}/p27^{-/-}$ mice. Furthermore, $Pten^{+/-}/p27^{-/-}$ mutants develop prostate carcinoma at complete penetrance within 3 months from birth. These cancers recapitulate the natural history and pathological features of human prostate cancer. Thus, PTEN and $p27^{KIP1}$ play a pivotal cooperative role in tumour suppression in epithelial tissues, and combined loss of PTEN and $p27^{KIP1}$ are key events in the pathogenesis of epithelial cancers as suggested by the short latency and complete penetrance of these tumours in *Pten/p27* double-mutant mice.

Rb-deficient Mice (Inherited Childhood Retinoblastoma)

Childhood retinoblastoma has both an inherited and sporadic condition component: about 60% of cases are nonhereditary and 40% are inherited. In familial retinoblastoma, young children can develop multifocal bilateral tumours of the retina. These individuals inherit one defective autosomal allele through the germ line and the wild-type allele is lost due to somatic mutation leading to tumorigenesis. Individuals who inherit an inactivated copy of *Rb* gene have an estimated 90% probability of developing retinoblastomas at an early age. *Rb* is also commonly mutated in sporadic retinoblastomas, lung, breast, prostate and bladder cancer.

Rb is a nuclear phosphoprotein, which shares homology with p107 and p130, two other members of the same family of proteins. The Rb protein is an important regulator of the G1 checkpoint. It binds to members of the E2F family of transcription factors needed for the expression of S phase early genes and prevents S phase entry. Hyperphosphorylation of Rb by G1/S cyclin-dependent kinases release E2Fs from Rb-mediated repression and enables the progression of the cell cycle.

In order to study the role of the *Rb* gene in tumour suppression, several groups have inactivated Rb in the mouse germ-line. Rb null mice die at mid-gestation and show aberrant haematopoiesis and severe defects in the nervous system. However, $Rb^{+/-}$ mice and chimaeric mice derived from Rb null ES cells develop tumours of the pituitary gland. In the tumours arising in the $Rb^{+/-}$ mice the *Rb* wild-type allele is lost. Despite the important role of Rb in G1/S transition, the Rb null mice can progress throughout early stages of development. One could postulate that other members of the Rb family of proteins (p107 and p130) might compensate for the lack of Rb. However, p107 and p130 null mice develop normally without any developmental defect. Double-deficient p107 and p130 mutants show neonatal lethality. The $p107/Rb^{-/-}$ embryos show earlier lethality (11.5 d.p.c.) than $Rb^{-/-}$ embryos, which die at mid-gestation (14.5 d.p.c.). These observations suggest that the Rb family of proteins exert overlapping functions as well as distinct functions. This example illustrates the importance of the *in vivo* analysis in the mouse in dissecting the relative contribution of family members to controlling molecular pathways important for cancer pathogenesis.

In order to overcome the embryonic lethality of the Rb null mice and to study the effect of Rb mutation in specific organs/cells, mice carrying a conditional allele of *Rb* (with *loxP*, or Frt) have been generated. Mice carrying *loxP* or Frt modified *Rb* allele were viable and showed no cancer predisposition. By contrast, pituitary expression of the Cre recombinase in mice carrying a conditional allele of *Rb* resulted in the rapid onset of intermediate lobe tumours of the pituitary gland in 100% of the cases. No other abnormalities were observed in these mice. Although this may not be regarded as a relevant mouse model of human cancer, the Rb conditional model system is nevertheless a straightforward example of the utility of conditional gene inactivation for the *in vivo* analysis of the molecular basis of tumorigenesis.

Along the same line, in order to address the mutational requirement of Rb, p107 and p53 in tumorigenesis, mice where the targeted loss of Rb was directed to the photoreceptor cell compartment of the retina were generated using the conditional Cre–*loxP* system. The targeted loss of *Rb* gene was achieved by expressing the Cre under the control of interphotoreceptor retinoid binding protein (IRBP) promoter. IRBP is expressed in both rod and cone cells of the photoreceptor layer as well as in the pineal gland. Mice with conditional loss of Rb were crossed with p107 and p53 null mice to obtain compound mutant mice with all the possible gene combinations. The resulting mice developed two types of tumours: anterior lobe of the pituitary gland and pineal gland tumours. All the tumours underwent Cre-mediated loss of *Rb*. The pineal gland tumours were expected, since Cre expression was directed to the pineal gland. However, the tumours of pituitary glands were unforeseen and are very likely due to the

ectopic expression of the *Cre* transgene. The onset of tumours in both cases was accelerated in the $p53^{+/-}$ mice and the tumours showed loss of the wild-type allele. Pineal gland tumours also developed in $Rb^{+/-}$, $p53^{-/-}$ mice with loss of the wild-type *Rb* allele.

Recently a mouse model for medulloblastomas has been generated using a similar approach: both Rb and p53 have been knocked out through a Cre–*loxP*-mediated approach in the cerebellar external granular layer cells. The resulting mice developed highly aggressive embryonal tumours of the cerebellum with typical features of medulloblastomas. This demonstrates the importance of the loss of Rb function in medulloblastomas and supports the hypothesis that medulloblastomas arise from cells located in the cerebellar external granular layer.

All together, these models underscore the crucial importance of conditional knockout mutants to study the mechanisms of tumorigenesis and in particular the function of tumour-suppressor genes which are required for early developmental stages.

PML-deficient Mice (Acute Promyelocytic Leukaemia)

PML (promyelocytic leukaemia gene) originally termed *myl*, was cloned by virtue of its involvement in t(15;17)(q22;q11.2-q12) associated with acute promyelocytic leukaemia (APL), where *PML* fuses to the retinoic acid receptor alpha gene (*RARα*). PML is the most frequent partner of *RARα* in APL. It is important to note that unlike other tumour-suppressor genes, *PML* was not originally found mutated in cancer. Instead, it is the evidence accumulated *in vitro* and *in vivo* in mouse models that has uncovered the tumour-suppressive role of PML, in turn prompting further studies of this gene in tumours other than APL.

PML belongs to a family of proteins characterized by the presence of a RING-B-box-coiled-coil (RBCC) motif. The PML protein is typically found concentrated in discrete nuclear structures along with other proteins variably named PML nuclear bodies (NB) or PODs (PML oncogenic domains). In APL, PML, together with the other nuclear body components, is delocalized from the NB into aberrant nuclear structures as results of its heterodimerization with the translocation fusion gene product PML–RARα. This observation led to the hypothesis that PML function might be impaired in APL leading to leukaemogenesis. (See also chapter on *Haematological Malignancies in Cancer Research*.)

In order to address this issue, PML knockout ($PML^{-/-}$) mice have been generated. Like p53 null mice, PML null mice were developmentally normal. They succumbed, however, to spontaneous infections that prevented a long-term follow-up for tumour incidence. However, when challenged with tumour-promoting activity such as 12-*O*-tetradecanoylphorbol-13-acetate (TPA) and with carcinogens such as dimethylbenzanthracene (DMBA), $PML^{-/-}$ mice developed skin papillomas and carcinomas as well as B and T lymphomas and histiocytomas at a much higher frequency than in wild-type controls. Therefore, PML can suppress tumorigenesis *in vivo*. In addition, cells of $PML^{-/-}$ mice are protected from multiple caspase-dependent apoptotic signals such as Fas, tumour necrosis factor (TNF), ceramide, interferon (INF) and ionizing radiations, demonstrating an essential role for PML in multiple apoptotic pathways. Furthermore, as observed in cells from Bloom syndrome patients, an increased rate of sister chromatid exchange is observed in PML null cells implicating PML and the PML–NB in the maintenance of genomic stability. In this respect it is important to note that PML and BLM were found to co-localize in the NB. All together, these data indicate that PML might exert its tumour-suppressive activity at multiple levels, by controlling cell growth, cell survival and genomic stability. Based on these findings it may be logical to propose that PML–RARα oncoprotein antagonizes the normal function of PML. In APL, this may be facilitated by the reduction to heterozygosity of the normal *PML* allele. Indeed, the progressive reduction of the dose of *PML* resulted in a dramatic and dose-dependent increase in the incidence of leukaemia, and in an acceleration of leukaemia onset in *PML–RARα* transgenic mice. Thus, *in vivo*, PML is haploinsufficient in antagonizing the leukaemogenic potential of PML–RARα. This also suggests that PML's reduction to heterozygosity as a consequence of its involvement in the t(15;17) translocation may be a critical event in APL pathogenesis. In conclusion, mouse modelling has allowed the identification of a new tumour-suppressor gene, prompting a comprehensive analysis of the role of PML in tumorigenesis.

BIOLOGICAL LIMITATIONS AND TECHNICAL CONSIDERATIONS

As mentioned, in several cases, the mutation of a certain gene in mice does not necessarily give rise to the same tumour observed in humans. A number of explanations have been proposed. The difference might result, for instance, from subtle differences in growth control pathways in mouse and human cells. The susceptibility to loss of the remaining wild-type allele may not be the same in human and murine cells. The shorter life span in mice may affect the tumour susceptibility in certain tissues. Moreover, the high incidence of lymphomas and sarcomas in various mouse models generated may simply reflect the fact that the C57BL/6 strain into which the mutated allele are usually crossed is highly susceptible to these tumour types. Thus, the tumour-suppressor inactivation might exacerbate a pre-existing predisposition. In fact, crossing

the tumour-suppressor mutation into inbred strains of different backgrounds has revealed strain-specific differences in tumour development and spectra. The availability of markers spanning the mouse genome will allow the identification of new tumour-modifier genes and cooperating oncogenes or tumour-suppressor genes. Nevertheless, these differences have to be taken into account when comparing different studies and differences between human and mouse phenotypes. (See also chapter on *Human Tumours in Animal Hosts*.)

Combining multiple cancer-associated mutations or multiple inactivated genes in a single animal has proven to be a very useful tool to study cooperative effects and define pathways involved in tumour pathogenesis; however, the number of mutations that can be introduced in a single animal are limited by Mendelian laws: the introduction of homozygous mutation in more than three genes at the same time requires an enormous amount of interbreeding, rendering the experiment time consuming and expensive. A possible way to overcome this problem would be to introduce several mutations by targeting multiple genes in a single ES cell clone. This approach, however, is complicated by the limited number of selectable markers that can be used at once to enrich for the targeted ES cells. Moreover, ES cells cannot be subjected to multiple rounds of manipulations and passages in culture without losing their totipotency and acquiring karyotypic abnormalities. Along a different line, it has to be remembered that the generation of knockout mice is a very costly operation. The high cost of the generation of mouse models has to be added to the elevated cost of animal maintenance and husbandry. Finally, the phenotypic characterization of the knockout mice and their pathological analysis needs to be performed in an appropriate qualified setting.

Although there are limitations to the large-scale and effective application of these techniques for the study of cancer pathogenesis and other human disease, still further applications of this methodology have been underutilized and will generate in the future exciting findings. As an example, the Cre–*loxP* system can be used to introduce chromosomal rearrangements such as deletions, duplications and inversions into the mouse genome. This is of importance in view of the fact that in many tumours these rearrangements accompany more discrete mutations in tumour-suppressor or oncogenes. Recently, rearrangements encompassing up to three-quarters of chromosome 11 have been introduced in ES cells by this approach. The Cre–*loxP* system can also be used to recreate reciprocal chromosomal translocations in ES cells or *in vivo*, thus providing mouse models of human cancer associated with these genetic lesions. Finally, gene targeting can also be utilized to introduce point mutations and isoform specific deletions that will be critical, in the future, to refine our understanding of gene function and its relationship with the genesis of cancer.

PERSPECTIVES

The development of a second generation of gene-targeting technologies such as conditional or inducible systems has made it possible to introduce in the mouse germ line more subtle gene alterations that can provide functional information beyond that provided by the generation of null alleles. The comparison between the effects of tumour-suppressor losses in different mouse strains will allow the discovery of novel tumour modifiers. The identification of modifier genes may indeed be one of the most definitive ways to determine the critical biochemical pathways responsible for cancer pathogenesis. Standard biochemical studies tell us how genes and gene pathways can intermingle, and have proven and pointed out the intricacy of these cellular pathways. The mouse genetics can be effective in simplifying and thus shedding light on this complexity. There has been an explosive growth of accessible genetic databases and available collections of genomic and expressed gene reagents and rapidly evolving tools for detecting genomic alterations and gene expression patterns. As an example, gene expression profiling by microarrays and protein arrays will allow the determination of the genes or pathways affected in a particular tissue of a knockout mouse. Spectral karyotyping (SKY), multiple fluorescence *in situ* hybridization (M-FISH), multicolor chromosome banding and other labelling techniques and strategies are recent technical developments in the field of molecular cytogenetics. The use of these techniques in mouse models will allow screening for specific chromosomal aberrations. The presence of recurrent abnormalities associated with a specific tumour type will allow the identification of additional genetic events in tumour progression and metastasis. All these novel tools, if applied on a comparative basis on multiple mouse models, will allow the definition of pathways specifically or universally affected as a consequence of oncogenic activities and loss of tumour-suppressor gene functions.

REFERENCES

Di Cristofano, A., *et al*. (1998). PTEN is essential for embryonic development and tumor suppression. *Nature Genetics*, **19**, 348–355.

Di Cristofano, A., *et al*. (2001). Pten and p27^{KIP1} cooperate in prostate cancer tumor suppression. *Nature Genetics*, **27**, 134–135.

Harbour, J. W. and Dean, D. C. (2000). Rb function in cell-cycle regulation and apoptosis. *Nature Cell Biology*, **2**, 65–67.

Lozano, G. and Liu, G. (1998). Mouse models dissect the role of p53 in cancer and development. *Seminars in Cancer Biology*, **8**, 337–344.

Marino, S., *et al*. (2000). Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external

granular layer cells of the cerebellum. *Genes Development*, **14**, 994–1004.

May, P. and May, E. (1999). Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene*, **18**, 7621–7636.

Muller, U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mechanical Development*, **82**, 3–21.

Rego, E. M., *et al*. (2001). The molecular pathogenesis of acute promyelocytic leukemia. *Seminars in Hematology*, **38**, 54–70.

Wang, Z., *et al*. (1998). Role of PML in cell growth and the retinoic acid pathway. *Science*, **279**, 1547–1551.

Wang, Z. G., *et al*. (1998). PML is essential for multiple apoptotic pathways. *Nature Genetics*, **20**, 266–272.

FURTHER READING

Reviews

Cohen-Tannoudji, M. and Babinet, C. (1998). Beyond 'knock-out' mice: new perspectives for the programmed modification of the mammalian genome. *Molecular and Human Reproduction*, **4**, 929–938.

Gao, X., *et al*. (1999). Advanced transgenic and gene-targeting approaches. *Neurochemical Research*, **24**, 1181–1188.

Hooper, M. L. (1998). Tumour suppressor gene mutations in humans and mice: parallels and contrasts. *EMBO Journal*, **17**, 6783–6789.

Muller, U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mechanical Development*, **82**, 3–21.

Osada, J. and Maeda, N. (1998). Preparation of knockout mice. *Methods in Molecular Biology*, **110**, 79–92.

Takeda, J., *et al*. (2000). Conditional gene targeting and its application in the skin. *Journal of Dermatological Science*, **23**, 147–154.

Books

Hogan, B., Beddington, R., Costantini, E. and Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd edn. Cold Spring Harbour Press, NY.

Jackson, I. and Abbott, M. (eds) (2000). *Mouse Genetics and Transgenics: A Practical Approach*. (Oxford University Press, Oxford).

Wasserman, P. M. and De Pamphilis, M. L. (1993). *Methods in Enzymology, Vol. 225: Guide to Techniques in Mouse Development*. Academic Press, San Diego.

Web sites

Transgenic animal web: http://www.med.umich.edu/tamc/esoutline.html.

Internet resources for transgenic and targeted mutation research (The Jackson Laboratory): http://tbase.jax.org/docs/databases.html.

The Cancer Center at Washington University Medical Center: http://medicine.wustl.edu/~escore/es_core.htm.

UCI transgenic mouse facility: http://darwin.bio.uci.edu/~tjf/index.html.

The whole mouse catalog: http://www.rodentia.com/wmc/.

The mouse models of human consortium MMHCC: http://mmhcc.nci.gov/

Chapter 58

Human Tumours in Animal Hosts

Robert Clarke
Georgetown University School of Medicine, Washington, DC, USA

CONTENTS

- Animal Models in Cancer Research
- Principles of Establishment: Issues Relating to Experimental Design
- General Applications
- Biological and Technical Limitations
- Conclusions and Future Prospects

ANIMAL MODELS IN CANCER RESEARCH

Animal models have been used in cancer research for many years. Often these provide the only way to model effectively the complex interactions between host and tumour. For example, while many new technologies are being developed, it remains virtually impossible to predict accurately the pharmacokinetic and toxicological properties of experimental antineoplastic agents from only *in vitro* studies. Assessing the activity of new drugs, or of the full tumorigenic or metastatic properties of cells, also remains beyond the capability of *in vitro* technologies. Thus, for the moment, it remains necessary to use animal models for many studies in cancer research. Nonetheless, it is to be hoped that, at some point in the future, *in vivo* models will be superseded by new technologies that will no longer require the use of live animals.

While the research community currently retains the privilege of using live animals, the longevity of this privilege is uncertain. Animal rights organisations continue to campaign for the elimination of *in vivo* studies. In recent years, many governments have enacted legislation to control more carefully the use of animals in research. Most of these changes have been beneficial and have significantly improved the conditions under which animals are maintained and studied. Further legislation and limitations will almost certainly follow in the coming years. It is imperative that investigators are aware of the legislation to which they are subject. In some cases, such legislation may be less than ideal, such that ensuring the minimum level of regulatory compliance may not be optimal. Ultimately, if the research community is to maintain an adequate ability to use *in vivo* models, at least until viable and validated alternatives become available, its members must be seen to use animals rationally and responsibly.

The use of live animals has clear ethical and scientific implications. Many research institutions are required to establish an oversight body to regulate the use of animals by its members. Investigators must justify the use of animals, their choice of species, strain and number of animals, and to document their search for alternate *in vitro* systems. Where surgical or other invasive techniques are required, individuals must demonstrate that they are adequately trained/experienced in the procedures, and provide for appropriate anaesthesia and analgesia. Some investigators using animals are required to obtain a specific licence from their government or other oversight body.

The precise definition of what constitutes an 'appropriate use of animals' varies considerably among investigators and across international boundaries. Many of the differences are cultural. However, for scientists some issues should remain constant. These include a high level of awareness, and an unrelenting commitment to the health and welfare of animals under their care. Such issues are not only ethically valid, but are demanded by the nature of science. An appropriate environment, and the provision of high-quality animal care, are central to the performance of reliable and reproducible science. Quality of diet, environmental hygiene, density of animal housing, health monitoring of animals and various other concerns, which are well beyond the scope of this article, are some examples (Schiffer, 1997). Specific requirements vary among species. There are many excellent sources for specific information on animal care and welfare (see Further Reading and Websites).

Ultimately, animal models represent the most complex tool available for research scientists. They can provide information and knowledge well beyond that produced by some of our most complex machinery. As such, at the very least, the animals in our care deserve a greater degree of consideration than the most expensive and complex piece of equipment designed and built by humans.

Biology of Graft Models

The grafting of tissues into another host may be across species (xenograft) or to another animal of the same species (allograft). Immunocompromised rodents will maintain both allografts and xenografts. While tumour cells are more frequently transplanted, some normal tissue xenografts also can be successfully initiated.

Generally, the biology of xenografts is comparable to that seen *in vitro* or in the original host. Continual passage of some xenografts can produce cells better acclimated to their new host, and take rates may increase and/or latency decrease. Most xenografted cell lines exhibit comparable genetic stability *in vivo* as they do *in vitro* in the absence of any specific selective pressure. The removal of endocrine stimuli (Clarke *et al*., 1989), immunological effectors or the imposition of other selective pressures, e.g. drug treatments, can produce phenotypic, and probably epigenetic, changes in some xenografts. This has proven to be an effective way to generate variant cell lines with desired phenotypes, so they can be compared with wild-type cells to identify the genes and proteins conferring the variant phenotype.

Given the diversity and complexity of human tumour xenografts, it is not possible to discuss adequately all of the issues related to their development and use in cancer research. This chapter will focus on several general topics related to experimental design, including choice of model, number of animals required and assays for tumorigenicity, metastasis and the screening of experimental antineoplastic agents. The major immunodeficient rodent models are described, since these are the hosts for the xenografts.

PRINCIPLES OF ESTABLISHMENT: ISSUES RELATING TO EXPERIMENTAL DESIGN

Xenografts can be initiated from suspensions of cells grown *in vitro* or from effusions in patients, or from tissue pieces obtained directly from patients or other animals. The size of the tumour piece, or the number of cells inoculated, varies with the characteristics of the graft cells/tissues. Xenografts from many solid tumours can be initiated or maintained by placing pieces as small as 1 mm^3 into the appropriate site in the recipient animal. As a general guide, 1×10^6–5×10^6 cells provide a reasonable inoculum if the cells are being transferred from cells proliferating *in vitro*. Investigators should determine the optimal inoculum for their chosen model either from the literature or by experimentation.

It is also important to know the biological characteristics of the model being used. For example, what is the anticipated tumour doubling time (T_D), are the growth properties adequately reproducible, is the tumour sufficiently invasive or metastatic, or might extension of the primary tumour cause neurological problems such that the animals can no longer obtain food and water? Rapidly proliferating or highly metastatic tumours could create health problems before the desired experimental endpoints are reached. For many established models, the critical biological characteristics are well known and should be widely reported. Where these are not known, a small pilot study in a few animals can help reduce total animal use and save considerable resources (see the next sub-section).

Generating an appropriate experimental design is among the most important aspects of working with xenografts and other animal models. As with all experiments, careful consideration of controls and experimental groups are essential. However, other issues also are important, including statistical power of the study design, choice of model, selection of recipient and site of inoculation. These issues are briefly discussed in the following sections.

Animal Numbers and Statistical Power

One of the more common errors in designing animal studies is the tendency of some investigators to leave issues of data analysis until the experiment is completed and the data have been collected. Unfortunately, only then is it realized that the study was underpowered, or that the statistical analysis of the data is not straightforward. Considerable resources and animals can be spared by consultation with a biostatistician prior to starting an *in vivo* study. Clearly, this has substantial ethical, scientific and economic advantages.

The number of animals required for hypothesis testing will reflect the characteristics of the model, and particularly the magnitude of the expected response. Essentially, the ability of a study design to detect accurately differences among experimental/control groups is its statistical power. Studies are usually considered to be adequately powered when they have an 80% chance or better of detecting a true difference with a $p < 0.05$. An underpowered study may produce an inconclusive result, and perhaps require additional studies that ultimately use more animals than a single but appropriately powered experiment. Alternatively, an overpowered study may use more animals than necessary to test the hypothesis, while identifying statistically significant differences of uncertain or questionable biological relevance.

Some treatments or procedures result in the death of animals before the desired endpoints are reached. Thus, power can be lost during the performance of the study. This also needs to be considered during the initial experimental design. The censoring of data occurring from loss of animals also has implications for the statistical approaches used to analyze the data (Hanfelt, 1997).

In xenograft studies, it may be tempting to try to increase power by inoculating more sites per animal. Although this can increase power, the issue is complicated. There is a potential for multiple tumours in one animal to behave more like each other than like tumours in other animals of the same experimental group. Thus, a doubling of the number of sites may not produce a doubling in statistical power. Readers are directed to an excellent review of this, and other issues of statistical analysis and experimental design relevant to animal model studies, by Hanfelt (1997).

Choice of Model

Several aspects of experimental design impact upon the choice of model. Perhaps the most important is the ability of the model to facilitate testing of the investigator's hypothesis. Animal models represent specific phenotypic/genetic characteristics that occur in a potentially comparable human condition. No matter how well an experiment may be designed, if the model does not adequately reflect the critical aspects of the human disease, any resulting data may have little relevance and the hypothesis may remain untested. Thus, the utility/applicability of a model depends on how well it reflects the most appropriate characteristics of the original disease/disease process.

Perhaps one of the first questions cancer researchers should ask during the experimental design phase is whether their choice of model(s) will clearly allow them to address their hypothesis in a way that has some relevance for the human disease. One possible example of when this tenet was inadequately followed was the use, for many years, of the murine P388 and L1210 lymphocytic leukaemias to screen new antineoplastic agents. Many drugs active in this *in vivo* screen showed limited activity against solid cancers in humans. Following considerable criticism and debate, the National Cancer Institute (Bethesda, MD, USA) decided to instigate a broader initial screen using a panel of cell lines with comprising representatives of each of the major cancers. Many of these cell lines also are useful as xenografts, allowing agents active *in vitro* to be subsequently evaluated *in vivo* in the same models. To the extent that the selected cell lines for each cancer adequately reflect key aspects of that cancer's biology, this approach should allow for the identification of potent agents with even limited applicability. Thus, one potential advantage is that agents with specific patterns of activity, perhaps active against some but not all cancers, may be identified.

The P388 and L1210 models, while they remain useful and widely used, provide another example of issues relating to model selection. Many cytotoxic agents are designed to be, or are expected to be, more active against proliferating rather than resting cells. Drugs active against a leukaemia with a short tumour T_D, high growth fraction and relatively short cell cycle time are expected to appear less potent in solid tumour models with a long T_D, small growth fraction and long cell cycle time. Thus, screening against the leukaemias could identify false positives in the context of solid tumours. Where specific pharmacological or kinetic requirements of drug action are known, clearly these should be effectively reflected in the biologies of the models selected for *in vivo* evaluation.

It is not always necessary, and it may be the exception rather than the rule, that a single model reflects all the characteristics of the original disease. For example, the divergent characteristics of breast tumours provide important clues to their biology and how they should be treated. Some breast tumours express receptors for the steroid hormones oestradiol and progesterone. These tumours respond well to antioestrogens and other endocrine therapies. In marked contrast, tumours that do not express these receptors rarely respond to endocrine manipulations. While tumours that respond to antioestrogens can still metastasize in patients, most antioestrogen-responsive human breast cancer xenografts are poorly metastatic or non-metastatic. These models are not invalid because of their poor metastatic potential. Rather, it limits the aspects of the human disease that they represent and, consequently, the hypotheses they can be used to test.

When selecting experimental *in vivo* models, it is often apparent that no single model may adequately reflect all aspects of the human disease required for testing an investigator's hypothesis. Thus, several models may be required. The choice to use several models requires careful evaluation, since there is a need to reduce the number of animals used in research. Cost also may be an issue. However, an inadequately tested hypothesis in a small study could lead to a need for more animals than a carefully designed, larger single study. Clearly, a realistic appreciation of each model's advantages and limitations, require careful consideration.

Immune-deficient Rodent Models as Xenograft Recipients

Although there are more than 30 loci at which mutations produce an immunocompromised phenotype, relatively few are widely used in cancer research. The most familiar is the nude mouse, homozygous for the *nu* mutation on mouse chromosome 11. Originally identified in Glasgow, Scotland, in 1962, the predominating immunodeficiency was described shortly thereafter. The first report of the ability of these mice to sustain human tumour xenografts was published by Rygaard and Povlsen in 1969. Another widely used single mutation immunodeficient model is the severe combined immunodeficiency (SCID) mutation. Several combined mutation models also are popular, and recent years have seen the emergence of additional immunodeficient models. A brief description of

the major characteristics of selected models follows. Several have been reviewed in detail elsewhere (Shultz, 1989; Clarke, 1996).

The different immunobiologies of these other rodent models provide viable alternatives to the nude mouse. However, it remains controversial whether there are substantial differences in their respective abilities to maintain xenografts (Fodstad, 1991). Some investigators find SCID or *bg/nu/xid* mice to have a reproducibly higher overall take rate than *nu/nu* mice, but many find these models essentially equivalent. Graft take rates may vary by cancer type. Several mouse models reject leukaemic cell line xenografts, but the further immunosuppression produced by gamma-irradiation can be beneficial in helping to alleviate rejection (Ghetie *et al*., 1996). However, the effects can be highly strain/mutation deficient. Irradiation of SCID mice can overcome the defective V(D)J recombination (Livak *et al*., 1996). While this can increase B-cell lymphoma tumour take, metastasis and response to immunotoxins, the response of xenografts to chemotherapy is reduced (Ghetie *et al*., 1996). Thus, the respective immunobiologies may have significant implications for the choice of model and strain. A general representation of the major immunedeficiencies of several common models is shown in **Figure 1** (adapted from Clarke, 1996).

Besides maintaining neoplastic cells, immunocompromised rodents also can allow the study of normal tissues, the effects of therapies on normal cell function or the events associated with virally or chemically induced carcinogenesis. For example, several models can be 'reconstituted' with human cells. Reconstitution of SCID mice with normal peripheral blood lymphocytes from Epstein–Barr-positive human donors develop fatal B cell lymphomas. Human skin grafts in SCID mice can model the vascular leak syndrome seen in some human immunotoxin trials.

Nude (nu/nu) Mice

More than 70 congenic strains of nude mice have been generated, and strain-dependent differences in their respective immunologies and abilities to support xenografts have been described (Maruo *et al*., 1982; Clarke, 1996). The mice are often called athymic nude mice, although a rudimentary thymus is present in some individuals. B cell maturation is defective but normal virgin B cells are present. T-lymphocyte levels are very low

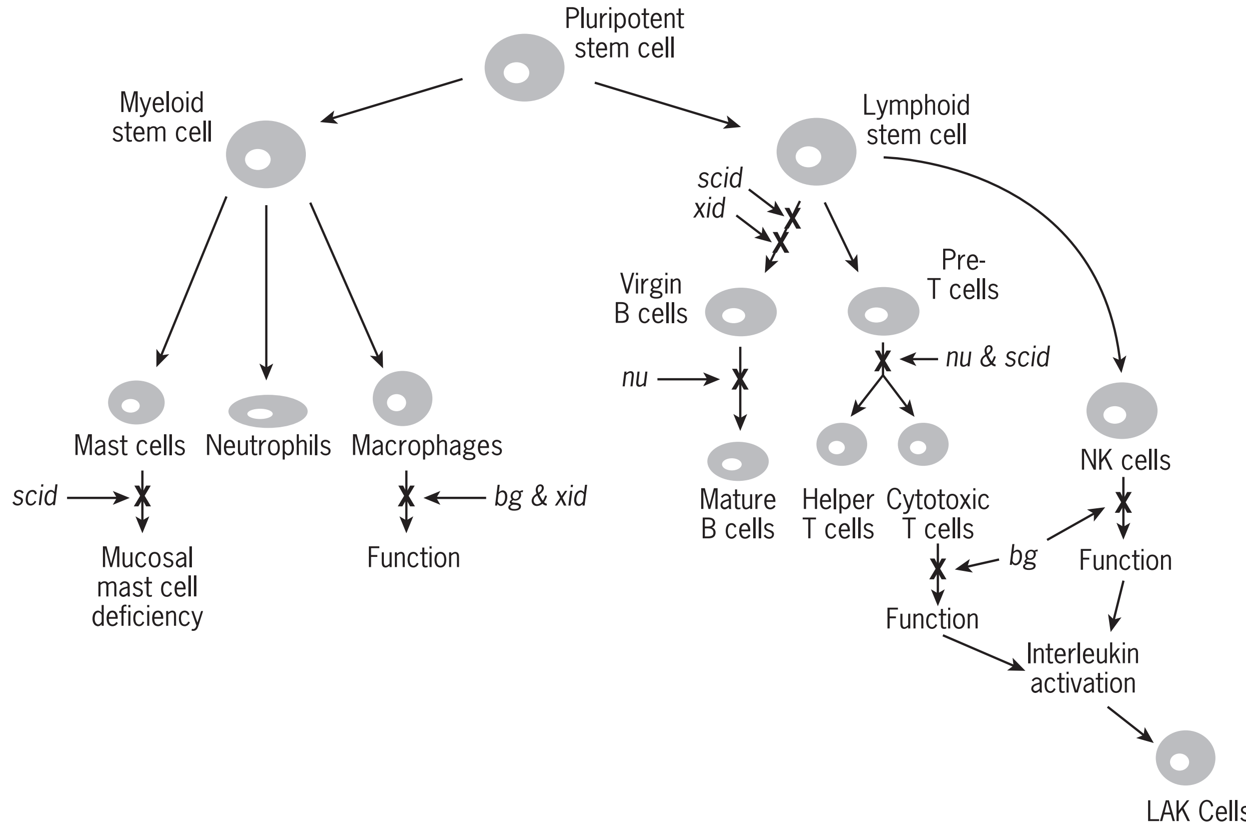


Figure 1 Immunobiologies of several rodent models. (Adapted from Clarke, 1996.) This is a general representation. The degree to which each cell type is affected varies by mutation and often by individual animal. The location of the defects is not intended to be precise, rather it should provide a simple means to assess the likely changes in different models.

or undetectable in most strains, as are responses to T cell-dependent antigens. Substantial immune function can be restored by reconstitution with T cells. Nonetheless, serum IgM levels are similar to, or higher than, those in *nu/+* mice. Occasionally, circulating IgG and IgA are detected in individual animals.

While humoral immunity is compromised, nude mice retain several key mediators of cell-mediated immunity. T cell-independent antigen responses are often normal, and the levels of natural killer (NK) cell activity are greater than that in normal and *nu/+* mice. There is no apparent defect in tumoricidal macrophages, and nude mice can produce lymphokine activated killer (LAK) cell numbers comparable to +/+ mice. Thus, nude mice are relatively robust and easier to maintain than some of the more severely immunocompromised rodent models.

The high NK cell activity in nude mice may contribute to the poor take rate of some primary human xenografts. This can be increased by embedding the transplant tissues in an artificial basement membrane (Fridman *et al*., 1991), which probably protects the tumours from lysis by cell mediated effector cells, and provides a structural/dimensional environment containing attachment molecules and mitogens (Clarke, 1996).

While mice are the most common species in which the nude mutation is used, a comparable mutation also exists in rats. The rat nude mutation (*rnu*) produces athymic nude rats, which appear to have an immunobiology broadly comparable to that of nude mice (Hougen and Klausen, 1984). Nude rats also will support tumour xenografts, but are more expensive to maintain. However, the size of the rat may make these a useful alternative, particularly where orthotopic transplantation or other surgical studies are difficult or impractical in mice.

Combined Beige (bg)/nu Mutations

To eliminate the higher NK cell activity seen in nude mice, double congenic *bg/nu* mice have been generated (Karre *et al*., 1980). The *bg* mutation, located on mouse chromosome 13, produces a reduced synthesis of pigment granules in melanocytes which results in a light coat colour. Analogous to the multisystem autosomal recessive disease Chédiak–Higashi syndrome in humans, the main immune-deficiency in *bg/bg* mice is a block in NK function. Other immunodeficiencies produced include an impaired production of cytotoxic T cells and defects in macrophages and granulocytes.

The *bg* mutation significantly reduces NK activity when introduced along with the nude mutation. However, residual NK activity is detectable in *bg/nu* mice, which is clearly higher than that seen in *bg/bg* mice. A further contribution of the *bg* mutation is a clotting disorder, seen in both *bg/bg* and *bg/nu* mice. This can be problematic when the experimental design requires surgery, e.g. for orthotopic implantation (Clarke, 1996).

Combined bg/nu/xid Mutations

A further reduction in immune competence was obtained by adding the X-linked immunodeficiency (*xid*) to that of the *nu/bg* genotype (Andriole *et al*., 1985). The major contribution of the *xid* mutation to the phenotype is an impaired development of B cells. B cell colonies are not detected in *in vitro* assays, probably because of an inability of otherwise normal immature B cells to respond to early activation signals. Immunoglobulin levels, particularly IgM and IgG_3, are low. Mice bearing the *bg/nu/xid* mutations have characteristics of each of the individual mutations. Thus, they are essentially athymic, have intermediate NK cell activities, low numbers of LAK cells, and have defects in both B and T cell maturation. While more immunodeficient than the single or double congenic strains, the clotting disorder conferred by the *bg* mutation also is apparent (Clarke, 1996).

Severe Combined Immune Deficiency (SCID) Mutation

The *scid* locus on mouse chromosome 16 is equivalent to the human locus at chromosome 8q11. The *scid* mutation disrupts the process of rearrangement of genes encoding antigen-specific receptors on B and T cells (Bosma and Carroll, 1991). This reflects an inability to effectively join the cleaved variable region segments catalysed by the immunoglobulin V(D)J recombinase. Affected mice have small lymphoid organs and defective differentiation/maturation of lymphocytes. While the myeloid lineage cells are normal, pre-B and B cells cannot be detected and remaining T cells are defective. IgG 2a, 2b and 3a, IgM and IgA are rarely detected. However, individual mice may produce detectable levels of two or more IgG isotypes and/or IgM. There are no major defects in macrophages, NK cells and LAK cells but a lack of specific lymphokines required for their development from progenitors can produce a deficiency in mucosal mast cells.

Some earlier SCID models were described as ‘leaky.’ These mice often produce detectable numbers of functional B and T cells. In many SCID mouse strains, individuals become ‘leaky’ with age, i.e. by 10–14 months of age. As with most immunocompromised rodent models, there can be some variability in tumour take rates among SCID mice of different backgrounds and from different vendors.

Other Immunodeficient Models

In recent years, other gene mutations have been shown to affect immunity in rodents. Their use as xenograft hosts, relative to the more established models, remains to be firmly established. However, it seems likely that additional models will continue to emerge. For example, null mutations in the *RAG-1* and *RAG-2* genes in humans produce a severe immunodeficiency characterized by lack of mature

B and T cells (Omenn syndrome). The deficiency is a result of a loss of V(D)J recombination activity, and comparable mutations in these genes in mice produce immunocompromised animals. These mice do not become leaky with age. Mice bearing the non-obese diabetes mutation (NOD) also exhibit an unusual T cell ontogeny. When both the NOD and *scid* mutations are present, the mice develop thymic lymphomas. While the mice have a relatively short life span, when a strain that does not exhibit reintegrations with the Emv30 retrovirus is used, the frequency of large lymphomas is reduced. More recently, NOD/LtSz-Rag1^{null} mice have been generated. These mice have an increased life span, reflecting the later onset of lymphomas, have no mature B or T cells, and exhibit low levels of NK cell activity. These mice will support grafts of human haematopoietic stem cells and lymphoid cells. Engrafted T cells can be infected with the HIV virus (Shultz *et al*., 2000).

Sites of Inoculation and Assays for Tumorigenesis and Metastasis

Important experimental endpoints include determining factors that influence the tumorigenic or metastatic potential of cells. These also are important characteristics of new cell line models. The metastatic cascade is a complex process, requiring cells to invade, migrate, intravasate, survive in the circulation, extravasate and, finally, to colonize distant sites successfully. While some aspects can be modelled *in vitro,* e.g. invasiveness and motility, several different assays have been used to measure metastasis *in vivo*. Direct inoculation of tumour cells into the circulation is among the more common approaches. The route of systemic inoculation can affect the likely pattern of metastasis. Intracardiac injection increases the probability of obtaining bone metastases (Thompson *et al*., 1999). This can be important for studying cancers such as breast cancer, where bone metastasis is a common feature of the human disease.

Systemic inoculation does not fully replicate the entire metastatic process, e.g. the initial steps of local invasion, migration and intravasation are not required. These approaches are often referred to as representing experimental metastasis. More rigorous estimates of metastatic potential, at least for solid tumour models, require the demonstration of an ability to metastasize from a solid tumour.

The site of inoculation becomes an important issue in assessing metastatic or tumorigenic potential. Despite the widespread use of subcutaneous inoculations, there is substantial and compelling evidence that the biology of grafts at orthotopic sites (site of the original tumour) can be more biologically relevant (Morikawa *et al*., 1988).

Perhaps the most widely used model for studying metastasis has been the B16 mouse melanoma cell line, which forms pigmented metastases in several tissues. Most non-melanoma models do not produce such easily detectable metastases. However, cells can be tagged with expressed genes producing marker proteins, e.g. enhanced green fluorescent protein, β-galactosidase. Tumours tagged with these genes should be readily detected, but for the models to be useful, their phenotypes also should not be markedly affected by high levels of marker gene expression.

For some assays there are practical restrictions on the site of inoculation. Ease and accuracy of solid tumour measurement are critical and can restrict inoculation sites. An immune-privileged site may be required. The sub-renal capsule has been used successfully. Nonetheless, where possible, orthotopic transplantation is preferred for most studies. Tumour take also can be increased by embedding cells/tissues in an artificial basement membrane (Matrigel; Collaborative Research, Bedford, MA, USA) (Fridman *et al*., 1991).

An individual cell line may be tumorigenic or metastatic in one immune-deficient model and not in another. This might occur most frequently when assessing metastatic potential but variability among cell lines for tumorigenicity endpoints also may occur. Where *in vivo* growth is assessed as part of the characterization of a new cell line, more than one model should be used. This is more important when a cell line appears non-tumorigenic or is not metastatic. Thus, the designation of a cell line as non-tumorigenic or non-metastatic should indicate in which model(s) these endpoints were assessed.

GENERAL APPLICATIONS

Xenograft models have been widely used in many cancer research-based studies. As indicated above, such studies can allow for evaluations on endpoints of tumorigenesis and metastasis, e.g. in the characterization of new experimental models. One of the most common uses of human tumour xenografts is for the evaluation of the antineoplastic activity of novel therapeutic strategies. This can include irradiation regimens, cytotoxic, endocrine or gene therapies, or nutritional studies. For drug-based therapies, xenograft models can be used to determine pharmacokinetics, toxicology and activity. In principle, investigators perform in animals the equivalent of a clinical trial in humans. The aims include the prediction of safe and appropriate starting doses, and the identification of potential toxicities, for subsequent human trials. When successfully designed and implicated, these animal studies should significantly reduce human suffering without excessive and unnecessary burden to the animals. The following sections will briefly describe some of the more important issues relating mostly to drug-based studies.

Screening of Experimental Agents *In Vivo*: Dose Selection and Toxicology

The determination of safety and antineoplastic activity in animal models is effectively a prerequisite for the preclinical evaluation of experimental agents. Determination of the starting dose or dose range is an important consideration. Since the presence of tumours can often alter toxicity, it is not uncommon for rodent toxicological studies to be performed in tumour-bearing mice, although this is not essential. This can be useful where subsequent studies are to be performed at the estimated maximum tolerated dose (MTD).

The issues of toxicology and establishing a starting dose for *in vivo* studies are closely related. Many such studies are based on estimations of activity at or below a drug's MTD. To estimate the MTD, one simple approach is to do a pilot dose range finding study. A broad range of doses is evaluated using only one or two animals per dose. This limits the number of animals exposed to toxic/lethal doses. A more definitive subsequent study can then be done based on the data from the pilot analysis, and a MTD estimated, if necessary (UKCCCR, 1998; Clarke, 1997a). In such toxicological studies, change in animal body weights is often used as an indicator of toxicity. While a loss of body weight is frequently used as the endpoint, a more sensitive measure is the detection of a reduction in the rate of weight gain in young animals. A loss of 10% body weight should be an adequate estimate of toxicity in most studies using adult animals.

Some drugs will produce a severe but transient toxicity. Drug-related deaths that occur within several hours of administration should not be used to establish the MTD. These toxicities often reflect sensitivity to the peak plasma concentration resulting from the bolus. The vehicle for drug delivery also may be an important contributor to toxicity. The volume in which the drug is delivered will vary with route of administration, but should be the smallest volume practicable for reproducibility of administration. More detailed discussions of toxicologic studies can be found elsewhere (LoRusso *et al*., 1990; Clarke, 1997a; UKCCCR, 1998).

Screening of Experimental Agents *In Vivo*: Endpoints for Assessing Activity

Determining the *in vivo* activity of experimental agents is done using a variety of endpoints. These are generally either excision assays, where the tumours are treated *in vivo* and the response measured either *ex vivo* following transplantation into a new untreated host; or in *in situ* assays, where the tumours are both treated and their response measured *in vivo* in the same host (Clarke, 1997a). Excision assays remove the contribution of reproductively dead cells to tumour volume, but the tumour microenvironment is lost if activity is assessed *in vitro*. *In situ* assays also can be complicated where the kinetics of regrowth are variable or unpredictable. Nonetheless, these are generally preferred to excision assays.

The *in situ* techniques should be performed on proliferating tumours. Growth delay is commonly assessed, and is defined as the time difference between the tumours in treated and control groups to reach a predetermined size (Clarke, 1997a). Since tumours are of the same size, the data are not confounded by the gompertzian or exponential–quadratic kinetics involved (Hanfelt, 1997). For growth delay assays, a predetermined endpoint size is required. This should be sufficient to produce tumours that are easily measured, but not so large that they include areas of substantial hypoxia or necrosis. Generally, the best endpoint size is close to the size at treatment, e.g. twice the treatment volume, which reduces any effects of treatment on growth rate. Data can be analysed as specific growth delay, total cell kill, cell kill/dose or net cell kill (Clarke, 1997a).

Tumour growth delay estimates often compare treated (T)/control (C) or $T-C$ values, with a T/C of $\leq 42\%$ usually taken as indicating activity. Whether these are based on mean or median values, longitudinal analyses may be a more effective use of the data. Other approaches for statistically exploring growth delay data are described elsewhere (Clarke, 1997a; Hanfelt, 1997)

Overall survival also has been widely used as an endpoint for antineoplastic activity. When the lethal tumour burden of the model is well established, e.g. as with the P388 and L1210 growing as ascites, survival estimates are reliable. However, many solid tumours exhibit a poor or unpredictable metastatic potential. Tumour burden as a percentage of total body weight can readily exceed that observed in humans. When drugs are administered intravenously and the P388/L1210 models are growing as ascites, a percentage increased life span (%ILS) of $>40\%$ is often used as an estimate of adequate activity. %ILS is estimated as

$$\%ILS = 100(T - C)/C$$

Where there is heterogeneity in survival data, this may be better estimated as previously described:

$$\%ILS = 100(HR - 1)$$

where HR is the hazard ratio obtained by Cox proportion hazards regression analyses (Clarke, 1997a).

While survival may appear the best defined endpoint, and is the definitive endpoint in many human trials, few institutions now allow death to be used routinely as an endpoint without considerable justification. This is not surprising, given the discomfort and suffering often associated with the survival endpoint. It is difficult to justify for solid tumours, since most can be used in growth delay or excision assays.

Where possible, survival as an endpoint could be restricted to ascites models. It also may be possible to substitute morbidity for death. This requires knowledge of the time from the onset of morbidity to death, and evidence that this period be consistent. To allow others to use the same endpoints and models, the criteria used to define morbidity, and the verification of its applicability as a surrogate for death, should be reported (Clarke, 1997a).

When the endpoints are 'cure' or the proportion of 'long-term survivors,' the time point at which 'cure' or 'survival' is attributed must be defined. Even a rapidly proliferating tumour with $T_D = 48$ h could require up to 16 weeks post-treatment to establish 'cure.' The proportion of long-term survivors is also often reported. A reasonable definition of this endpoint is the proportion of animals that survive to three times the mean or median survival of the appropriate control groups.

Studies on Diet and Cancer

It is now evident that diet and nutrition may play an important role in the development, progression and management of many cancers. Interest in dietary supplements, nutrients and micronutrients continues to increase, as does the need adequately to identify and determine their importance and place in cancer prevention and therapy. Many such studies require the feeding of complex but controlled diets to rodents, in a manner that allows for appropriate hypothesis testing (Clarke, 1997b).

Unfortunately, the regular rodent chow diets are not controlled in terms of a reproducible source for nutrients and micronutrients. Thus, when performing nutritional studies, or studies on the effects of specific nutritional components, it is necessary to use an appropriately formulated, semipurified diet. While there are many choices available, those formulated by the American Institute of Nutrition's *Ad Hoc* Committees on Standards for Nutritional Studies in Rodents have become widely accepted. The modified AIN-76A diet has been updated and reformulated to reflect better the varying needs of the animals at different stages of growth. The resulting AIN-93G, designed for growth, pregnancy and lactation, and AIN-93M, designed for the maintenance of adult rats (Reeves *et al.*, 1993), also are more appropriate for long-term carcinogenesis studies.

Many laboratory chows also contain high isoflavone concentrations, and could contribute to some of the variations in tumorigenesis reported among laboratories (Thigpen *et al.*, 1999). Perhaps the most widely studied of these isoflavones is genistein, which is reported to have antineoplastic activity in several cell line and xenograft models, including mammary, liver, prostate and colon cancer models. It also can act as an antioxidant and is an inhibitor of protein tyrosine kinases, angiogenesis and topoisomerase activity. Effects on serum cholesterol and bone mineral density in rodents also have been reported (Thigpen *et al.*, 1999). These activities could confound data interpretation in endocrine, antineoplastic activity and other studies. Where this is a major concern, or may be suspected of contributing to outcome, investigators may need to consider using a semipurified diet, such as the appropriate AIN-93 modification.

BIOLOGICAL AND TECHNICAL LIMITATIONS

There are several limitations to the use of xenograft models. Some ethical considerations have already been alluded to above. Clearly, there should be restrictions on the use of live animals, particularly where this use may involve prolonged discomfort, pain and/or death. A detailed discussion of the major issues is beyond the scope of this chapter, but some should be self-evident. For example, the use of animal models must be carefully considered and justified, no viable alternatives should be available, the minimum number of animals required for adequate hypothesis testing should be used and every precaution should be taken to eliminate/minimize pain or discomfort. All animals should receive the highest quality of humane care by appropriately trained individuals.

While the *in vivo* environment may be more physiologically relevant than growth *in vitro*, it cannot eliminate many of the problems associated with the use of human cell lines. For example, some cell lines are genetically or phenotypically unstable. The adaptation for growth *in vitro,* or many years of culture on plastic, may produce changes that do not occur in tumours arising in patients.

For grafts to be successful, the hosts must be severely immunocompromised. The immunobiology of the recipient hosts may be very different from the original human host. Consequently, some immune effects may be missed and the consequences of others exaggerated by the immunity present in the animal model.

While metastasis is a critical component of the biology of cancer in humans, relatively few xenografts accurately reproduce the frequency and distribution of metastases. Whether this is a problem with the animal hosts or the cells/tissue grafted is not always clear.

There are other obvious potential problems associated with the use of non-human species to attempt to predict the safety and activity of an experimental agent or regimen in humans. For example, the pharmacokinetics of a drug in rodents may be very different to that in humans. Generally, it is necessary to adjust for interspecies differences, particularly in their metabolic rates. One approach is simply to correct for body weight, administering the same amount of drug on a milligrams of drug per kilogram body weight basis. However, this is often considered inadequate, with estimates based on body surface area (milligrams of drug

per square metre) providing more comparable data. The approximate correction values for milligrams per kilogram, based on surface area, have been estimated as (Freireich *et al*., 1966)

mouse $\rightarrow$ rat (rat dose = mouse dose in mg kg^{-1} $\times$ 0.5)

rat $\rightarrow$ human (human dose = rat dose in mg kg^{-1} $\times$ 0.14)

Starting doses for human trials are often selected based on one-tenth MTD in rodents as the approximate starting dose for a Phase I trial in humans. Using a dose that produces comparable pharmacokinetics can increase the predictability of the mouse xenograft-to-human tumour model. However, in a recent study of 25 drugs, a starting dose of one-tenth the MTD in mice was shown to be safe in humans. Rodent toxicological studies can accurately predicted the dose-limiting toxicity. Toxicological studies in mice also produced better estimates than those in rats (Newell *et al*., 1999). Thus, carefully designed studies in mice can produce data that have direct relevance for the design of Phase I clinical trials.

Other limitations or concerns may relate to endocrinological investigations. Hormonal supplementation of rodents has been widely used to study the regulation of tumorigenesis, e.g. in breast and prostate cancers. Several endocrine agents are known to influence the cell-mediated immunity effector cells present in many immune-compromised rodent models. Oestrogens produce a biphasic effect on NK cell activity, and tamoxifen can induce NK cell activity. Oestrogens also may alter B cell function, increase IgM secretion and inhibit T-suppressor function and T-helper maturation. Medroxyprogesterone acetate alters the T4+/T8+ ratio, whereas lynoestrenol stimulates active T rosetting and phagocytosis by monocytes. Thus, the effects of endocrine treatments on immune function may require special attention in some study designs, and could provide a limitation in data analysis and interpretation.

CONCLUSIONS AND FUTURE PROSPECTS

Despite their limitations, graft models of human cancers have provided considerable information. Data from these studies have significantly reduced the risks to humans in Phase 1 clinical trials and provided invaluable information in the design and testing of new cancer therapies. For the moment, there are no viable alternatives for many of these types of studies. Consequently, investigators must strive to refine, reduce and replace the use of animal models. At some point in the future, this should lead to a time when there will no longer be either a need, or a place, for the use of live animals in cancer research.

REFERENCES

Andriole, G. L., *et al*. (1985). Evidence that lymphokine-activated killer cells and natural killer cells are distinct based on an analysis of congenitally immunodeficient mice. *Journal of Immunology*, **135**, 2911–2913.

Bosma, M. J. and Carroll, A. M. (1991). The SCID mouse mutant: definition, characterization, and potential uses. *Annual Review of Immunology*, **9**, 323–350.

Clarke, R. (1996). Human breast cancer cell line xenografts as models of breast cancer: the immunobiologies of recipient mice and the characteristics of several tumorigenic cell lines. *Breast Cancer Research and Treatment*, **39**, 69–86.

Clarke, R. (1997a). Issues in experimental design and endpoint analysis in the study of experimental cytotoxic agents *in vivo* in breast cancer and other models. *Breast Cancer Research and Treatment*, **46**, 255–278.

Clarke, R. (1997b). Animal models of breast cancer: experimental design and their use in nutrition and psychosocial research. *Breast Cancer Research and Treatment*, **46**, 117–133.

Clarke, R., *et al*. (1989). Progression from hormone dependent to hormone independent growth in MCF-7 human breast cancer cells. *Proceedings of the National Academy of Sciences of the USA*, **86**, 3649–3653.

Fodstad, O. (1991). Tumorigenicity and dissemination of human tumors in congenitally immune-deficient mice. *Journal of the National Cancer Institute*, **83**, 1419–1420.

Freireich, E. J., *et al*. (1966). Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemotherapy Reports*, **50**, 219–244.

Fridman, R., *et al*. (1991). Enhanced tumor growth of both primary and established human and murine tumor cells in athymic mice after coinjection with Matrigel. *Journal of the National Cancer Institute*, **83**, 769–774.

Ghetie, M. A., *et al*. (1996). Effect of sublethal irradiation of SCID mice on growth of B-cell lymphoma xenografts and on efficacy of chemotherapy and/or immunotoxin therapy. *Laboratory Animal Science*, **46**: 305–309.

Hanfelt, J. (1997). Statistical approachs to experimental design and data analysis of *in vivo* studies. *Breast Cancer Research and Treatment*, **46**, 279–302.

Hougen, H. P. and Klausen, B. (1984). Effects of homozygosity of the nude (rnu) gene in an inbred strain of rats: studies of lymphoid and non-lymphoid organs in different age groups of nude rats of LEW background at a stage in the gene transfer. *Laboratory Animals*, **18**, 7–14.

Karre, K., *et al*. (1980). *In vitro* NK-activity and *in vivo* resistance to leukemia: studies of beige, beige/nude and wild type hosts on C57BL background. *International Journal of Cancer*, **26**, 789–797.

Livak, F., *et al*. (1996). Transient restoration of gene rearrangement at multiple T cell receptor loci in gamma-irradiated scid mice. *Journal of Experimental Medicine*, **184**, 419–428.

LoRusso, P., *et al*. (1990). Antitumor efficacy of PD115934 (NSC 366140) against solid tumours of mice. *Cancer Research*, **50**, 4900–4905.

Maruo, K., *et al*. (1982). Strain-dependent growth of a human carcinoma in nude mice with different genetic backgrounds: selection of nude mouse strains useful for anticancer agent screening system. *Experimental Cell Biology*, **50**, 115–119.

Morikawa, K., *et al*. (1988). Influence of organ environment on the growth, selection, and metastasis of human colon carcinoma cells in nude mice. *Cancer Research*, **48**, 6863–6871.

Newell, D. R., *et al*. (1999). Evaluation of rodent-only toxicology for early clinical trials with novel cancer therapeutics. *British Journal of Cancer*, **81**, 760–768.

Reeves, P. G., *et al*. (1993). AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition *Ad Hoc* Writing Committee on the reformulation of the AIN-76A rodent diet. *Journal of Nutrition*, **123**, 1939–1951.

Schiffer, S. P. (1997). Animal welfare and colony management in cancer research. *Breast Cancer Research and Treatment*, **46**, 313–331.

Shultz, L. D. (1989). Single gene models of immunodeficiency diseases. In: Wu, B. and Zheng, J. (eds), *Immune-deficient Animals in Experimental Medicine.* 19–26 (Karger, Basel).

Shultz, L. D., *et al*. (2000). NOD/LtSz-Rag1^{null} mice: an immunodeficient and radioresistant model for engraftment of human hematolymphoid cells, HIV infection, and adoptive transfer of NOD mouse diabetogenic T cells. *Journal of Immunology*, **164**, 2496–2507.

Thompson, E. W., *et al*. (1999). LCC-15-MB: a human breast cancer cell line from a femoral bone metastasis. *Clinical and Experimental Metastasis*, **17**, 193–204.

Thigpen, J. E., *et al*. (1999). Phytoestrogen content of purified, open- and closed-formula laboratory animal diets. *Laboratory Animal Science*, **49**, 530–536.

United Kingdom Co-ordinating Committee on Cancer Research (UKCCCR) (1998). Guidelines for the welfare of animals in experimental neoplasia. *British Journal of Cancer*, **77**, 1–10.

FURTHER READING

Arnold, W., *et al.* (1996). *Immunodeficient Animals: Models for Cancer Research*. (Karger, Basel).

Botting, J. H. (1992). *Animal Experimentation and the Future of Medical Research*. (Portland Press, London).

Gad, S. C. and Weil, C. S. (1988). *Statistics and Experimental design for Toxicologists*. (Telford Press, Telford, NJ).

Gart, J. J., *et al.* (1986). *The Design and Analysis of Long-Term Animal Experiments*. (International Agency for Research on Cancer, Lyon).

Joint Steering Committee of the EORTC and CRC (1990). General guidelines for the preclinical toxicology of new cytotoxic anticancer agents in Europe. *European Journal of Cancer*, **26**, 411–414.

Kallman, R. F. (1987). *Rodent Tumor Models in Experimental Cancer Therapy*. (Pergamon Press, New York).

Rygaard, J. and Poulsen, C. O. (1969). Heterotransplantation of a human malignant tumor to 'nude' mice. *Acta Pathologica, Microbiologica et Immunologica Scandinavica*, **77**, 758–760.

Woodman, D. D. (1997). *Laboratory Animal Endocrinology. Hormonal Action, Control Mechanisms and Interactions with Drugs*. (John Wiley & Sons, Chichester).

World Cancer Research Fund and American Institute for Cancer Research. *Food, Nutrition and the Prevention of Cancer: A Global Perspective* (1997). (American Institute for Cancer Research, Washington, DC).

Zurlo, J., *et al.* (1994). *Animals and Alternatives in Testing*. (Mary Ann Liebert, New York).

Web Sites

American Association for Laboratory Animal Science:
http://www.aalas.org

Association for the Assessment and Accreditation of Laboratory Animal Care:
http://www.aaalac.org

Canadian Council on Animal Care:
http://www.ccac.ca

Home Office (UK), Animals, Byelaws & Coroners Unit:
http://www.homeoffice.gov.uk/ccpd/abcu.htm

Humane Society of the USA:
http://www.hsus.org

National Research Council, Institute of Laboratory Animal Resources:
http://www.nap.edu/readingroom/books/labrats

National Institutes of Health Mammary Gland Biology Web Site:
http://mammary.nih.gov/models/index.html

Norwegian Reference Centre for Laboratory Science & Alternatives:
http://oslovet.veths.no

The Jackson Laboratory, Mouse Genome Informatics:
http://mgd.hgmp.mrc.ac.uk

US Department of Agriculture, Animal and Plant Health Inspection Service:
http://www.aphis.usda.gov

Chapter 59

Mammary Tumour Induction in Animals as a Model for Human Breast Cancer

Jose Russo and Irma H. Russo
Fox Chase Cancer Center, Philadelphia, PA, USA

CONTENTS

- Biological Basics
- Principles of Establishment
- General Applications
- Pathogenesis of Rat Mammary Tumours
- Biological and Technical Limitations
- Acknowledgements

BIOLOGICAL BASICS

Breast cancer is the most frequent spontaneous malignancy diagnosed in women in the Western world and it is also observed in rodent experimental models. In the United States alone 186 000 new cases were diagnosed in 1999; its incidence has been increasing for several decades. Early diagnosis has improved the rates of cure and prolonged survival, but stage-specific survival rates have increased only slightly since the mid-1970s, and breast cancer remains second only to lung cancer as a cause of death (Wingo *et al*., 1998).

From all experimental systems available for the study of mammary cancer, rodent models have been particularly useful, because these species develop spontaneous mammary tumours. In several strains of female rats they are the most common hormone-dependent spontaneous neoplasms developed (Welsh, 1987).

The susceptibility of the rodent mammary gland to develop neoplasms has made this organ a unique target for testing the carcinogenic potential of specific chemicals. Several carcinogens which induce mammary tumours in both mice and rats have been extensively studied in both species (Welsh, 1987). Tumours induced by administration of chemical carcinogens constitute useful tools for dissecting the multistep process of carcinogenesis, which involves initiation, promotion and progression, and serve as a baseline for testing the carcinogenic potential of chemicals in risk assessment. Chemically induced mammary tumours are, in general, hormone-dependent adenocarcinomas. Their incidence, number of tumours per animal and even tumour type, are influenced by the age of the host at the time of carcinogen exposure, reproductive history, endocrinological milieu and diet, among other factors. These factors, in turn, influence the development and degree of mammary gland differentiation (Russo, I. H., *et al*., 1989), which are subject to a multiplicity of endocrine stimulatory and inhibitory influences from embryogenesis onward (Daniel and Silberstein, 1987). If these influences are not exerted in proper temporal, sequential and quantitative relationships, normal development, differentiation and function are adversely affected. Mammary gland development, in turn, cannot be separated from its ageing, a process that markedly influences the incidence of spontaneous tumours in all strains of rats, as well as in mice infected with the mouse mammary tumour virus (MMTV).

The ideal animal tumour model should mimic the human disease. This means that the investigator should be able to ascertain the influence of host factors on the initiation of tumorigenesis, mimic the susceptibility of tumour response based on age and reproductive history and determine the response of the tumours induced to chemotherapy. The utilization of experimental models of mammary carcinogenesis in risk assessment requires that the influence of ovarian, pituitary and placental hormones, among others and also overall reproductive events, and age at menarche and menopause, are taken into consideration, since they are important modifiers of the susceptibility of the organ to neoplastic development. Several species, such as rodents, dogs, cats and monkeys, have been evaluated for these purposes; however, none of them fulfil all the criteria specified above. Rodents, however, are the most widely used models, and therefore this chapter will concentrate in discussing the rodent model.

PRINCIPLES OF ESTABLISHMENT

The mammary gland, a specialized accessory gland of the skin that characterizes the mammalian species, is

a frequent source of tumours or neoplasms. The terms tumour and neoplasm are applied indistinctly to either benign or malignant lesions, since tumour (from the Latin *tumere*, to swell), means any pathological enlargement or new growth, also called neoplasm (from the Greek *neos*, new, +*plasma*, formation); however, none of them defines the true nature of a given growth. Benign tumours are those that do not invade adjacent tissues, do not metastasize to distant sites and can be cured by local excision. Malignant tumours, or cancer, are neoplasms characterized by their ability to invade, metastasize and ultimately cause the death of the host. They are called 'carcinomas' when they are derived from epithelial cells and 'sarcomas' if they are mesenchymal in origin.

Spontaneous mammary tumours are frequently observed in long-term rodent studies (Rao *et al*., 1987). In mice the development of 'spontaneous' mammary tumours is linked to the infection of female mice with either an exogenous MMTV or a less virulent endogenous provirus. A third strain of MMTV transmitted through the milk and through the germ line has also been identified in the European mouse strain GR. The exogenous MMTV is an RNA virus first recognized to be transmitted through the milk of A and C3H strain mothers, the Bittner factor. DBA and RIII are also inbred strains of mice that harbour the highly oncogenic MMTV transmitted through the milk. Foster-nursed neonate mice (i.e. C3Hf or DBAf) become free of the milk-transmitted MMTV, although they retain the genetically transmitted MMTV, which induces mammary tumours late in life. In high-incidence strains of mice, tumours develop as a multistep process initiated in pre-neoplastic lesions, the hyperplastic alveolar nodules, which evolve from pregnancy-dependent to pregnancy-independent adenocarcinomas. Nulliparous mice develop a low incidence of mammary tumours. Out of a total of 1361 female B6Cf$_1$/CrlBR mice, only five adenomas (0.4%), four fibroadenomas (0.3%), 10 adenocarcinomas (0.9%) and seven carcinomas (0.6%) developed by the end of a 24-month follow-up. Some mammary tumours that develop in females of susceptible strains, such as C3H, A, DBA, CBA and certain sub-strains of Balb/c, are strongly hormone dependent in terms of their initiation. Multiple pregnancies enhance tumour development, and final tumorigenic response is greater in multiparous than in nulliparous animals. In RIII, BR6, DD, and GR mice mammary tumours develop during the first pregnancy, but they regress during lactation. In some strains of mice the growth of mammary tumours is stimulated by chronic administration of oestrogens, certain steroidal contraceptives, progesterone, PRL and epidermal growth factor (EGF), whereas hormone deprivation, induced by hypophysectomy, ovariectomy, ovariectomy–adrenalectomy, and siladenectomy, suppress mammary carcinogenesis (Welsh, 1987).

In the rat, the majority of spontaneously developed tumours, with the exception of leukaemia, are neoplasms of endocrine organs or of organs under endocrine control. Spontaneous mammary tumours develop in females of various strains of rats, such as August, Albany–Hooded, Copenhagen, Fisher, Lewis, Osborne–Mendel, Sprague–Dawley, Wistar and Wistar/Furth (Welsh, 1987). Spontaneous mammary tumours are third in incidence among spontaneous tumours found in the Fisher 344 rat used in the National Cancer Institute (NCI)/National Toxicology Program (NTP) carcinogenicity bioassays (Rao *et al*., 1987). They are predominantly benign tumours, i.e. fibroadenomas (**Figure 1**), fibromas and more rarely adenomas. Malignant tumours such as adenocarcinomas are rare, although they are the most frequent tumours induced by chemical carcinogens (**Figure 2**). The development of spontaneous tumours varies as a function of strain, age and endocrine influences. Mammary gland tumours develop in older females; they are more frequent in multiparous than in nulliparous rats. As in mice, hormone withdrawal inhibits tumour development, and hormone supplementation, such as chronic administration of oestrogens, increases the incidence of adenocarcinomas, whereas chronic administration of prolactin or of growth hormone stimulates benign tumour growth (Welsh, 1987).

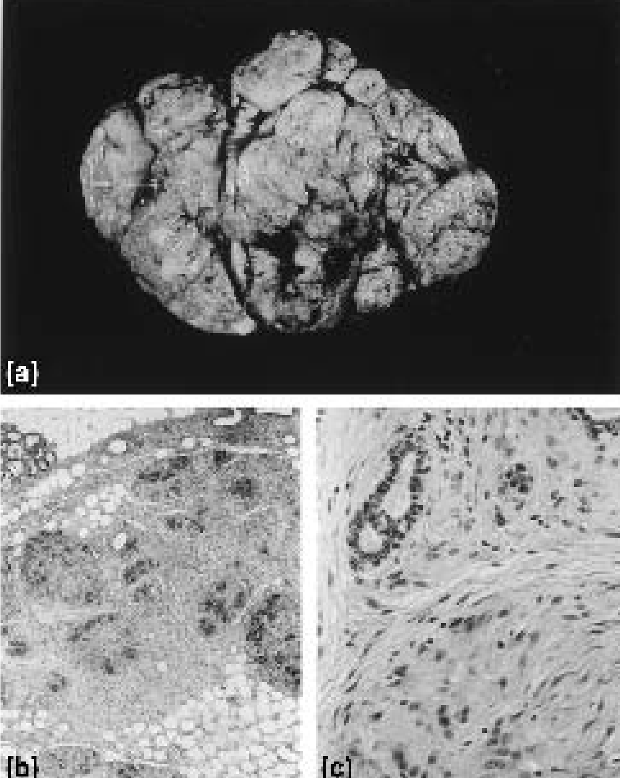
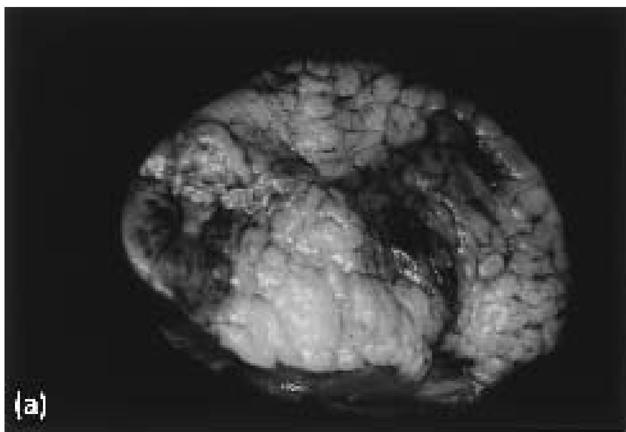


Figure 1 Spontaneous fibroadenoma in a female rat: (a) gross appearance of the dissected tumour; (b) and (c), histological sections of the same tumour reveal the predominance of stromal cells with few ductules interspersed among collagenous fibres. (b) and (c) haematoxylin and eosin stain, ×4 and ×25, respectively.



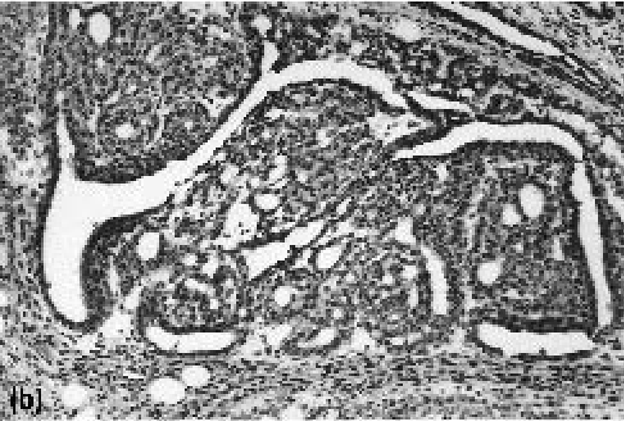


Figure 2 Adenocarcinoma induced by DMBA in virgin rats. (a) Gross appearance; (b) histological section of a papillary adenocarcinoma stained with haematoxylin and eosin, ×25.

The long latency period for spontaneous tumour development, up to 2 years in susceptible strains to develop a 50–70% tumour incidence, limits the usefulness of this model for experimental studies. (See also chapters *Transgenic Technology in the Study of Oncogenes and Tumour Suppressor Genes*; *Gene Knockouts in Cancer Research*.)

GENERAL APPLICATIONS

Chemically Induced Mammary Tumorigenesis

The potential of chemicals to induce cancer was recognized almost two centuries ago as an occupational disease, when a high incidence of skin cancer was linked to exposure to coal tar. Although it has not been proved that human breast cancer is caused by a given chemical or physical genotoxic agent, the human population is exposed to a large number of environmental chemicals, such as polycyclic aromatic hydrocarbons, nitrosoureas and aromatic amines, that have been demonstrated to be carcinogenic in experimental animal models and to induce mutagenesis and neoplastic transformation of human breast epithelial cells *in vitro* (Russo, J. *et al*., 1993). Although a specific aetiological agent or the conditions that might explain the initiation and progression of breast cancer in humans have not been identified, experimental animal models have proved to be useful tools for answering specific questions on the biology of mammary cancer relative to their validity to the human disease, and also for assessing the risk of breast cancer posed by toxic chemicals. *In vivo* experimental animal models provide information not available in human populations; they are adequate for hazard identification, dose–response modelling, exposure assessment and risk characterization, the four required steps for quantifying the estimated risk of cancer development associated with toxic chemical exposure. The utilization of experimental models of mammary carcinogenesis in risk assessment requires that the influence of ovarian, pituitary and placental hormones, among others, and also overall reproductive events, are taken into consideration, since they are important modifiers of the susceptibility of the organ to undergo neoplastic transformation.

Chemically induced mammary tumours develop by a multistep process. The initial step is a biochemical lesion caused by the interaction of the carcinogen with cellular DNA. In this interaction the DNA is damaged, and if the damage is not repaired efficiently, the result is a mutation, chromosomal translocation, inactivation of regulatory genes or more subtle changes not well identified as yet. Neoplastic development requires that the lesion becomes fixed, aided by cell proliferation, progressing to a third stage of autonomous growth, resulting in cancer, when the lesion acquires the capacity to invade and metastasize. Several carcinogens that induce mammary tumours in rodents have been identified and extensively studied for more than 50 years in mice and for more than 30 years in rats (Welsh, 1987). Mammary carcinomas have been induced in strains of mice with low spontaneous mammary cancer incidence with 3,4-benzopyrene, 3-methylcholanthrene (MCA), 1,2,5,6-dibenzanthracene, 7,12-dimethylbenz[*a*]anthracene (DMBA) and urethane. Most of the mammary tumours induced in mice by chemical carcinogens are adenoacanthomas and type B adenocarcinomas. They develop after a relatively long period of time, and their induction requires multiple applications. Enhanced tumorigenicity has been obtained with prolonged hormonal stimulation. However, the hormone responsiveness of chemically induced mammary tumours in mice has not been as thoroughly studied as it has been in the rat.

The most frequently utilized rat mammary carcinogens are DMBA and *N*-methyl-*N*-nitrosourea (MNU), although MCA, 2-acetylaminofluorene, 3,4-benzopyrene, ethylnitrosourea, and butylnitrosourea have also been extensively utilized as mammary cancer inducers (Zwieten, 1984; Welsh, 1987). The majority of rat mammary carcinomas induced by either DMBA or MNU are hormone dependent. Maximum tumour incidence is elicited when the carcinogens are administered to young virgin females

with an intact endocrine system. These models of hormone-dependent tumours constitute useful tools for dissecting the multistep process of carcinogenesis, and serve as a baseline for testing the carcinogenic potential of chemicals in risk assessment.

Radiation-induced Mammary Tumorigenesis

Ionizing radiation is probably the most widely acknowledged and studied human carcinogen (Zwieten, 1984). Exposure to radiation, either accidentally or for therapeutic reasons, has long being associated with a greater incidence of neoplasms, namely haematopoietic, gonadal and of the breast. The female breast is one of the tissues with the highest sensitivity to radiation carcinogenesis. Breast cancer developed in irradiated women shows a strong association with young age at the time of exposure, an association not observed in irradiated rodents, but similar to what has been observed in chemically induced mammary carcinogenesis in rats (Russo, J. *et al.*, 1977, 1979). Since controversy exists concerning the shape of the dose–response curve, the effects of fractionated irradiation and the effect of low levels of radiation, animal studies are necessary to address these issues. The rat model has been widely used in this regard, mainly since the demonstration in the early 1950s that a single lethal dose of X-rays to female Holtzman rats (a Sprague–Dawley stock) maintained by temporary parabiosis induced an increased number of benign and malignant mammary tumours within 6 months of exposure. Sublethal doses of different types of radiation, including X-rays and neutrons, have been shown to induce mammary tumour development, often within 1 year, with linear dose–effect relationships for neutrons over the total dose range and for X-rays down to dose levels of 0.2 Gy. Irradiation of animals with fractionated doses of γ-radiation has resulted in linear–quadratic dose–response curves.

Although most studies have utilized whole-body irradiation, localized irradiation also induces mammary tumours in the rat within the irradiated field. This effect occurs also in women, but reportedly not in several other animal species studied, e.g. mice, dogs and guinea pigs. In rats, mammary carcinomas can be induced by whole-body or segmental radiation with either X-rays, γ-rays or neutrons (Zwieten, 1984). Several studies utilizing a variety of fractionated irradiation protocols, i.e. at 12-h intervals for 60 days, semi-weekly for up to 16 weeks and monthly for up to 10 months, have shown in general, no increase in tumour latency, incidence or total number of mammary tumours, and no sparing or enhancing effect on mammary tumour development when compared with animals exposed to single doses of radiation. Some investigators, however, have reported an increased number of mammary carcinomas in animals receiving fractionated doses. Sprague–Dawley and Lewis rats are the most susceptible to radiation-induced tumorigenesis. AxC, Fisher, Long–Evans and Wistar/Furth strains are also susceptible, but to a lesser degree. The mammary tumours developed by irradiated rats are, in general, hormone-dependent adenocarcinomas or fibroadenomas. The hormonal status of the female rat is of paramount importance in determining the outcome of irradiation of the mammary gland. Ovariectomy completely prevents, and oestrogen treatment enhances, radiation-induced mammary tumour formation. The latency period for tumour development is shortened and tumour incidence is increased considerably in oestrogen-treated rats. The number of cribriform-type adenocarcinomas, and the number of tumours per tumour-bearing rat are also increased. Radiation and oestrogens, namely 17β-oestradiol (E_2) or diethylstilboestrol (DES), have been reported to exert either an additive or synergistic effect. The effect of E_2 administration and irradiation on mammary tumorigenesis has been reported to be equal for hormone administration 1 week before or beginning 12 weeks after irradiation, but no additive effect has been observed when hormone administration was begun 24 weeks after irradiation. The amplification of radiation-induced mammary tumorigenesis by oestrogens has been attributed by several investigators to the effects of this hormone on the pituitary, an interpretation supported by the observations that DES treatment of ACI rats and E_2 treatment of rats of three different strains, for example, result in increased incidence of pituitary tumours, accompanied by a marked increases in plasma prolactin (PRL) levels. The development of malignant mammary tumours in these rats appeared to be associated with the extent of increase in plasma PRL. Mammary tumour incidence and number or type of mammary tumours are not modified by irradiation during pregnancy, lactation or postlactational regression in comparison with irradiation in the virginal state, in contrast to what has been reported in chemically induced mammary carcinogenesis (Russo, J. *et al.*, 1979). Furthermore, radiation-induced mammary tumours developed in rats do not exhibit the age dependency observed in women or in chemically induced rat mammary carcinomas (Russo, J. *et al.*, 1977, 1979; Russo, J. and Russo, I. H., 1980a,b). They do not exhibit topographic selectivity in their development, since they arise randomly in thoracic and abdomino-inguinal regions (Russo, J. and Russo, I. H., 1987; Welsh, 1987). Further studies are needed to clarify these differences in tumour incidence between radiation- and chemically induced mammary tumours in rats for validating this model for risk assessment. The reason for the differences in physiological influences on the inductive action of chemicals and irradiation in rat mammary gland is not clear. It has been speculated that radiation-induced changes might occur in a specific stem cell population maintained throughout the reproductive life, whereas chemically induced changes depend upon the number and rate of turnover of other types

of mammary gland cells (Russo, J. and Russo, I. H., 1980a).

Genetic Background and Mammary Carcinogenesis

Genetic differences among individuals may affect their susceptibility to the carcinogenic effect of chemicals. Inheritance may well predispose an individual to develop certain specific types of cancer. These influences have been carefully dissected in rodent experimental animal models (Isaacs, 1986). In carefully designed experiments, it has been demonstrated that the susceptibility of rats to the chemical carcinogens 2-acetylaminofluorene (AAF), DMBA and MNU is genetically determined (Isaacs, 1986). Isaacs demonstrated that Buffalo, Lewis, Wistar/Furth and inbred Sprague–Dawley rats, all strains of Wistar genetic background, are highly susceptible to chemically induced carcinogenesis, whereas the non-Wistar-derived strains Fischer, August, ACI and Copenhagen are of low susceptibility. However, there are exceptions to this rule, since the Wistar-derived inbred WN strain is of low susceptibility, and the non-Wistar-derived Osborne–Mendel strain is highly susceptible. Of the commonly used strains, Sprague–Dawley and Wistar–Furth are the most susceptible and Fischer 344 and ACI rats show intermediate susceptibility. Copenhagen rats are essentially completely resistant even to the direct application of DMBA to the gland, although they do develop fibrosarcomas in response to parenteral DMBA. Extensive analyses comparing DMBA tumorigenesis, mammary gland growth rate, serum hormone levels and DMBA toxicokinetics in female rats of several strains and F_1 hybrids between the strains have not found major differences that correlated with susceptibility to tumorigenesis, since both susceptible and resistant strains develop similar percentages of malignant changes (60% in resistant, 80% in susceptible). However, macroscopically detectable tumours developed in 70% of susceptible and only in 10% of resistant glands. These findings have been confirmed by transplantation experiments between resistant or susceptible strains into F_1 hybrids, and between the two strains, and by direct exposure to DMBA. These observations indicate that genetic factors govern the progression from microscopic to macroscopic tumour, rather than from normal to histologically malignant epithelium, and furthermore, that resistant rats possess a dominant suppressor allele for the gene governing susceptibility. In contrast, tumour induction by DES is demonstrable in the ACI but not in the Sprague–Dawley strain of rats, although a co-carcinogenic effect of DES with DMBA can be shown in the latter. Both malignant and benign tumours are increased by the combined treatment, but there is a relatively greater increase in benign tumours. The response of target organs other than the mammary gland to DES is also different in these two strains of rats, but the mechanisms are not known.

PATHOGENESIS OF RAT MAMMARY TUMOURS

Chemical carcinogen induction of mammary tumours in rodents is one of the most widely studied and useful models of mammary carcinogenesis (Huggins *et al*., 1961; Russo, J. *et al*., 1977; Russo, I. H. and Russo, J., 1978; Russo, J. and Russo, I. H., 1978; Zwieten, 1984; Russo, I. H. *et al*., 1989). For those reasons, this section will focus mainly on that model. The two most widely utilized experimental systems of mammary tumorigenesis are the induction of rat mammary tumours by administration of either the indirectly acting polycyclic hydrocarbon DMBA, given intragastrically (i.g.) to Sprague–Dawley rats (Huggins *et al*., 1961), or the direct acting carcinogen NMU, given intravenously (i.v.) or subcutaneously (s.c.) to Sprague–Dawley or Fischer 344 rats, respectively (Gullino *et al*., 1975). A single i.g. dose of 80–100 mg of DMBA per kg body weight induces tumours with latencies that generally range between 8 and 21 weeks. The final tumour incidence reaches 100% when the carcinogen is administered to intact virgin rats in their peak of maximal susceptibility, that in Sprague–Dawley rats occurs between the ages of 40 and 60 days (Russo, J. *et al*., 1983). NMU, given in a single i.v. dose of 25 or 50 mg/kg body weight, yields tumours with similar latency and incidence (Gullino *et al*., 1975). In a comparative study between the carcinogenic potential of DMBA, administered i.g. at a dose of 20 mg/kg, and NMU, given i.v. at a dose of 50 mg/kg, it was demonstrated that both induce approximately equal tumour incidence and number of tumours per animal, with approximately equal latency, but a somewhat greater percentage of NMU-induced tumours were histologically malignant.

The analysis of the factors that modulate the tumorigenic response of the mammary gland, especially in short-term studies, have revealed that the most sensitive and reliable end points are both tumour latency and tumour histological type. Tumour latency is, in general, inversely related to carcinogen dose, whereas overall tumour incidence and number of malignant tumours per animal are directly related, especially when relatively early end points are used. No such a relationship has been found for benign tumours. The number of tumours per rat or per group and the number or incidence of malignant tumours are additional end points useful in analysis of data. Comparison of data among laboratories requires strict standardization of the experimental conditions, since considerable variations in tumour incidence and latencies between laboratories and between experiments in the same laboratory are frequently seen.

The induction of mammary carcinomas in the rat requires that the carcinogen acts on a specific compartment

of the mammary gland, the terminal end bud (TEB), a club-shaped undifferentiated structure found at the peripheral margins of the developing mammary parenchyma in young virgin rats (**Figure 3a** and **b**). Although TEBs are present in the six pairs of mammary glands, tumour development does not occur as a random event. Tumour incidence in animals treated with the carcinogen between the ages of 20 and 180 days is greater in those glands located in the thoracic region, whereas glands located in the abdominal and inguinal areas develop a smaller number of tumours (**Figure 4**). In addition to differences in tumour incidence as a consequence of the topographic location of the gland, there are differences in tumour type, which seem to vary with the age of the animal at the time of carcinogen treatment. Ductal and papillary adenocarcinomas are more frequent in both thoracic and abdominal glands of younger animals, whereas adenocarcinomas with a tubular pattern are found mostly in abdominal glands and in older animals. The development of the rat mammary gland occurs through a combined process of branching and differentiation of the parenchyma, mainly in those ducts ending in TEBs that progressively divide and differentiate into alveolar buds (ABs) (**Figure 3c**). These structures in turn differentiate into lobules. Although this pattern of development is common to the six pairs of mammary glands, it does not occur simultaneously in all of them, but varies in relation to the topographic location of each specific pair. Individual structures, i.e. TEBs, ABs and lobules, appear similar in morphology in all the glands, but their relative number and

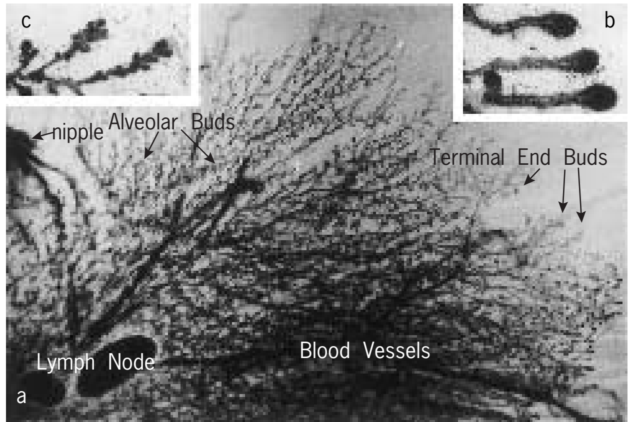


Figure 3 (a) Whole mount preparation of the abdominal mammary gland of a 55- day-old virgin rat. Arrows indicate the location of terminal end buds at the periphery of the gland parenchyma (inset b), alveolar buds (inset c) and the nipple. Toluidine blue, $\times 4$. (b) Enlargement of terminal end buds indicated in (a), $\times 10$. (c) Alveolar buds from the area shown in (a), $\times 10$.

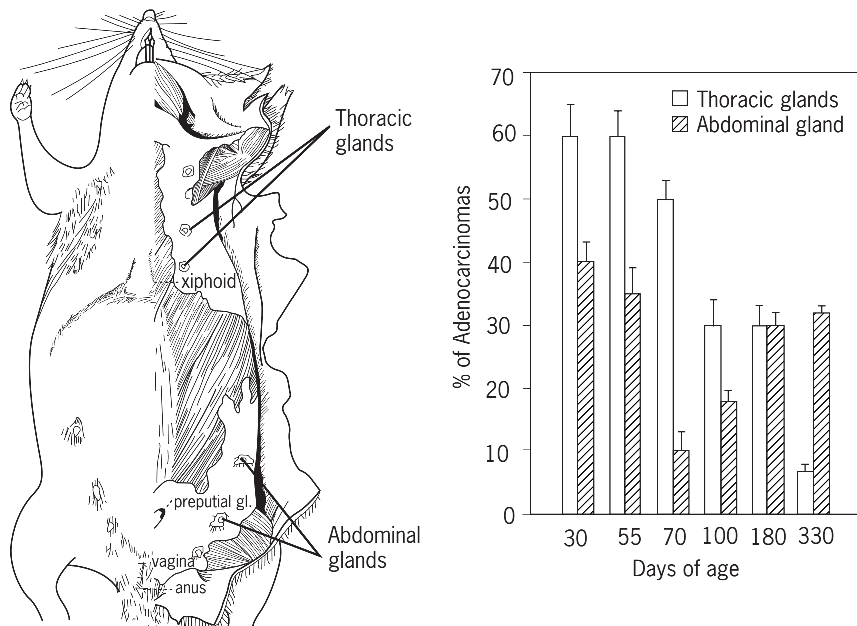


Figure 4 Drawing of a rat showing the distribution of the six pairs of mammary glands. The fourth pair, or abdominal mammary gland, is the most frequently utilized for morphological, cell kinetic and tumorigenic studies because of its large size and ease of access. However, they are less susceptible to be transformed by chemical carcinogens, as indicated in the histogram at the right-hand side, which shows that the higher incidence of adenocarcinomas occurs in thoracic mammary glands in those animals inoculated with the carcinogen at ages younger than 100 days. These results were obtained by inoculating virgin Sprague–Dawley rats intragastrically with a single dose of 8 mg of 7,12-dimethylbenz[a]anthracene (DMBA) (Sigma Chemical, St. Louis, MO, USA) per 100 g body weight. The carcinogen was dissolved in corn oil at a concentration of 16 mg mL^{-1} by heating in a water-bath at 95°C for 30 min.

the general architecture of the organ vary considerably from one pair of glands to another. The most notable ones are the thoracic mammary glands, since each single gland is composed of two different layers separated by connective and muscular tissue; one layer is composed of more numerous ABs and small lobules, whereas the adjacent one is more extensive and contains thin long ducts ending in prominent TEBs. The abdominal glands have a markedly reduced number of TEBs, which are located exclusively in the most distal portion of the gland, whereas the middle and proximal portions show a much more differentiated appearance. The difference in number of TEBs in thoracic versus abdominal mammary glands is significant. With ageing, TEBs decrease progressively and their reduction is proportional in all the glands. This reduction is mostly due to either their regression to TDs or to a greater differentiation to ABs and lobules (**Figure 5**). The higher incidence of ductal carcinomas observed in thoracic glands is attributed to the difference in degree of development of the undifferentiated layer of this gland in comparison with the glands located in other topographic areas (Russo, I. H. *et al*., 1989).

Mammary Gland Differentiation as a Modulator of Carcinogenic Response

Mammary cancer in experimental models is the result of the interaction of a carcinogen with the target organ, the mammary gland. This target, however, is extremely complex, since the mammary gland does not respond to the carcinogen as whole, but only specific structures within the gland are affected by given genotoxic agents. The knowledge of the architecture and cell kinetic characteristics of the mammary gland at the time of carcinogen administration constitutes a necessary initial step for understanding the pathogenesis of the disease. It is also required for distinguishing those changes induced by the carcinogen from changes reflecting normal gland development, especially when evaluating early tumorigenic response in short-term studies (Russo, J. *et al*., 1977; Russo, J. and Russo, I. H., 1978; Russo, I. H. and Russo, J. 1978).

The susceptibility of the mammary gland to DMBA- or NMU-induced carcinogenesis is strongly age dependent; it is maximum when the carcinogens are administered to virgin females between the ages of 40 and 60 days, that is, soon after vaginal opening and during early sexual maturity (**Figure 5**) (Russo, J. *et al*., 1977). Active organogenesis and high rate of proliferation of the glandular epithelium are characteristics of that period, in which there is also high DMBA activation (**Figure 6**) (Russo, I. H. and Russo, J., 1978; Russo, J. and Russo, I. H., 1978, 1980a,b). Its significance, however, is uncertain, since NMU, which is also most effective at that age, does not require activation. The incidence of DMBA-induced tumours reaches 100% when the carcinogen is administered to rats aged 30–55 days, but the highest number of tumours per animal is observed when the carcinogen is given to animals between the ages

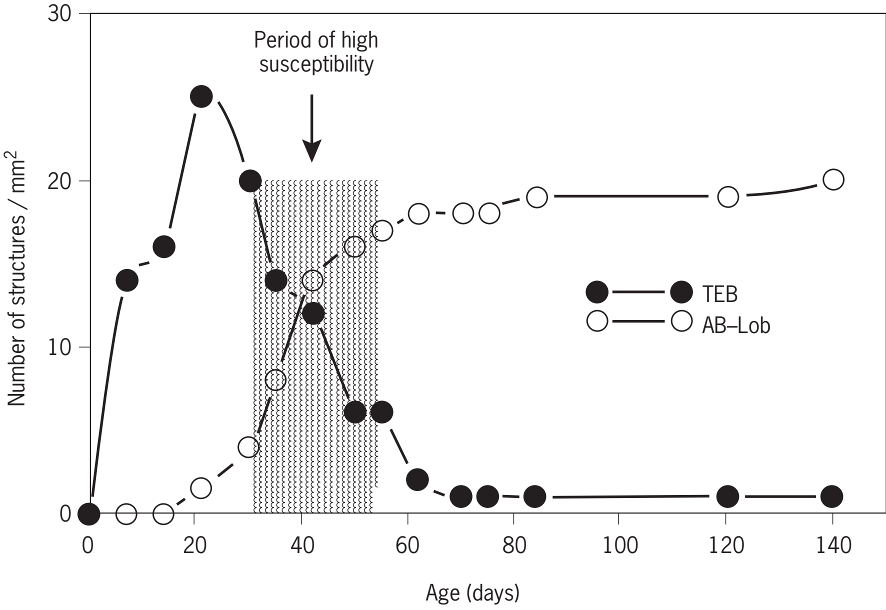


Figure 5 Histogram representing the number of structures per mm^2 (ordinate) of the rat mammary gland in relation to age (abscissa). Hatched area, period of highest susceptibility of the mammary gland to neoplastic transformation by chemical carcinogens. (From Russo, J. and Russo, I. H., 1994, *Cancer Epidemiology Biomarkers and Prevention*, **3**, 353–364.)

Structure	DNA-LI	Tc	Cell cycle	GF	[^{3}H]DMBA	Lesions
TEB	34%	11.65 h	G_2, M, G_1(0.7 h), S	0.55	6.8 ± 2.8	Carcinoma
AB	4%	28.18 h	G_2 M, S, G_1(17 h)	0.13	1.3 ± 0.8	Cysts Hyperplastic Alveolar Nodules Adenomas Fibroadenomas
Lobule	0.1%	49.63 h	S, G_2 M, G_1(35 h)	0.0049	0.9 ± 0.5	None

Figure 6 Correlation between type of structure, terminal end bud (TEB), alveolar bud (AB) and lobule with the DNA-labelling index (DNA-LI), length of the cell cycle in hours (Tc), cell cycle and growth fraction (GF). Nuclear uptake of [^{3}H]DMBA, detected by autoradiography and expressed as the number of grains per nucleus, is directly proportional to the DNA-LI and GF and inversely proportional to the Tc of each specific structure. The development of carcinomas also correlated with high DNA-LI, Tc and DMBA binding, whereas benign lesions or the absence of neoplasms were seen in an inverse relation. (From Russo, J. and Russo, I. H., 1994, *Cancer Epidemiology Biomarkers and Prevention*, **3**, 353–364.)

of 40 and 46 days, coincident with the period in which the mammary gland exhibits a high density of highly proliferating TEBs (**Figures 3, 5 and 6**). This high susceptibility is attributed to the specific characteristics of the mammary gland prevailing during that period of life. Administration of DMBA to virgin rats induces the largest number of transformed foci when TEBs are decreasing in number owing to their differentiation into ABs (**Figures 3 and 5**). These structures, instead of differentiating into ABs, become progressively larger owing to epithelial proliferation, with multi-layering, secondary lumen formation and early papillary projections to the widened lumen; at this stage, transformed TEBs are called intraductal proliferations (IDPs) (**Figure 7**). Their confluence leads to the formation of microtumours that histologically are classified as adenocarcinomas, first intraductal, which progress to invasive tumours, developing various patterns such as cribriform, comedo or papillary types (**Figures 2 and 7**) (Russo, J. *et al*., 1977).

Even though TEB differentiation into AB is inhibited by carcinogen treatment, not all the TEBs present in the mammary gland at the time of DMBA administration progress to IDPs. Some of them still differentiate into ABs but their number is always smaller than that of control animals. Occasional lobular development is observed, although it is negligible. Some TEBs become smaller with an atrophic appearance being called at this stage terminal ducts (TDs). TDs are also susceptible to neoplastic transformation and are the main target of carcinogens in older animals (Russo, I. H., *et al*., 1989). Those TEBs which were already differentiated into ABs and early lobular structures before DMBA administration do not develop carcinomas. Most of them either remain unmodified or undergo dilatation of the lumen, giving rise to hyperplastic lesions, such as alveolar bud hyperplasia (**Figure 7**). Others exhibit epithelial proliferation, forming tubular adenomas, give rise to cystic dilatations, or fibroadenomas (**Figure 7**). When DMBA is inoculated into old virgin females, ranging in age from 180 to 330

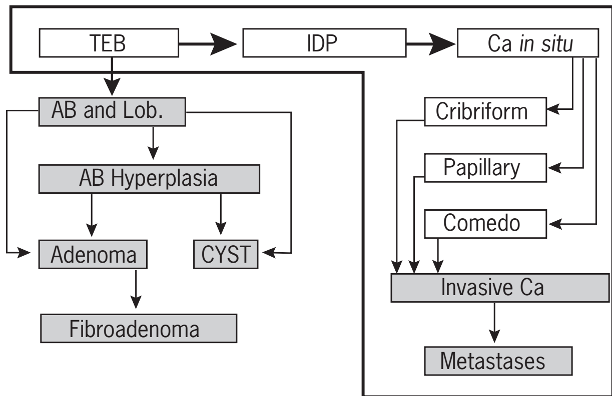


Figure 7 Pathogenesis of chemically induced rat mammary tumours. The undifferentiated terminal end bud (TEB) affected by the carcinogen progresses to intraductal proliferation (IDP), and *in situ* ductal carcinoma (Ca *in situ*) that exhibit various histopathological types. Further tumoural growth and coalescence of neighbouring lesions originate invasive carcinomas (Invasive Ca), which might become metastatic. When the carcinogen affects more differentiated structures, such as alveolar buds (AB) and lobules (Lob.) the lesions developed are more benign in nature and appear later than the ductal carcinomas. Alveolar bud hyperplasia (AB Hyperplasia).

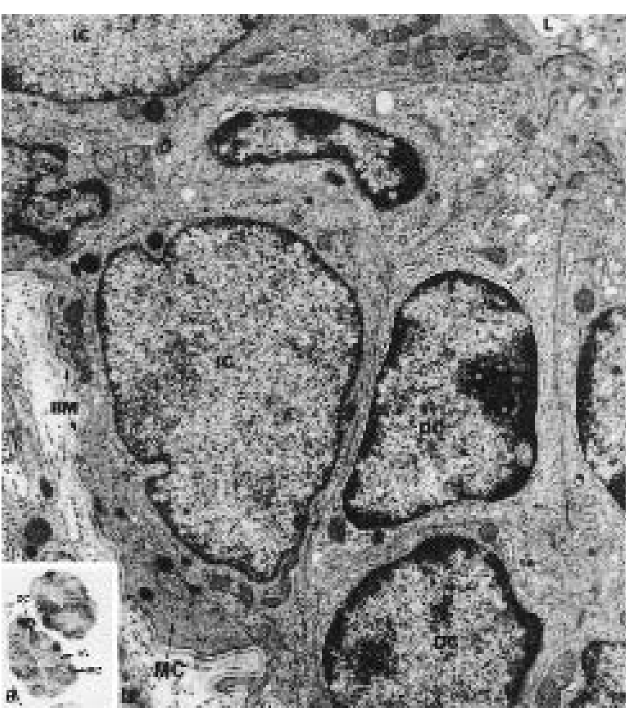


Figure 8 The rat mammary gland ductal system is lined by three types of cells, myoepithelial cells (MC), dark cells (DC) and intermediate cells (IC). (a) A 1-μm section of a rat mammary gland embedded in plastic and stained with toluidine blue, $\times$40. (b) Electron micrograph of the rat mammary gland showing the three cell types. The basement membrane (BM), that separates the epithelial cells from the stroma, is opposite to the lumen (L). Uranyl acetate–lead citrate, $\times$4000.

days, they develop tubular adenomas that exhibit focal areas of malignant transformation, giving rise to well-differentiated adenocarcinomas with a tubular pattern. These lesions develop predominantly in the abdominal glands, in which a higher incidence of tumours is observed in older animals (**Figure 4**). The observation that mammary carcinomas arise from undifferentiated structures of the gland, namely TEBs and TDs, whereas benign lesions such as adenomas, cysts and fibroadenomas (**Figures 1 and 2**) arise from structures that were more differentiated at the time of carcinogen administration, indicates that the carcinogen requires an adequate structural target, and the type of lesion induced is dependent upon the area of the mammary gland that the carcinogen affects (**Figure 7**). Hence the more differentiated the structure at the time of carcinogen administration, the more benign and organized is the lesion that develops.

Cell of Origin of Rat Mammary Carcinomas

In the rat mammary gland parenchyma, three types of cells have been described, dark, intermediate and myoepithelial cells (**Figure 8**). The distribution of cell populations in the mammary gland during carcinogenesis varies in TEBs and TDs (**Figure 9a and b**), starting as early as 24 h post-DMBA administration, but no changes in cell composition occur in more differentiated structures such as ABs and lobules. The changes taking place in TEBs and TDs are limited to the dark-cell type, the proportion of which decreases from 76 to 67% and to the intermediate cell, the proportion of which increases from 11 to 19%. Myoepithelial cells are unaffected. The trend is a progressive shift of cell population distribution, with a continuous decrease in dark cells and a concomitant increase in intermediate cells (**Figure 10**). By 14 days post-DMBA, the latter constitute about 40 and 50% of the proliferative compartment in TEBs and TDs, respectively. At this time the morphological manifestations of tumorigenesis have started to become apparent: increased number of epithelial cell layers, greater irregularity of the luminal border and progressively larger intercellular spaces, which in some cases form secondary lumina, are features indicative of the formation of an intraductal proliferation (IDP).

The basal lamina becomes distorted, thus rendering the identification of myoepithelial cells more difficult, and the surrounding stroma becomes fibrotic and infiltrated by inflammatory cells. Between 21 and 40 days post-DMBA

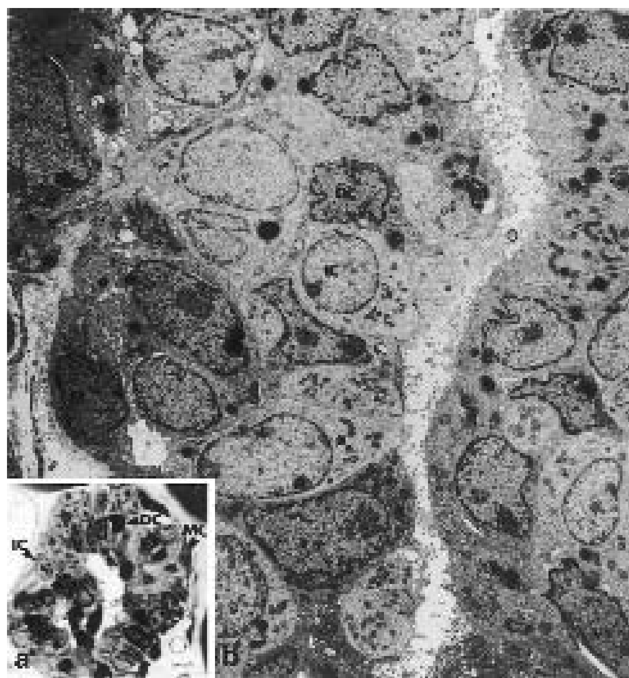


Figure 9 The terminal end bud (TEB) of the rat mammary gland is composed of the three cell types described in **Figure 8**. (a) A 1-μm section of the TEB embedded in plastic and stained with toluidine blue, ×40 (b) Electron micrograph of the TEB showing the three cell types. Uranyl acetate–lead citrate, ×2000. (Part (b) from Russo *et al*., 1983, *American Journal of Pathology*, **113**, 50–66.)

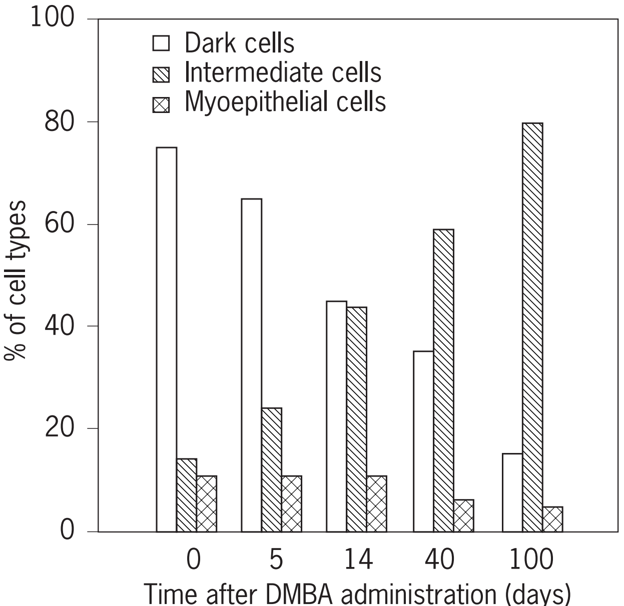


Figure 10 Histogram showing the relative proportion of myoepithelial, dark and intermediate cell types in the rat mammary gland of DMBA treated rats. (Adapted from Russo *et al*., 1983.)

treatment, the descending curve of dark cells crosses over the progressively ascending curve of intermediate cells. By 40 days, the tumours display a cell distribution of >65% intermediate cells and <35% dark cells (**Figure 10**). The proportion of intermediate cells continues to increase with tumour age, and at 70 days they represent >75% of the cells. Dark cells, at this point, have been reduced to <20%, whereas myoepithelial cells remain at approximately 5%. The 100-day-old tumours are dominated by intermediate cells comprising nearly 90% of the total number of cells (Russo, J. *et al*., 1983).

Cell Kinetics and Mammary Carcinogenesis

In every tissue, normal or abnormal, cell composition consists of a balance of three different cell populations: cycling cells, resting cells (cells in G_0), and dying cells (cell loss). In the mammary gland these three cell populations can be identified through the study of the cell cycle and determination of the growth fraction and the rate of cell loss. The growth fraction refers to the fraction of cycling cells, while the rate of cell loss refers to the fraction of cells that die or migrate to other tissues. Both cell cycle time and the growth fraction determine the number of cells produced per unit time, and the rate of cell loss determines the number of cells lost per unit time. The growth of normal cells involves a net increase in cell number resulting from more cells being born than dying. In the differentiated or in the adult tissue, in which growth has ceased, the number of cells produced per unit time is equal to the number of cells that die. (See also chapter *Models for Tumour Growth and Differentiation*.)

The higher susceptibility of the TEB to neoplastic transformation is attributed to the fact that this structure is composed of an actively proliferating epithelium, as determined by the mitotic index (MI), DNA-labelling indices (DNA-LI), length of the cell cycle (Tc), and growth fraction (GF) (**Table 1**, **Figure 6**) (Russo, J. and Russo, I. H., 1980a,b). Both the MI and DNA-LI are very high at the tip of TEBs, decreasing progressively toward the ductal or proximal portion of the gland and even further in ABs and lobules. TEBs are also characterized by having the highest growth fraction, which progressively diminishes in the more differentiated ABs and lobules (**Table 1**, **Figure 6**). By using these cell kinetic parameters we calculated the rate of cell loss in each one of the compartments of the mammary tree. Interestingly, the TEB is the structure not only with the highest proliferative ratio, but also with the lowest percentage of cell loss in comparison with other parenchymal structures. The rate of cell loss is very high in the lobular structures present in the mammary gland

Table 1 Cell kinetics in the mammary gland[a]

	Growth fraction	Rate of cell birth/hr	Rate of cell loss/hr
Y.V. TEB	0.55	472	44.96
Y.V. ducts	0.39	187	60.99
Y.V. lobules	0.13	46	87.00
Parous ducts	0.009	4.05	99.00
Parous lobules	0.004	0.96	98.00

[a]The growth of the mammary gland is the result of a balance between cell birth and cell loss. The growth fraction (GF) represents the total population of cycling cells, and is determined as the rate of ^{3}[H]thymidine incorporation after 5 days of continuous infusion (GF_5). It is maximum in the terminal end bud (TEB) of the young virgin (Y.V.) rat, decreasing progressively in ducts and lobules of the Y.V. animal. Pregnancy results in a more dramatic reduction on GF in both ducts and lobules.

of parous rats. This clearly indicates that the TEB of the young virgin female is the truly proliferating structure of the gland that reaches a steady state only after acquiring full differentiation. The differences in proliferative activity and growth fraction observed between TEBs and the more differentiated structures of the mammary gland are also reflected in variations in the length of Tc (**Figure 6**). Tc in TEBs of young virgin rats has an average length of 11 h, increasing to 20.81 h and 28.18 h in TDs and ABs, respectively. Further mammary gland differentiation as a consequence of ageing and pregnancy results in an even longer Tc, mainly due to a lengthening of the G_1 phase of the cell cycle (**Figure 6**) (Russo, J. and Russo, I. H., 1980a). The length of Tc also varies according to the cell type and to the specific compartment in which each given cell type is located in. The shortest Tc is observed in intermediate cells located in TEBs, whereas it lengthens when the same cell type is located in ABs or lobules. These differences in the length of Tc are due mainly to differences in the length of the G_1 phase of the cell cycle, whereas all the other phases remain constant. Intermediate-type cells located in TEBs have a Tc lasting 13 h. When the same cell type is located in more differentiated structures, such as ABs, it exhibits a lengthened Tc, lasting 34 h. These differences in cell kinetic parameters among the different cell types could explain the higher susceptibility of the intermediate cell of the TEBs to be affected by the carcinogen, which causes further expansion of the proliferative compartment of the intermediate cells and depression in the dark cell population after initiation of the carcinogenic stimulus (Russo, J. *et al*., 1983).

Mammary epithelial cells metabolize DMBA to polar metabolites with the formation of epoxides that cause DNA damage. When dissociated mammary epithelial cells of young virgin and of parous animals, that basically represent the cells of the TEBs and of the lobules, respectively, are grown *in vitro*, they exhibit different rates of formation of polar metabolites. TEB cells produce more polar and less phenolic metabolites than lobular cells, indicating that the former, in addition to their higher proliferative activity, are also producing more epoxides, as is manifested by a higher binding of DMBA to DNA (**Figure 6**). Autoradiographic studies performed *in vivo* confirm the observation that the greatest uptake of [^{3}H]DMBA occurs in the nucleus of epithelial cells of TEBs and the lowest in ABs and lobules, indicating that the highest DMBA–DNA binding is associated with the structure of the gland with the highest replicative properties. The ability of the cells to remove DMBA adducts from the DNA is an indication of their capability to repair the damage. TEB cells remove adducts formed less efficiently than lobular cells. This is attributed to the shorter G_1 phase of Tc and not to a lack of reparative enzymes (Tay and Russo, 1981).

These studies support the conclusion that the differentiation of the mammary gland modifies the following parameters: (1) glandular structure; (2) cell kinetics characteristics of the mammary epithelium, decreasing the growth fraction and lengthening the cell cycle, mainly the G_1 phase; (3) decreasing formation of polar metabolites and increasing phenolic metabolites; and (4) decreasing binding of the carcinogen. All the parameters listed above affect the susceptibility of the mammary gland to carcinogenesis, and should be taken into account when assessing chemicals for cancer risk.

BIOLOGICAL AND TECHNICAL LIMITATIONS

This chapter has described and compared experimental *in vivo* models currently used for assessing the mammary carcinogenic potential of given chemical and physical agents, and addressed the need for identifying those models that provide an answer regarding the mechanisms involved in cancer initiation and progression in the human population. The main problems identified in this study are the complexity and multistep nature of the carcinogenic process, the lack of understanding of the mechanism(s) involved in the initiation and progression of the disease and the lack of identification of a specific aetiological agent or agents of human breast cancer. The usefulness of chemically induced mammary tumours lies in their hormone dependence, the high frequency of adenocarcinomas histologically similar to human breast cancers, the possibility that they offer for dissecting the initiation, promotion and progression steps of carcinogenesis and the baseline information they provide in risk assessment. Chemically induced mammary tumours develop by a multistep process, which initiates as a biochemical lesion caused by the interaction of the carcinogen with cellular DNA. In this interaction the DNA is damaged, and if the damage is not repaired efficiently, the result is a mutation, chromosomal

translocation, inactivation of regulatory genes or more subtle changes not well identified as yet. Neoplastic development requires that the lesion becomes fixed, aided by cell proliferation, progressing to a third stage of autonomous growth, resulting in cancer, when the lesion acquires the capacity to invade and metastasize. Several carcinogens, such as 3,4-benzopyrene, MCA, 1,2,5,6-dibenzanthracene, DMBA and MNU, have been extensively studied in mice and/or rats. DMBA and MNU have been the most frequently used; the majority of the mammary tumours induced in rats by either one of these agents are hormone-dependent adenocarcinomas. The observation that maximum tumour incidence is elicited when the carcinogens are administered to young virgin female rats has led to important discoveries on the role of mammary gland development and differentiation in cancer initiation. Hormonally induced differentiation and the interplay of ovarian, pituitary and placental hormones have been identified as powerful modulators of the tumorigenic response elicited by genotoxic chemicals.

ACKNOWLEDGEMENTS

This manuscript was supported by Grant RO1 CA67238 from the NIH, PHS.

REFERENCES

Daniel, C. W. and Silberstein, G. B. (1987). Postnatal development of the rodent mammary gland. In: Neville, M. C. and Daniel, C. W. (eds). *The Mammary Gland. Development, Regulation and Function.* 3–36. (Plenum Press, New York).

Gullino, P. M., *et al*. (1975). *N*-Nitrosomethylurea as mammary gland carcinogen in rats. *Journal of the National Cancer Institute*, **54**, 401–410.

Huggins, C., *et al*. (1961). Mammary cancer induced by a single feeding of polynuclear hydrocarbons and its suppression. *Nature*, **1989**, 204–207.

Isaacs, J. T. (1986). Genetic control of resistance to chemically induced mammary adeno-carcinogenesis in the rat. *Cancer Research*, **46**, 3958.

Rao, G. N., *et al*. (1987). Influence of body weight on the incidence of spontaneous tumours in rats and mice of long-term studies. *American Journal of Clinical Nutrition*, **45**, 252–260.

Russo, I. H. and Russo, J. (1978). Developmental stage of the rat mammary gland as determinant of its susceptibility to 7,12-dimethylbenz(*a*)anthracene. *Journal of the National Cancer Institute*, **61**, 1439–1449.

Russo, I. H., *et al*. (1989). Morphology and development of rat mammary gland. In: Jones, T. C., *et al*. (eds), *Integument and Mammary Gland of Laboratory Animals*. 233–252 (Springer, Berlin).

Russo, J. and Russo, I. H. (1978). DNA labeling index and structure of the rat mammary gland as determinant of its susceptibility to carcinogenesis. *Journal of the National Cancer Institute*, **61**, 1451–1459.

Russo, J. and Russo, I. H. (1980a). Influence of differentiation and cell kinetics on the susceptibility of the rat mammary gland to carcinogenesis. *Cancer Research*, **40**, 2677–2687.

Russo, J. and Russo, I. H. (1980b). Susceptibility of the mammary gland to carcinogenesis. II. Pregnancy interruption as a risk factor in tumor incidence. *American Journal of Pathology*, **100**, 497–512.

Russo, J., *et al*. (1977). Pathogenesis of mammary carcinoma induced in rats by 7,12-dimethylbenz(*a*)anthracene. *Journal of the National Cancer Institute*, **59**, 435–445.

Russo, J., *et al*. (1979). Susceptibility of the mammary gland to carcinogenesis. I. Differentiation of the mammary gland as determinant of tumor incidence and type of lesion. *American Journal of Pathology*, **96**, 721–734.

Russo, J., *et al*. (1983). Susceptibility of the mammary gland to carcinogenesis. III. The cell of origin of mammary carcinoma. *American Journal of Pathology*, **113**, 50–66.

Russo, J., *et al*. (1993). A critical approach to the malignant transformation of human breast epithelial cells. *CRC Critical Reviews in Oncogenesis*, **4**, 403–417.

Tay, L. K. and Russo, J. (1981). 7,12-Dimethylbenz(*a*)anthracene (DMBA) induced DNA binding and repair synthesis in susceptible and non-susceptible mammary epithelial cells in culture. *Journal of the National Cancer Institute*, **67**, 155–161.

Welsh, C. W. (1987). Rodent models to examine *in vivo* hormonal regulation of mammary gland tumorigenesis. In: Medina, D., *et al*. (eds), *Cellular and Molecular Biology of Mammary Cancer*. 163–179 (Plenum Press, New York).

Wingo, P. A., *et al*. (1998). Cancer incidence and mortality, 1973–1995: a report card for the U.S. *Cancer*, **82**, 1197–1207.

Zwieten, M. J. (1984). Normal anatomy and pathology of the rat mammary gland. *Cancer Research*, **44**, 53–134.

FURTHER READING

Russo, I. H., *et al*. (1989). Endocrine Influences on mammary structure and development. In: Jones, T. C., *et al*. (eds), *Integument and Mammary Gland of Laboratory Animals*. 252–266 (Springer, Berlin).

Russo, J. and Russo, I. H. (1994). Toward a physiological approach to breast cancer prevention. *Cancer Epidemiology Biomarkers and Prevention*, **3**, 353–364.

Russo, J. and Russo, I. H. (1987). Biological and molecular bases of mammary carcinogenesis. *Laboratory Investigation*, **57**, 112–137.

Russo, J., *et al*. (1982). Differentiation of the mammary gland and susceptibility to carcinogenesis. *Breast Cancer Research and Treatment*, **2**, 5–37.

Russo, J., *et al*. (1989). Classification of neoplastic and non-neoplastic lesions of the rat mammary gland. In: Jones, T. C., *et al*. (eds), *Integument and Mammary Glands of Laboratory Animals*. 275–304 (Springer, Berlin).

Russo, J., *et al*. (1990). Comparative study of human and rat mammary tumourigenesis. *Laboratory Investigation*, **62**, 1–32.

Tay, L. K. and Russo, J. (1981). Formation and removal of 7,12-dimethylbenz(*a*)anthracene nucleic acid adducts in rat mammary epithelial cells with different susceptibility to carcinogenesis. *Carcinogenesis*, **2**, 1327–1333.

Chapter 60

Mathematical Models in Cancer Research

Mark A. J. Chaplain
University of Dundee, Dundee, UK

CONTENTS

- Biological Background
- Principles of Mathematical Modelling
- Applications of Mathematical Modelling to Tumour Growth
- Perspectives
- Acknowledgements

BIOLOGICAL BACKGROUND

The aim of this chapter is to introduce several mathematical models which have been developed to analyse theoretically the growth and development of solid tumours and other associated processes (such as angiogenesis). Strictly, the mathematical models we will present are neither *in vitro* nor animal models; perhaps the terms *in silico* (given that an important part of the modelling work is carried out on a computer) or *in mente* (given that much of the modelling must be carried out in one's mind) are more appropriate (cf. Little *et al*., 1998). Whatever the most appropriate description of this type of activity, it is clear that these are not *in vivo* models. These are theoretical models, which nonetheless can be used to shed light on a complicated phenomenon. It is hoped that the work in this chapter will be of direct relevance to and will link with the work in the following chapters: *Human Tumours in Animal Hosts*; *Mammary Tumour Induction in Animals as a Model for Human Breast Cancer*; *Angiogenesis Models*; *Models for Tumour Cell Adhesion and Invasion*; *Tumour Metastasis Models*; *Models for Tumour Cell–Endothlial Cell Interactions* and *Models for Immunotherapy and Cancer Vaccines*. Before introducing the mathematical models, we will first give a brief biological overview of the key stages of cancer progression on which we intend to focus and model with systems of mathematical equations. Specifically, we focus on the immune response to cancer (at an early avascular stage), angiogenesis, invasion and metastasis.

The development of a primary solid tumour (e.g. a carcinoma) begins with normal individual cells becoming transformed as a result of mutations in certain key genes. These transformed cells differ from normal ones in several ways, one of the most notable being their escape from the body's homeostatic mechanisms, leading to inappropriate proliferation. Individual tumour cells have the potential, over successive divisions, to develop into clusters (or nodules) of tumour cells. Subsequent growth and development require many coordinating factors to occur and these are described below. In the most general terms then, a neoplasm (tumour) may be defined as an abnormal mass of tissue whose growth exceeds that of normal tissue, is uncoordinated with that of the normal tissue and persists in the same excessive manner after cessation of the stimuli which evoked the change (MacSween and Whaley, 1992). A cancer, or malignant tumour, is a tumour that invades surrounding tissues, traverses at least one basement membrane zone, grows in the host tissue at the primary site and has the ability to grow in a distant host tissue, forming secondary cancers or metastases.

In the following section we describe the various stages of growth involved in tumour progression in more detail.

Early Solid Tumour Growth and the Immune Response

Tumours which originate spontaneously in humans or animals usually grow slowly. Many months, years or even dozens of years are required for their clinical manifestation (Uhr *et al*., 1997) and the precise nature of this phenomenon remains unclear. One of the reasons for the slow growth of tumours and, in some cases, for their regression, may be the reaction of the host immune system to the nascent tumour cells. Many authors point out that intensive lymphoid, granulocyte and monocyte infiltration in a tumour, especially pronounced at the early stages, correlates with a favourable prognosis (Lord and Burckhardt, 1984). However, the possible involvement of the immune system in the control of growth and development, including metastases in humans, is still a matter of some debate. Indeed, some researchers do not confirm the prognostic role of lymphoid infiltration in a tumour. Moreover, data have been reported which suggest a stimulating effect of certain cells of the immune system on the growth of tumours (Prehn, 1994). These apparent contradictions may

be associated with, and partly explained by the fact that carcinogenesis is a multi-stage process and that right from its inception there exists a complex tissue architecture within the tumour. Indeed, a tumour is a highly intricate, heterogeneous structure consisting of various cells which secrete a not inconsiderable quantity of biologically active compounds into the local environment (Franks and Teich, 1986).

The early stage of primary tumour formation often occurs in the absence of a vascular network and it is believed that this stage may last up to several years. The tumour nodule grows to an approximate size of 1–3 mm in diameter, containing close to 10^6 cells, and then growth slows and/or ceases. This limitation of growth is attributed by researchers to the competition between tumour cells for metabolites, the competition between tumour cells and cells of the immune system for metabolites and/or a direct cytostatic/cytotoxic effect produced by the tumour cells on each other. In addition, the early (avascular) stage and the subsequent stages of tumour growth are characterized by a chronic inflammatory infiltration of neutrophils, eosinophils, basophils, monocytes/macrophages, T-lymphocytes, B-lymphocytes and natural killer cells (Wilson and Lord, 1987). These cells penetrate the interior of the tumour and accumulate in it owing to attractants secreted from the tumour tissue and the high locomotive ability of activated immune cells (Ratner and Heppner, 1986). During the avascular stage, tumour development can be effectively eliminated by tumour-infiltrating cytotoxic lymphocytes (TICLs) (Loeffler and Ratner, 1989). The TICLs may be cytotoxic lymphocytes, natural killer cells and/or lymphokine-activated killer cells (Lord and Burckhardt, 1984; Wilson and Lord, 1987). The cytostatic/cytotoxic activity of granulocytes and monocytes/macrophages located in the tumour is determined less frequently. In some cases, relatively small tumours are in cell-cycle arrest or there is a balance between cell proliferation and cell death. This steady-state of fully malignant, but regulated, growth could continue for a period of many months or years (Uhr *et al.*, 1997). In many (but not all) cases such a latent form of small numbers of malignant tumours is mediated by cellular immunity (Uhr *et al.*, 1997). Clinically, such latent forms of tumours have been referred to as 'cancer dormancy'.

An important factor, which may influence the outcome of the interactions between tumour cells and TICLs in a solid tumour, is the spatial distribution of the TICLs. A thick shell of lymphoid infiltration is often revealed around the tumour and even near the central hypoxic zone. This would define an internal structure whereby the regions of cell proliferation and cell death alternate, with the TICLs located near the groups of dying tumour cells. In spite of some progress into the investigation of TICLs and their mechanisms of interaction with tumour cells, our understanding of the spatio-temporal dynamics of TICLs in small tumours and in micrometastases *in vivo* is still limited. It is perhaps not surprising, therefore, that this complicated picture has not yet received an adequate explanation. Certainly, other components of the immune system (e.g. cytokines) are involved in modulating the local cellular immune response dynamics. Many cytokines are produced during cell–cell interactions, which can be focused to perform their function over short ranges in space and over short intervals of time. Interestingly, strong local immune reactions are induced by the release of interleukin 1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12, IL-13, IL-15, G-CSF, GM-CSF, INF-α, TNF-α, INF-γ, etc. Each of these cytokines recruits and activates some distinct cell population, which could be tumour-infiltrating cells, or the tumour cells themselves. Besides immune reactions, other processes (e.g. cell proliferation, development, migration and apoptosis) are governed through complicated feedback loops. The experimental analysis of such *in vivo* functions requires gene transfer technology and adoptive cell transfer studies, among others. Such studies are, however, hampered by the availability of *in vitro* cloned TICLs and by the frustrating experience that these *in vitro* propagated killer cells perform relatively badly *in vivo*, owing to, for example, the down-regulation of their homing receptors and the up-regulation of many of the cytokines amongst other things.

Angiogenesis, Vascular Growth, Invasion and Metastasis

As we have seen in the previous section, repeated growth and proliferation of the tumour cells during the early phase of growth leads to the development of an avascular tumour consisting of approximately 10^6 cells. A solid tumour in this state cannot grow any further, owing to its dependence on diffusion processes as the only means of receiving nutrients and removing waste products. For any further development to occur the tumour itself must initiate angiogenesis – the recruitment of blood vessels. The tumour cells first secrete a variety of angiogenic factors which in turn induce endothelial cells in any neighbouring blood vessels to degrade their basal lamina and begin to migrate towards the tumour. As it migrates, the endothelium begins to form sprouts which can then form loops and branches through which blood circulates. From these branches more sprouts form and the whole process repeats, forming a capillary network which eventually connects with the tumour, completing angiogenesis and supplying the tumour with the nutrients it needs to grow further. There is now also the possibility of tumour cells finding their way into the circulation and being deposited in distant sites in the body, resulting in metastasis. The complete process of metastasis involves several sequential steps, each of which must be successfully completed by cells of the primary tumour before a secondary tumour

(a metastasis) is formed. A summary of the key stages of the metastatic cascade is as follows:

- cancer cells escape from the primary tumour;
- they locally degrade the surrounding tissue and continue migration;
- they enter the lymphatic or blood circulation system (intravasation);
- they must survive their journey in the circulation system;
- they must escape from the blood circulation (extravasation);
- the cancer cells (from the primary tumour) must then establish a new colony in distant organs;
- the new colony of cells must then begin to grow to form a secondary tumour in the new organ;
- the complete process of growth, angiogenesis, invasion can then begin all over again.

(See also the chapter *Extracellular Matrix: The Networking Solution*.)

It is highly likely that *in vivo*, the two processes of tumour-induced angiogenesis and degradation of the local surrounding tissue (extracellular matrix) by the tumour cells occur simultaneously. Once the capillary network has reached the tumour, vascular growth occurs and at this stage there is a complex interaction between the vascular network, the invading tumour cells and the extracellular matrix. However, for clarity of purpose, we describe each of the processes of angiogenesis and tumour invasion/matrix degradation separately.

Tumour-induced Angiogenesis

Angiogenesis (syn. neovascularization), the formation of blood vessels from a pre-existing vasculature, is a crucial component of many mammalian growth processes. It occurs in early embryogenesis during the formation of the placenta, after implantation of the blastocyst in the uterine wall. It also occurs, in a controlled manner, in adult mammals during tissue repair. By contrast, uncontrolled or excessive blood-vessel formation, is essential for tumorigenesis and is also observed in arthritis, abnormal neovascularization of the eye, duodenal ulcers and following myocardial infarction (Folkman, 1985, 1995; Folkman and Klagsbrun, 1987). These instances may be considered pathological examples of angiogenesis (Muthukkaruppan *et al*., 1982). In each case, however, the well-ordered sequence of events characterizing angiogenesis is the same, beginning with the rearrangement and migration of endothelial cells from a pre-existing vasculature and culminating in the formation of an extensive network, or bed, of new capillaries (Madri and Pratt, 1986). (See also the chapter *Angiogenesis Models*.)

The first event of tumour-induced angiogenesis involves the cancerous cells of a solid tumour secreting a number of chemicals, collectively known as tumour angiogenic factors (TAF) (Folkman and Klagsbrun, 1987), into the surrounding tissue. These factors diffuse through the tissue space creating a chemical gradient between the tumour and any existing vasculature. Upon reaching any neighbouring blood vessels, endothelial cells lining these vessels are first induced to degrade the parent venule basement membranes and then migrate through the disrupted membrane towards the tumour. Several angiogenic factors e.g. vascular endothelial growth factor (VEGF), acidic and basic fibroblast growth factor (aFGF, bFGF), angiogenin and others, have been isolated (Folkman and Klagsbrun, 1987) and endothelial cell receptors for these proteins have been discovered. Indeed, there is now clear experimental evidence that disrupting these receptors has a direct effect on the final structure of the capillary network (Hanahan, 1997).

The initial response of the endothelial cells to these angiogenic factors is a chemotactic one (Paweletz and Knierim, 1989), whereby the endothelial cells respond to the gradients of TAF and migrate towards the source of these secreted cytokines, the tumour. Following this, small, finger-like capillary sprouts are formed by accumulation of endothelial cells which are recruited from the parent vessel. The sprouts grow in length due to the migration and further recruitment of endothelial cells and continue to move toward the tumour directed by the motion of the leading endothelial cell at the sprout-tip. Further sprout extension occurs when some of the endothelial cells of the sprout-wall begin to proliferate. Cell division is largely confined to a region just behind the cluster of mitotically inactive endothelial cells that constitute the sprout-tip. This process of sprout-tip migration and proliferation of sprout-wall cells forms solid strands of endothelial cells amongst the extracellular matrix. The cells continue to make their way through the extracellular matrix which consists, among other things, of interstitial tissue, collagen fibres, fibronectin and laminin (Liotta *et al*., 1983; Paweletz and Knierim, 1989). Interactions between the endothelial cells and the extracellular matrix are very important and directly affect cell migration. In particular, the specific interactions between the endothelial cells and fibronectin, a major component of the extracellular matrix, have been shown to enhance cell adhesion to the matrix.

Cultured endothelial cells are known to synthesize and secrete cellular fibronectin and the expression of this secreted fibronectin by the endothelial cells in cultures (*in vitro*) closely reflects the distribution of pre-existing fibronectin observed in matrices *in vivo* (Hynes, 1990). The fibronectin which is produced and secreted by endothelial cells does not diffuse, but remains bound to the extracellular matrix (Hynes, 1990), its central function being the adhesion of cells to the matrix. It is a major ligand (both cellular fibronectin and plasma fibronectin) between cells and matrix materials in many situations. Endothelial cells use fibronectin for attachment to the matrix via integrins, a family of cell-surface receptors (Hynes, 1990; Alberts *et al*., 1994). It has been verified experimentally that fibronectin stimulates directional

migration of a number of cell types (including endothelial cells) in Boyden chamber assays. These results have demonstrated that fibronectin promotes cell migration up a concentration gradient and the results of McCarthy and Furcht (1984) have further demonstrated that this is a response of the cells to a gradient of adhesiveness of bound fibronectin, termed haptotaxis (Carter, 1967). Therefore, in addition to the chemotactic response of the endothelial cells to the TAF, there is a complementary haptotactic response to the fibronectin present within the extracellular matrix. (See also the chapter *Modelling Tumour–Tissue Interactions*.)

Initially, the sprouts arising from the parent vessel grow essentially parallel to each other. It is observed that once the finger-like capillary sprouts have reached a certain distance from the parent vessel, they tend to incline toward each other (Paweletz and Knierim, 1989), leading to numerous tip-to-tip and tip-to-sprout fusions known as anastomoses. Such anastomoses result in the fusing of the finger-like sprouts into a network of poorly perfused loops or arcades. Following this process of anastomosis, the first signs of circulation can be recognized and from the primary loops, new buds and sprouts emerge repeating the angiogenic sequence of events and providing for the further extension of the new capillary bed. The production of new capillary sprouts from the sprout-tips is often referred to as sprout branching, and as the sprouts approach the tumour, their branching dramatically increases until the tumour is eventually penetrated, resulting in vascularization (Muthukkaruppan *et al*., 1982).

Tumour-induced angiogenesis provides the crucial link between the avascular phase of solid tumour growth and the more harmful vascular phase, wherein the tumour invades the surrounding host tissue and blood system (Chaplain, 1996). However, these apparently insidious features of tumour-induced angiogenesis are now being used to combat cancer growth and the clinical importance of angiogenesis as a prognostic tool is now well recognized. Anti-angiogenesis strategies are also being developed as a potentially powerful, noninvasive weapon against the spread of cancer.

Matrix Degradation and Cancer Invasion

A crucial part of the invasive/metastatic process is the ability of the cancer cells to degrade the surrounding tissue or extracellular matrix (ECM) (Liotta *et al*., 1983; Stetler-Stevenson *et al*., 1993). As mentioned in the previous section, the extracellular matrix is a complex mixture of macromolecules, some of which, like the collagens, are believed to play a structural role and others, such as laminin, fibronectin and vitronectin, are important for cell adhesion, spreading and motility. We note that all of these macromolecules are bound within the tissue, i.e. they are nondiffusible. The ECM can also sequester growth factors and itself be degraded to release fragments which can have growth-promoting activity. Thus, while ECM may have to be physically removed in order to allow a tumour to spread or intra- or extravasate, its degradation may in addition have biological effects on tumour cells.

A number of matrix degradative enzymes (MDEs) such as the plasminogen activator (PA) system and the large family of matrix metalloproteinases (MMPs) have been described (Matrisian, 1992; Mignatti and Rifkin, 1993). While no MDE is completely specific for one element of the ECM, some broad preferences are expressed, for example the gelatinases (two members of the MMP family) preferentially cleave the laminar collagens IV and V and denatured fibrillar collagens I, II and III but can also digest vitronectin and laminin, at least *in vitro*. Both PAs and the MMPs have been repeatedly implicated in all of the above steps of tumour invasion and metastasis (Stetler-Stevenson *et al*., 1996; Chambers and Matrisian, 1997). Regulation of matrix-degradative activity is highly complex. In both these enzyme systems (PAs/MMPs) there exist several endogenous inhibitors (Kleiner and Stetler-Stevenson, 1993), and the enzymes are often secreted as inactive precursors which must themselves be partially degraded to reach full activity. More than one cell type may be involved in the activation of any one enzyme (Kleiner and Stetler-Stevenson, 1993). (See also the chapter *Invasion and Metastasis*.)

Having described in some detail the key processes that we are going to model using mathematical techniques, we now turn our attention to the aspect of the mathematical modelling itself. In the following sections, we introduce and describe several mathematical models for the growth of an avascular solid tumour, the immune response to cancer, tumour-induced angiogenesis and tumour invasion/metastasis.

PRINCIPLES OF MATHEMATICAL MODELLING

It is difficult to control experimentally all of the interacting elements in a tumour. Furthermore, complex biological systems, such as the immune system and a cancer *in vivo*, do not always behave or act as predicted by experimental investigations *in vitro* (Prehn, 1994). Mathematical modelling and computer simulations of these models can be helpful in understanding important features in these complex inter-related systems. In order to understand how one goes about formulating a mathematical model and then using it to shed light on a given complex biological process, it is useful to have in mind the following algorithm:

- identify the key biological processes and the key biological variables (e.g. various tumour cell types, cytokines, matrix macromolecules);
- identify interactions between these variables (e.g. which cells produce which cytokines, etc.);
- translate these interactions into a mathematical model using the appropriate mathematical tools (e.g. ordinary or partial differential equations; cellular automata);

- carry out computer simulations of the model using, where possible, experimentally measured parameter values (e.g. cytokine production rates, cell migration rates);
- Analyse the results of the model which should, if the model is a good one, at least reproduce observed events. A good model will, of course, go further and will provide extra information on (perhaps only qualitative) or explanations of how the biological process functions.

While the first four points of the above algorithm are relatively straightforward, and indeed so is the first statement of the fifth point, it is the final step of producing a model which sheds extra light on a process which is the most challenging. However, once a reliable model has been developed, it is then relatively easy to perform a range of 'mathematical experiments' on the system (e.g. by making changes to the system parameters) and test various hypotheses associated with the system or to examine the risk associated with these changes.

In developing mathematical models of cancer growth, we are interested in particular in the rates of change of variables with time and how the amount (e.g. concentration, density) of the variables change from point to point in space. The changes in the variables that we study depend on their spatial position and on time and mathematically we use tools known as differential equations to describe these changes.

A variable, V, say (e.g. concentration of a cytokine or the number of cancer cells per unit area/volume, i.e. cell density), will therefore be a function of its spatial coordinates (x, y, z) and time t. We denote this mathematically by writing $V(x, y, z, t)$. To construct a differential equation model we make use of the calculus (which enables us to describe the rate of change of variables) and the law of the conservation of mass. Applying these two rules we can formulate a generic 'conceptual' equation describing the rate of change of a substance with respect to time, its migration and its production/decay. Thus we have:

$$\text{rate of change of variable} = \text{spatial migration} + \text{production} - \text{decay/uptake}$$

By solving the above equation mathematically, at any given time one can predict how much of the variable (e.g. number of cells, amount of cytokine) is present in one's system. One may draw an analogy between the above equation describing a biological population (e.g. of cells, amount of a cytokine) and the population of a country – at any given time the number of people in the country can be calculated from the number of immigrants/emigrants (cf. migration/diffusion) and the number of births/deaths (cf. production/decay).

APPLICATIONS OF MATHEMATICAL MODELLING TO TUMOUR GROWTH

In the following sections we present three specific mathematical models which examine the three important processes of the immune response to cancer, tumour-induced angiogenesis and tumour invasion and metastasis. In each case, the mathematical model is an attempt to capture the essential biological processes, while retaining a certain simplicity. Each process has already been described in some detail in the first section and the reader can refer back to this section if necessary to follow the details of the mathematical equations. In addition, full details of each model can be found in certain key references which are cited.

Modelling the Immune Response to Cancer

We consider a simplified process of a small, growing, avascular tumour which produces signals (e.g. secretes cytokines) which attract a population of lymphocytes. These signals are recognized by lymphocytes and the tumour cells (TCs) are then directly attacked by tumour-infiltrating cytotoxic lymphocytes (TICLs) (Ioannides *et al.*, 1993). The humoral immune response, the interactions between cells of the immune system and dynamics of cytokines in the tumour, will be accounted for indirectly, because data related to these processes are often contradictory and many important mechanisms of cytokine dynamics in tumours are unknown. Local interactions between TICLs and tumour cells *in vivo* may be described by the following simplified kinetic scheme (see Kuznetsov *et al.* (1994) and Chaplain *et al.* (1998) for full details):

$$E + T \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \quad \begin{array}{l} \xrightarrow{k_2 p} E + T^* \\ \xrightarrow{k_2(1-p)} E^* + T \end{array}$$

where E, T, C, E^* and T^* are the local concentrations of TICLs, TCs, TICL–TC complexes, inactivated TICLs and 'lethally hit' or 'dead' TCs, respectively. The parameters k_1, k_{-1} and k_2 are non-negative kinetic constants defined as follows: k_1 and k_{-1} describe the rate of binding of TICL to TC and detachment of TICL from TC without damaging cells; k_2 is the rate of detachment of TICL from TC, resulting in either irreversibly programming TC for lysis with probability p or inactivating TICLs with probability $(1-p)$. Substantiation of the above scheme has already been discussed at greater length by Kuznetsov *et al.*,

(1994). It can be shown (Chaplain *et al.*, 1998) that it is sufficient to study the interaction of the three populations T, E and C (the other two can then be deduced from these). Using the law of mass action, and under the assumption that the tumour cells and the TICLs can diffuse, but the complexes cannot, the above scheme can be translated into the following system of differential equations:

$$\frac{\partial E}{\partial t} = \overbrace{\frac{\partial^2 E}{\partial x^2}}^{\text{diffusion}} + \overbrace{\sigma + \frac{\rho ET}{\eta + T}}^{\text{recruitment}} - \overbrace{\sigma E}^{\text{death}} - \overbrace{\mu ET + \epsilon C}^{\text{complexing}} \qquad [1]$$

$$\frac{\partial T}{\partial t} = \omega \overbrace{\frac{\partial^2 T}{\partial x^2}}^{\text{diffusion}} + \overbrace{\alpha(1-\beta T)T}^{\text{proliferation}} - \overbrace{\phi ET + \lambda C}^{\text{complexing}} \qquad [2]$$

$$\frac{\partial C}{\partial t} = \mu ET - \chi C \qquad [3]$$

In addition to kinetic interactions, we have also assumed that both the TICLs and the tumour cells can diffuse throughout the tissue and this cell migration is represented by the first two terms in eqns [1] and [2], respectively. The other terms in the model, in addition to those derived from the kinetic interaction scheme, account for (1) a recruitment of TICLs from the host tissue, under the assumption that the host immune response provides a constant supply σ of TICLs; (2) an additional recruitment of TICLs modelling the migratory response of the TICLs to the secretion of cytokines by the complexes (represented by the term $\rho ET/(\eta+T)$; (3) the natural death of the TICLs (represented by the term $-\sigma$); and (4) the proliferation of the tumour cells (represented by the logistic growth function $\alpha(1-\beta T)T$).

All of the parameters in the model equations can be estimated from actual experiments (Kuznetsov *et al.*, 1994; Chaplain *et al.*, 1998). With this information one can analyse eqns [1]–[3] numerically and simulate the interactions between a small avascular tumour and the immune system represented by lymphocyte infiltration. The results are presented in the next section.

Interpretation of Model Results

We simulated eqns [1]–[3] in a one-dimensional spatial domain $0 \leq x \leq 1$ representing a unit of tissue. **Figure 1** shows the initial conditions used in the model simulations. The tumour cells are initially located in a small cluster representing a small avascular tumour in the left-hand region of the domain between $x = 0$ and 0.2. The TICLs are initially located in the right-hand region of the domain between $x = 0.2$ and 1. Between the two cell populations is an initial group of TICL–TC complexes.

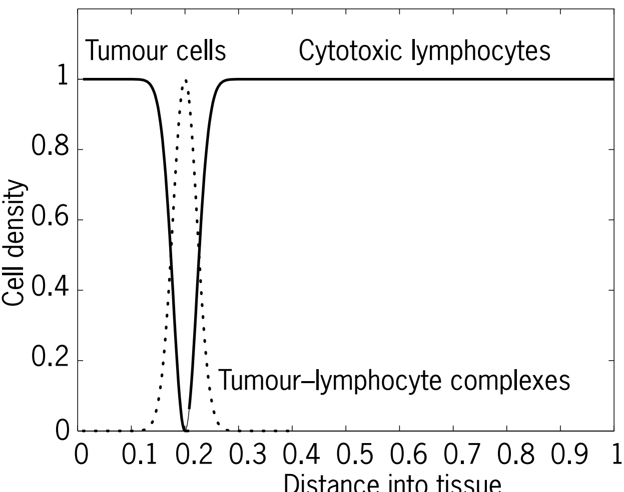


Figure 1 Initial conditions used for the tumour cells, tumour–lymphocyte complexes and the tumour-infiltrating cytotoxic lymphocytes. The tumour cells are initially confined to the left-hand side of the domain between $x = 0$ and 0.2 representing a small tumour; the cytotoxic tumour-infiltrating lymphocytes initially occupy the right-hand region of the domain between $x = 0.2$ and 1. The tumour–lymphocyte complexes are initially distributed between the two and are shown by the dotted curve.

Figures 2, **3** and **4** show the distributions of each variable (TICL, TC and complexes) throughout the domain at increasing times. One can see that as the cells interact and migrate, they produce rich spatio-temporal dynamics resulting in a highly heterogeneous distribution of lymphocytes, tumour cells and complexes throughout tissue. This heterogeneity is certainly observed in actual small solid tumours, infiltrated by lymphocytes (cf. Kuznetsov *et al.*, 1994).

Figures 5 and **6** show the (scaled) total number of lymphocytes and tumour cells within the tissue. From these results, we clearly see that, although these is a very heterogeneous spatial distribution of cells in the tissue, the total number of lymphocytes increases over time to a maximum, while the total number of tumour cells decreases to a minimum. This results in the tumour being kept under control by the immune system (cf. tumour dormancy). However, one can show (results not presented here) that by changing one of the key parameters of the model by a small amount (the parameter P), the tumour escapes the immune system control and completely invades the tissue, increasing its total number of cells as it proceeds. This result has certain implications for maintaining tumour dormancy via immunotherapy treatment (cf. Uhr *et al.*, 1997).

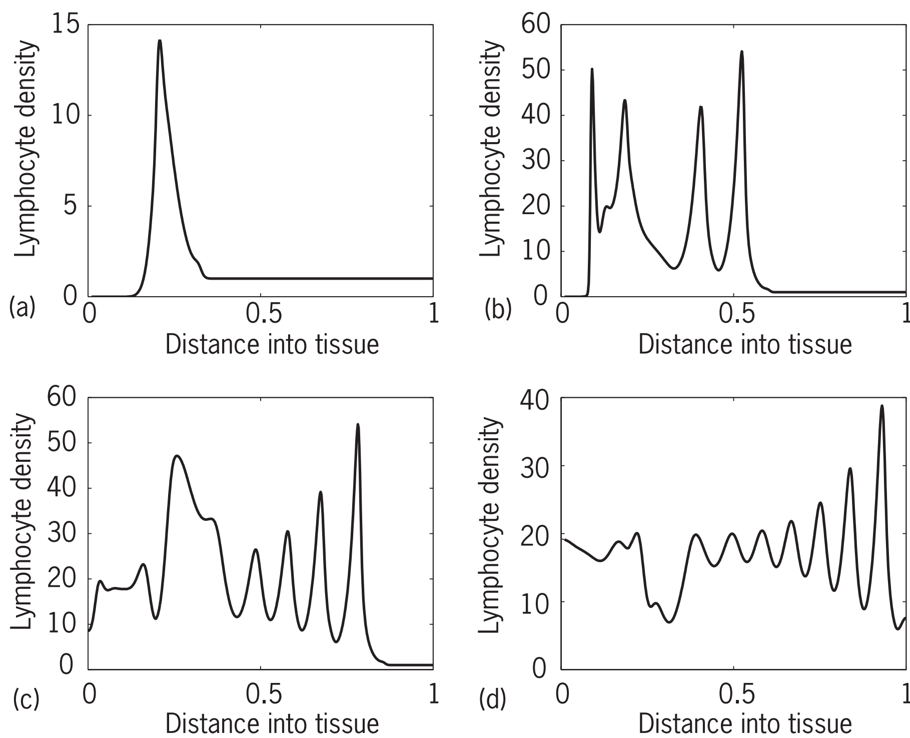


Figure 2 Spatio-temporal evolution of the tumour-infiltrating cytotoxic lymphocytes (TICLs) within the tissue at times corresponding to (a) 100, (b) 400, (c) 700 and (d) 1000 days, respectively. As time evolves, the interactions between the lymphocytes and the tumour cells become more complicated and the lymphocyte distribution within the tissue is highly heterogeneous and almost 'spatially chaotic'.

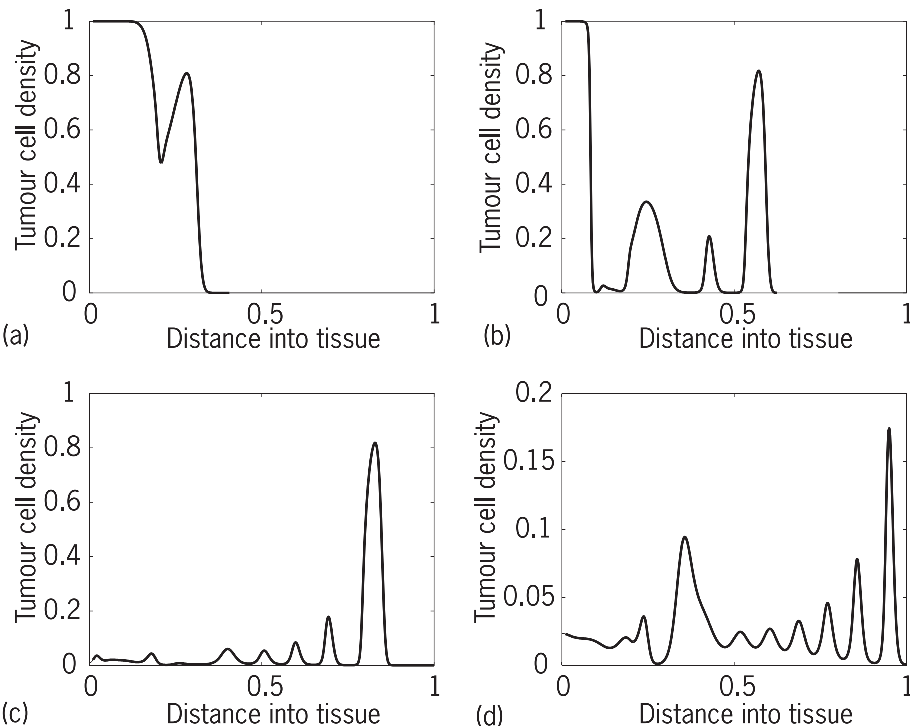


Figure 3 Spatio-temporal evolution of the tumour cells within the tissue at times corresponding to (a) 100, (b) 400, (c) 700 and (d) 1000 days, respectively. The tumour cells initially invade the tissue moving from left to right and are then attacked by the lymphocytes (see **Figure 2**). The complicated spatial interactions between the tumour cell and the lymphocytes result in isolated groups or solitary pulses of tumour cells invading the tissue rather than the whole tumour mass.

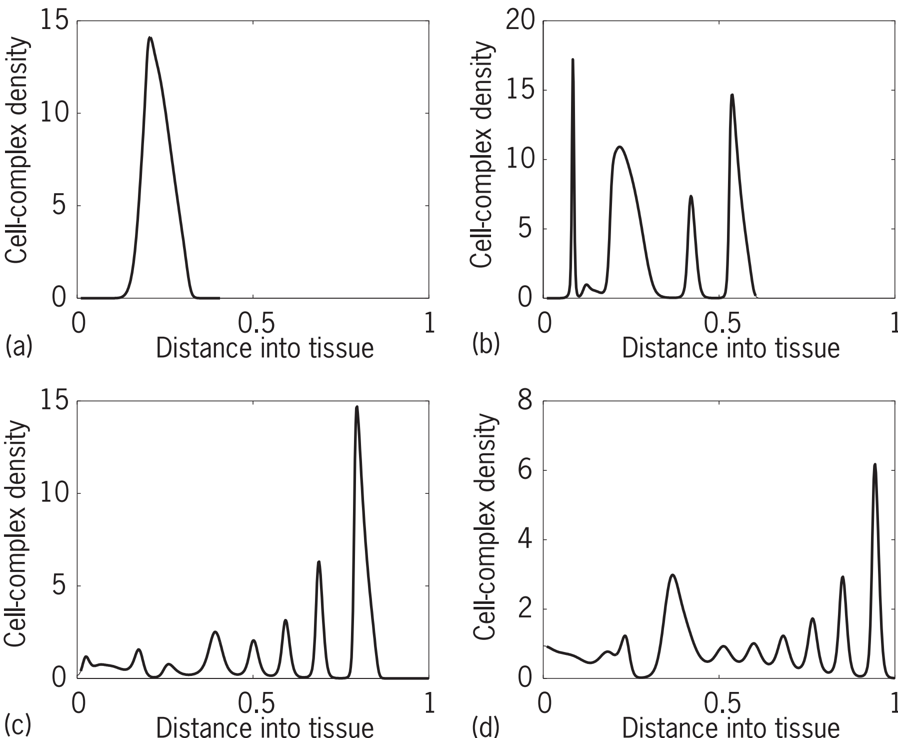


Figure 4 Spatio-temporal evolution of the tumour cell-TICL complexes within the tissue corresponding to the evolution of both TICLs and tumour cells in the previous two figures. The times correspond to (a) 100, (b) 400, (c) 700 and (d) 1000 days, respectively.

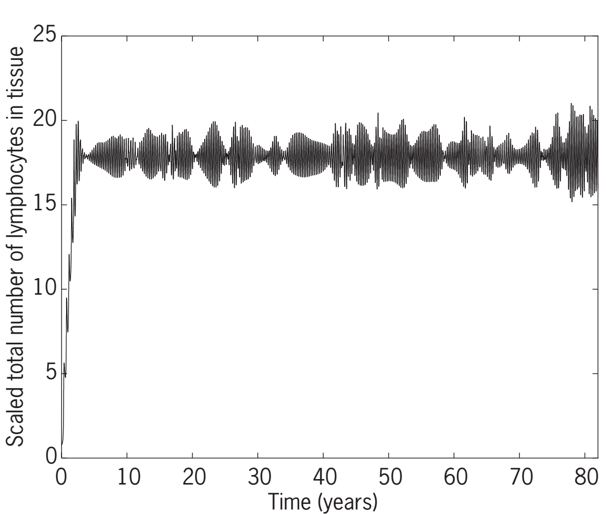


Figure 5 Plot of the total number of lymphocytes within the tissue over a period of 80 years. As can be seen, the lymphocyte number initially increases from a relatively small number to a stable but oscillatory distribution around 18. At this level the immune response is sufficient to suppress the growth of the tumour, as can be seen from the results in **Figure 6**.

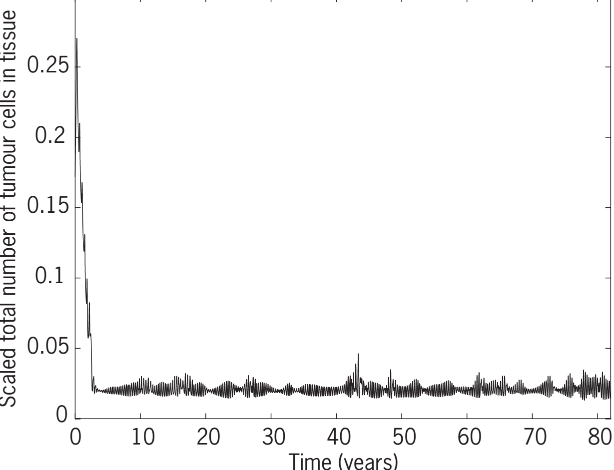


Figure 6 Plot of the total number of tumour cells within the tissue over a period of 80 years. The number of tumour cells initially decreases as the tumour is attacked and is infiltrated by the cytotoxic lymphocytes. Eventually the total number of tumour cells settles down to an oscillatory behaviour around a stationary level of approximately 0.02. Although the tumour is not eradicated completely, this simulation represents the case where the immune response is sufficient to keep the tumour in a harmless, dormant state.

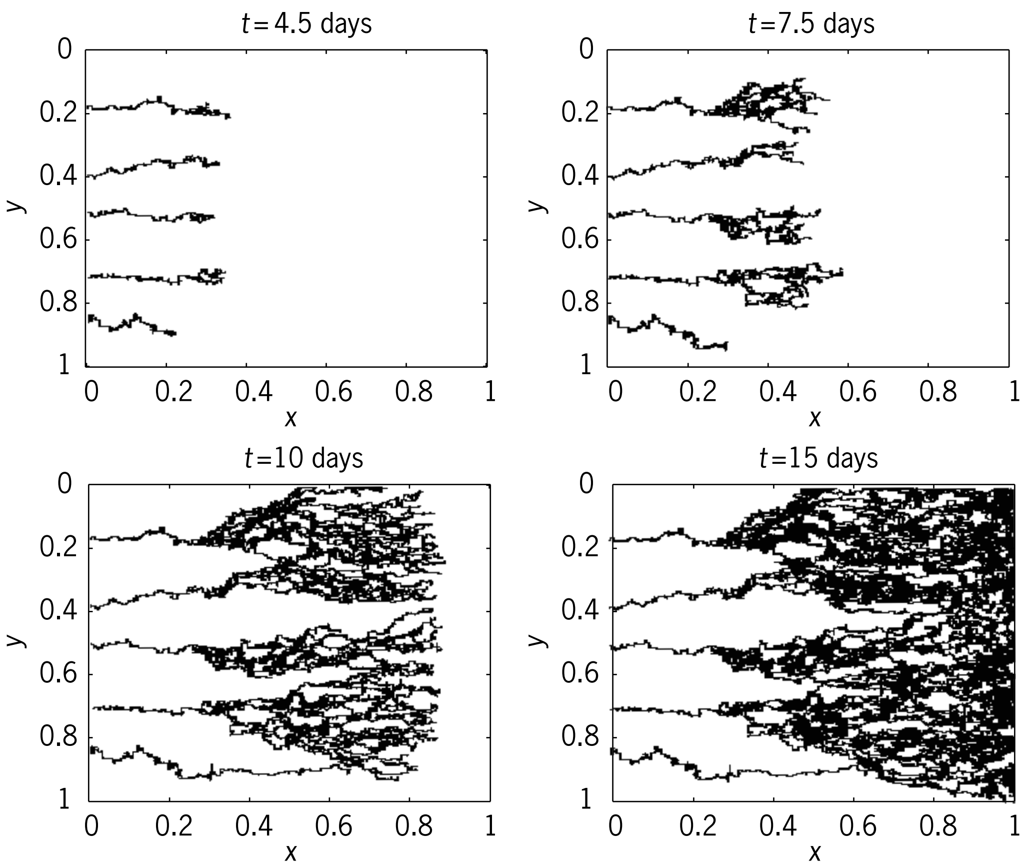


Figure 7 Spatio-temporal evolution of a theoretical capillary network generated from the mathematical model. The figure shows the migration of endothelial cells at sprout-tips in response to a line source of tumour cells. The parent vessel is located along the left-hand boundary while the tumour cells are located along the right-hand boundary. The cells respond chemotactically to TAF gradients and haptotactically to fibronectin gradients as well as proliferating. As the sprouts grow and migrate, a branched, connected network is formed which connects with the tumour between a time of 10 and 15 days and vascularization is achieved.

Modelling Tumour-induced Angiogenesis

The mathematical model in this section describes the interactions between endothelial cells n, angiogenic factors (cytokines) c and an adhesive matrix macromolecule such as fibronectin f. Full details of the model can be found in papers by Chaplain and Anderson (1997, 1999), Anderson and Chaplain (1998) and Schor *et al*. (1999). The basic assumptions of the model are as follows: (1) the endothelial cells migrate randomly, respond to the cytokines via chemotaxis and respond to the matrix macromolecule via haptotaxis; (2) the cytokine is produced by tumour cells and diffuses into the matrix where it is taken up by endothelial cells through binding; (3) fibronectin is present in the matrix, is synthesized by endothelial cells and is bound by the cells via integrins; there is also some degradation due to degradative enzymes. (See also the chapters *Extracellular Matrix: The Networking Solution*; *Angiogenesis Models*.)

The complete system of equations describing the interactions of the endothelial cells, the angiogenic factors and fibronectin is therefore given by

$$\frac{\partial n}{\partial t} = \overbrace{D_n \nabla^2 n}^{\text{random motility}} - \overbrace{\chi \nabla \cdot \left(\frac{n}{k_1 + c} \nabla c \right)}^{\text{chemotaxis}} - \overbrace{\rho \nabla \cdot (n \nabla f)}^{\text{haptotaxis}} \qquad [4]$$

$$\frac{\partial f}{\partial t} = \overbrace{\omega n}^{\text{production}} - \overbrace{\mu n f}^{\text{binding/degradation}} \qquad [5]$$

$$\frac{\partial c}{\partial t} = - \overbrace{\lambda n c}^{\text{uptake}} \qquad [6]$$

The model as presented describes the so-called basic ‘tissue response unit’ (Chaplain and Anderson, 1999; Schor *et al*., 1999), whereby individual cells respond to

their local tissue milieu through interactions with cytokines and matrix macromolecules. Once again, as with the model of the previous section, all parameters in the model are estimated from experimental data found in the literature. One can use the model to simulate the basic interactions between the endothelial cells, the matrix and a generic angiogenic cytokine. The mathematical technique of discrete biased random walks pioneered by Chaplain and Anderson (Chaplain and Anderson, 1997; Anderson and Chaplain, 1998) enables us to follow the paths of individual endothelial cells. This technique also enables us to include explicitly in the model crucial processes such as sprout branching, anastomosis and localized cell proliferation. The theoretical results of our numerical simulations are presented in the next section.

Interpretation of Model Results

The following figures show the theoretical capillary networks generated and predicted by the mathematical model of eqns [4]–[6].

Figure 7 shows the evolution of a theoretical capillary network at various stages in its development between 4 and 15 days. The parent vessel (cf. a limbal vessel) is located along the left-hand edge of the domain, while the tumour cells are located along the right-hand edge. The size of the domain is 2.5×2.5 mm. As can be seen, the capillary sprouts grow, elongate, branch and fuse together as they progress through the tissue towards the tumour cells before they finally connect up in the final figure. **Figure 8** shows the evolution of a theoretical capillary network in response to a small circular tumour implant. All parameter values were the same as those used to generate the results in **Figure 7**. The only difference is the 'shape' of the underlying concentration field of angiogenic factor.

These two figures are a mathematical simulation/representation of actual experiments carried out using the corneal assay technique (Muthukkaruppan *et al.*, 1982).

Figures 9 and **10** show the results of the model in three spatial dimensions in a cube of tissue of size

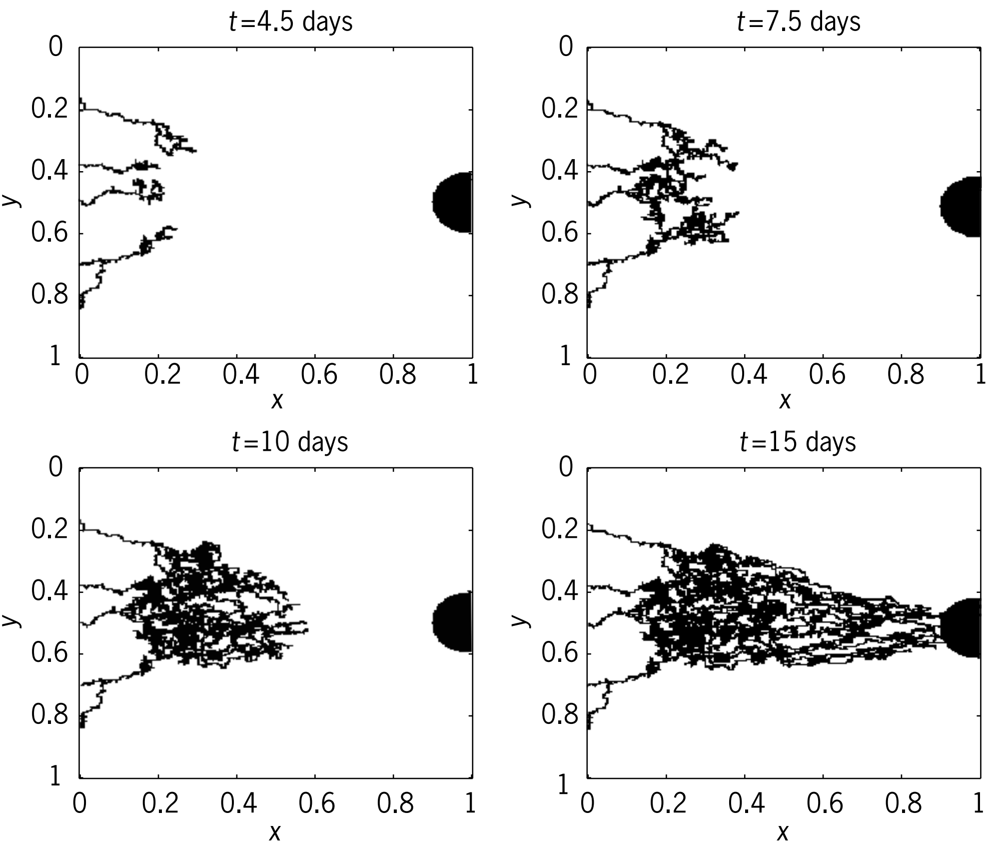


Figure 8 Spatio-temporal evolution of a theoretical capillary network generated from the mathematical model. The figure shows the migration of endothelial cells at sprout-tips in response to a circular source of tumour cells. The parent vessel is once again located along the left-hand boundary while the circular tumour implant is shown in black at the opposite boundary. The cells respond chemotactically to TAF gradients and haptotactically to fibronectin gradients as well as proliferating. As the sprouts grow and migrate, a branched, connected network is formed. Once again the network connects with the tumour around 15 days and vascularization is achieved.

$2.5 \times 2.5 \times 2.5$ mm. In these cases, the capillary networks generated are theoretical representations of the actual *in vivo* situation. In **Figure 9**, the network responds to a flat layer of tumour cells on the bottom face of the cube of tissue. In **Figure 10**, the network is reponding to a central spherical mass of tumour cells, representing a small, spherical tumour.

In all cases, the model generates quantitatively realistic networks in terms of the morphology, structure and growth rate.

Once again, one can change certain key parameters of the model (e.g. chemotactic response, haptotactic response of the endothelial cells) and can perform anti-angiogenic experiments (Orme and Chaplain, 1997). One can also use the model to calculate important variables such as flow rate through the network, vessel length, network area/volume and the fractal dimension of the network. All these variables are important in designing drugs to target tumours (Jain *et al.*, 1997) and, with some refinement, the model could be implemented to optimize anti-angiogenic strategies in terms of drug design or drug scheduling.

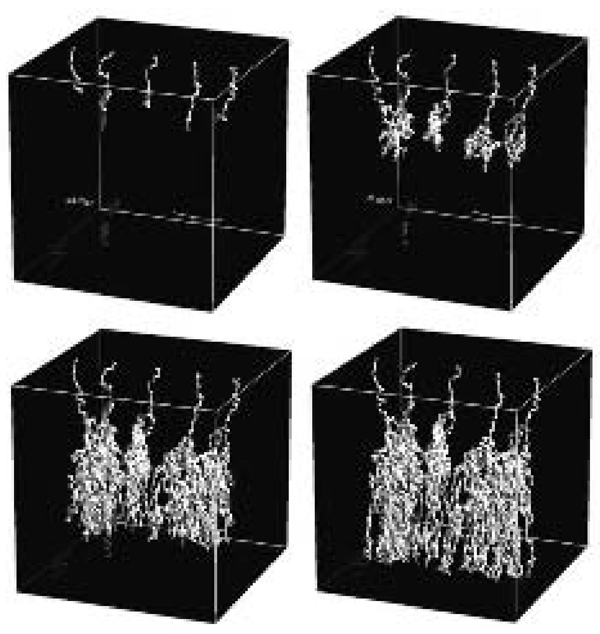


Figure 9 Spatio-temporal evolution of a three dimensional capillary network towards a single layer (plane) of tumour cells positioned on the bottom face of the cube. There is a gradient of TAF within the tissue, with a high concentration on the bottom face and a lower concentration on the top face. The endothelial cells respond chemotactically to this gradient as well as gradients of fibronectin in the tissue. A branching network is formed which eventually connects with the tumour on the bottom face.

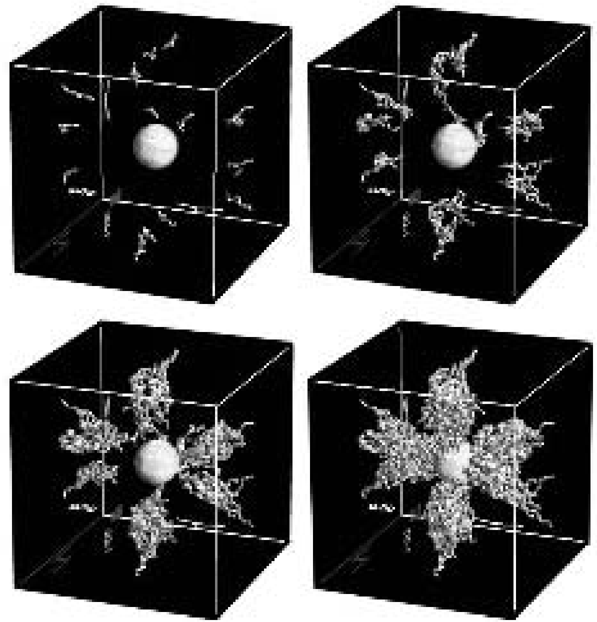


Figure 10 Spatio-temporal evolution of a three-dimensional capillary network towards a central spherical source of tumour cells (cf. avascular tumour or multicell spheroid). There is a gradient of TAF within the tissue, with a high concentration around the central spheroid decaying to lower concentrations in a radially symmetric manner on the boundaries. Once again the endothelial cells respond chemotactically to this gradient as well as gradients of fibronectin in the tissue. A branching network is formed which eventually connects with the avascular tumour in the centre, thus providing the tumour with extra nutrients and opening the way for tumour cells to escape into the bloodstream.

Modelling Tumour Invasion and Metastasis

The mathematical model in this section describes the basic interactions between tumour cells and the extracellular matrix during the process of tissue invasion. Full details of the model can be found in Anderson *et al.* (2000). The key variables in the model are the tumour cells n, matrix degrading enzymes e (MDE), and the matrix (ECM) m. The enzymes are assumed to be produced by the tumour cells, diffuse into the tissue and degrade the matrix; the tumour cells are assumed to migrate randomly and also via haptotaxis in response to gradients of the enzymes; and finally, the matrix is passively degraded by the enzymes. Hence the complete system of equations describing the interactions of the tumour cells, ECM and MDE is as

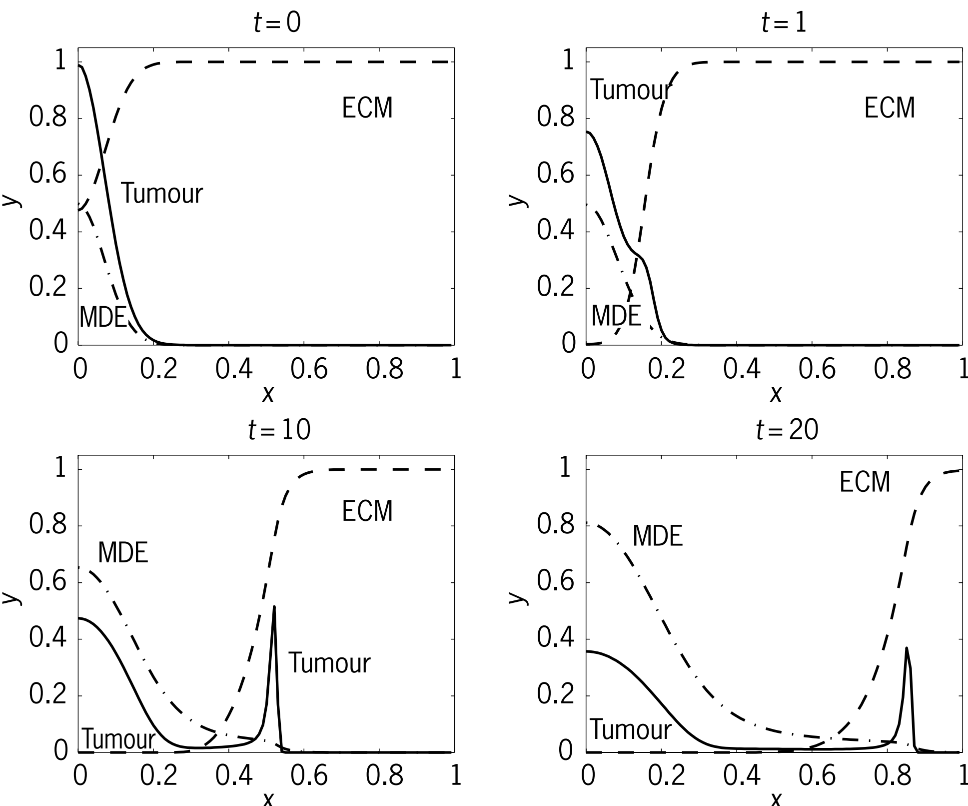


Figure 11 Spatio-temporal evolution of tumour cells invading tissue while secreting matrix-degrading enzymes. As the enzymes degrade the tissue (ECM) the tumour cells invade and form two distinct clusters, those mainly driven by diffusion and those driven by haptotaxis. The cluster of tumour cells at the front may break away completely from the original mass and either go undetected in surgery or connect with the lymph/circulatory systems and then go on to form metastases.

follows:

$$\frac{\partial n}{\partial t} = \overbrace{D_n \nabla^2 n}^{\text{random motility}} - \overbrace{\rho \nabla \cdot (n \nabla m)}^{\text{haptotaxis}} \quad [7]$$

$$\frac{\partial m}{\partial t} = - \overbrace{\delta e m}^{\text{degradation}} \quad [8]$$

$$\frac{\partial e}{\partial t} = \overbrace{D_e \nabla^2 e}^{\text{diffusion}} + \overbrace{\mu n}^{\text{production}} - \overbrace{\lambda e}^{\text{decay}} \quad [9]$$

(See also the chapters *Invasion and Metastasis*; *Modelling Tumour–Tissue Interactions*.)

As in the previous two models, once again we estimate the model parameters from the experimental literature. Using the model, we can therefore simulate the basic interaction between the cancer cells, the matrix and a generic matrix-degrading enzyme, e.g. a matrix metalloproteinase.

Interpretation of Model Results

The following figures show the theoretical invasive spread of cancer cells as predicted by the mathematical model of eqns [7]–[9].

Figure 11 shows the migration of a group of tumour cells as it invades a one-dimensional domain $0 \le x \le 1$ representing the host tissue. As can be seen as time evolves, the matrix is degraded by the enzyme, creating a gradient in matrix concentration. Through haptotaxis (directed migration in response to the gradient of matrix) a portion of the tumour almost 'breaks away' from the main tumour mass and continues to invade as a solitary front while a central mass of tumour is left behind.

Figures 12 and **13** show the results of model simulations in two space dimensions. In these figures we use the model to track the paths of individual invading cancer cells. Initially a small central mass of cancer cells are in the centre of the domain, surrounded by tissue (ECM). In the early stages of the process, all cells remain as a compact central mass but, as time goes on, individual cells break free from the central mass and penetrate more deeply into the tissue. These cells could potentially be missed by a surgeon during any operation to remove the primary tumour and could therefore lead either to recurrence of the primary cancer or escape and lead directly to metastases.

The results of the model show that the invasive capability of a cancer is predicted by the model and that cell–matrix interactions are crucial in this process. Once

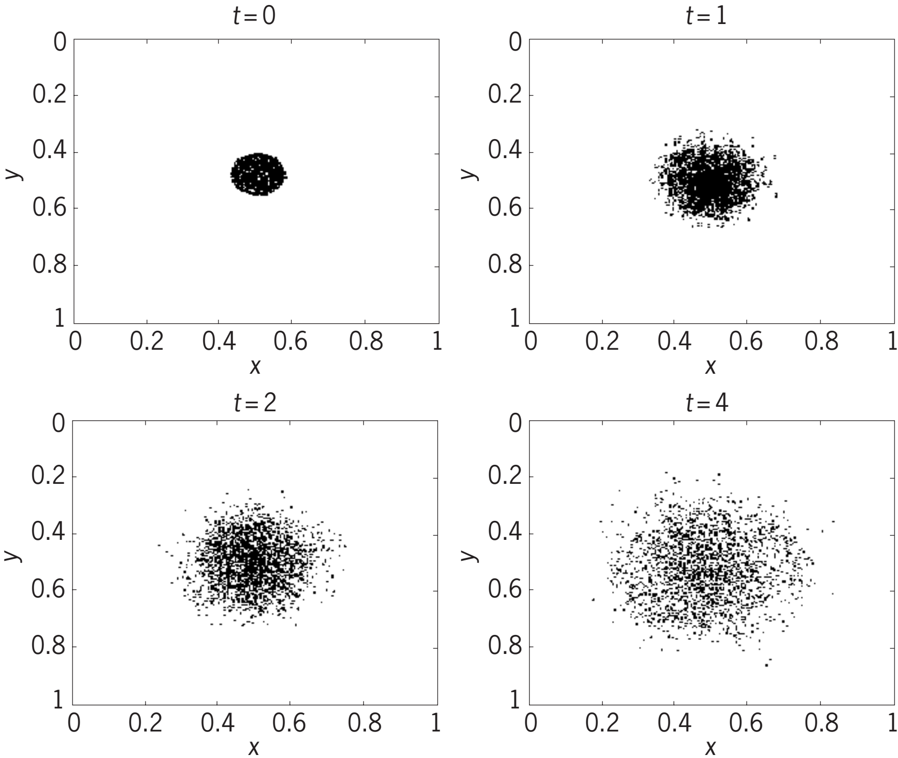


Figure 12 Spatio-temporal evolution of individual tumour cells invading tissue. The figure shows the early stages of invasion with tumour cells migrating from the centre ($x = 0.5$, $y = 0.5$) outwards into the tissue (ECM).

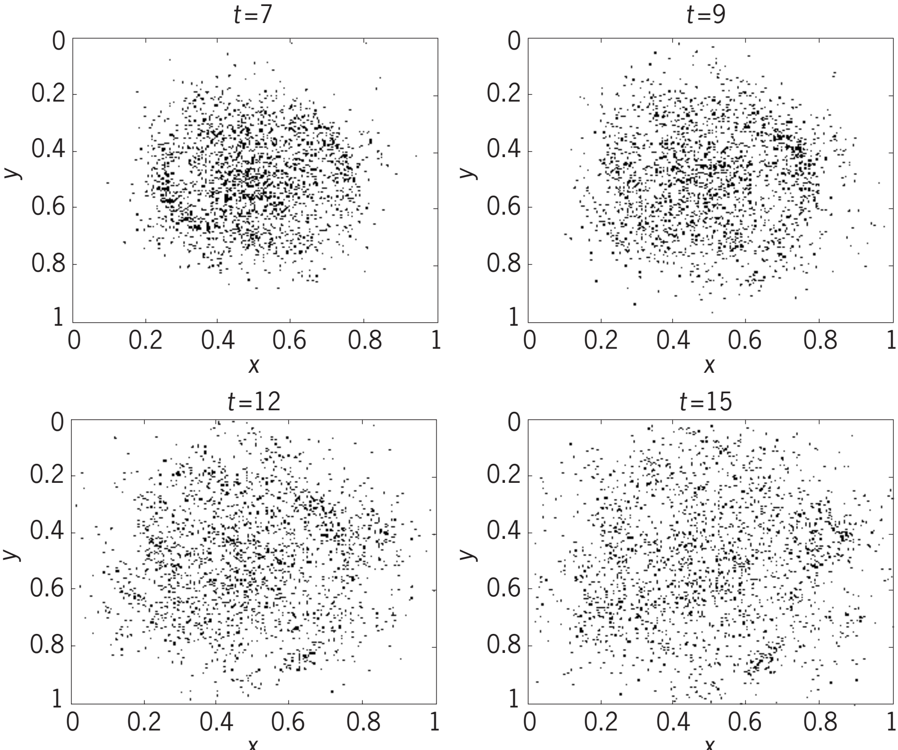


Figure 13 Spatio-temporal evolution of individual tumour cells invading tissue as in **Figure 12** but at later times. As can be seen, a certain number of individual cells break away from the main tumour mass and penetrate the tissue more deeply; some in fact reach the boundaries of the domain. These cells may then either go undetected in surgery or connect with the lymph/circulatory systems and then go on to form metastases.

again, by changing key parameters we can perform theoretical anti-invasive experiments, or predict how far the leading cancer cells penetrate into the tissue. This information could then be used in deciding how much tissue to resect in surgery.

PERSPECTIVES

In this chapter, we have attempted to show how mathematical techniques can be used to model certain key processes and key stages of cancer growth. The first model examined the immune response to cancer and showed that under certain conditions, the host immune response could be sufficient to suppress a small avascular tumour. However, although the tumour remained dormant in the sense that the total number of tumour cells was kept under control, the actual spatial distribution of the cells throughout the tissue was very heterogeneous and complex. Moreover, with a small change in one of the system parameters, the tumour would continue to grow and escape the immune response. More detailed analysis of the model may shed light on how to treat tumours using optimal immunotherapy techniques.

In the second model, the important process of tumour-induced angiogenesis was examined via a mathematical model. Using recently developed mathematical techniques, we were able to study the growth and form of a capillary network by tracking the motion of individual endothelial cells. Using experimental parameter values in our model and incorporating explicitly the processes of sprout branching, anastomosis and localized cell proliferation, we were able to generate quantitatively accurate structures in both two and three spatial dimensions. There is great potential for developing this model further and creating a 'mathematical angiogenesis assay'. This could be used to study and measure important quantities such as flow rates through the network (cf. chemotherapeutic drugs, anti-angiogenic drugs), network growth rate, network area/volume and network connectivity. Knowledge of these variables could then be used to optimize chemotherapy regimes.

In the third model, we examined the process of tumour invasion of the surrounding tissue. Once again, our mathematical model enabled us to focus on the key processes involved in invasion and to make certain predictions. The model enabled us to track individual cancer cells as they secreted matrix-degrading enzymes and migrated through the host tissue. The model naturally captured one of the key features of cancer invasion – that of the break-off of small clumps of cancer cells from the central mass. The model showed that the key process enabling this to happen was haptotaxis and so directed cell migration and therefore cell–matrix interactions are of crucial importance. In developing the model further we will be investigating ways of trying to localize the tumour cells and prevent small clusters from breaking away from the central mass; we will also be investigating how to predict the maximum distance that individual cells may penetrate into the tissue away from the central mass. Such investigations may eventually lead to more accurate surgery being carried out, with a greater survival rate for patients.

All the mathematical models presented are necessarily a simplification of the actual processes they seek to describe. Nevertheless, even in their present form they each shed light on complicated 'nonlinear' processes and make predictions about these processes. As new and more powerful mathematical and computational techniques become available, so the accuracy and predictive power of the models will increase, which in turn will bring them to bear on the clinical decision-making process and eventually benefit patients.

ACKNOWLEDGEMENTS

We gratefully acknowledge support of BBSRC grant MMI09008 and EPSRC Visiting Fellowship Grant GR/M29849.

REFERENCES

Alberts, B., *et al*. (1994). *The Molecular Biology of the Cell*, 3rd edn. (Garland Publishing, New York).

Anderson, A. R. A. and Chaplain, M. A. J. (1998). Continuous and discrete mathematical models of tumour-induced angiogenesis. *Bulletin of Mathematical Biology*, **60**, 857–899.

Anderson, A. R. A., *et al*. (2000). Mathematical modelling of tumour invasion and metastasis. *Journal of Theoretical Medicine*, **2**, 129–154.

Carter, S. B. (1967). Haptotaxis and the mechanism of cell motility. *Nature*, **213**, 256–260.

Chambers, A. F. and Matrisian, L. M. (1997). Changing views of the role of matrix metalloproteinases in metastasis. *Journal of the National Cancer Institute*, **89**, 1260–1270.

Chaplain, M. A. J. (1996). Avascular growth, angiogenesis and vascular rowth in solid tumours: the mathematical modelling of the stages of tumour development. *Mathematical and Computer Modelling*, **23**, 47–87.

Chaplain, M. A. J. and Anderson, A. R. A. (1997). Mathematical modelling of tumour-induced angiogenesis. *Invasive Metastasis*, **16**, 222–234.

Chaplain, M. A. J. and Anderson, A. R. A. (1999). On the growth and form of capillary networks. In: Chaplain, M. A. J., *et al*. (eds), *On Growth and Form: Spatio-temporal Pattern Formation in Biology*, 225–250 (Wiley, Chichester).

Chaplain, M. A. J., *et al*. (1998). Spatio-temporal dynamics of the immune system response to cancer. In: Horn, M. A., *et al*. (eds), *Proceedings of the Mathematical Models in Medical and Health Sciences Conference*. (Vanderbilt University Press, Nashville, TN).

Folkman, J. (1985). Tumor angiogenesis. *Advances in Cancer Research*, **43**, 175–203.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Medicine*, **1**, 21–31.

Folkman, J. and Klagsbrun, M. (1987). Angiogenic factors. *Science*, **235**, 442–447.

Franks, L. M. and Teich, N. (eds) (1986). *Introduction to the Cellular and Molecular Biology of Cancer*. (Oxford University Press, New York).

Hanahan, D. (1997). Signaling vascular morphogenesis and maintenance. *Science*, **227**, 48–50.

Hynes, R. O. (1990). *Fibronectins*. (Springer, New York).

Ioannides, C. G., and Whiteside, T. L. (1993). T-cell recognition of human tumours – Implications for molecular immunotherapy of cancer. *Clinical Immunology and Immunopathology*, **66**, 91–106.

Jain, R. K., *et al*. (1997). Quantitative angiogenesis assays: progress and problems. *Nature Medicine*, **3**, 1203–1208.

Kleiner, D. E. and Stetler-Stevenson, W. G. (1993). Structural biochemistry and activation of matrix metallo-proteases. *Current Opinions in Cell Biology*, **5**, 891–897.

Kuznetsov, V. A., *et al*. (1994). Nonlinear dynamics of immunogenic tumours: parameter estimation and global bifurcation analysis. *Bulletin of Mathematical Biology*, **56**, 295–321.

Liotta, L. A., *et al*. (1983). Tumour invasion and the extracellular matrix. *Laboratory Investigation*, **49**, 636–649.

Little, C. D., *et al*. (eds) (1998). *Vascular Morphogenesis: In Vivo, In Vitro, In Mente*. (Birkhauser, Boston).

Loeffler, D. and Ratner, S. (1989). *In vivo* localization of lymphocytes labeled with low concentrations of HOECHST-33342. *Journal of Immunological Methods*, **119**, 95–101.

Lord, E. M. and Burkhardt, G. (1984). Assessment of *in situ* host immunity to syngeneic tumours utilizing the multicellular spheroid model. *Cellular Immunology*, **85**, 340–350.

MacSween, R. N. M. and Whaley, K. (eds) (1992). *Muir's Textbook of Pathology*, 13th edn (Edward Arnold, London).

Madri, J. A. and Pratt, B. M. (1986). Endothelial cell–matrix interactions: *in vitro* models of angiogenesis. *Journal of Histochemistry and Cytology*, **34**, 85–91.

Matrisian, L. M. (1992). The matrix-degrading metalloproteinases. *Bioessays*, **14**, 455–463.

McCarthy, J. B. and Furcht, L. T. (1984). Laminin and fibronectin promote the directed migration of B16 melanoma cells *in vitro*. *Journal of Cell Biology*, **98**, 1474–1480.

Mignatti, P. and Rifkin, D. B. (1993). Biology and biochemistry of proteinases in tumour invasion. *Physiological Reviews*, **73**, 161–195.

Muthukkaruppan, V. R., *et al*. (1982). Tumor-induced neovascularization in the mouse eye. *Journal of the National Cancer Institute*, **69**, 699–705.

Orme, M. E. and Chaplain, M. A. J. (1997). Two-dimensional models of tumour angiogenesis and anti-angiogenesis strategies. *IMA Journal of Mathematics Applied in Medicine and Biology*, **14**, 189–205.

Paweletz, N. and Knierim, M. (1989). Tumour-related angiogenesis. *Critical Reviews in Oncology and Hematology*, **9**, 197–242.

Prehn, R. T. (1994). Stimulatory effects of immune reactions upon the growths of untransplanted tumours. *Cancer Research*, **54**, 908–914.

Ratner, S. and Heppner, G.H. (1986). Mechanisms of lymphocyte traffic in neoplasia. *Anticancer Research*, **6**, 475–482.

Schor, A. M., *et al*. (1999). Experimental assays relevant for angiogenesis. In: Chaplain, M.A.J., *et al*. (eds), *On Growth and Form: Spatio-temporal Pattern Formation in Biology*. 201–224 (Wiley, Chichester).

Stetler-Stevenson, W. G., *et al*. (1993). Tumor cell interactions with the extracellular matrix during invasion and metastasis. *Annual Reviews of Cell Biology*, **9**, 541–573.

Stetler-Stevenson, W. G., *et al*. (1996). Matrix metallo-proteinases and tumour invasion: from correlation to causality to the clinic. *Cancer Biology*, **7**, 147–154.

Uhr, J. W., *et al*. (1997). Cancer dormancy: opportunities for new therapeutic approaches. *Nature Medicine*, **3**, 505–509.

Wilson, K. M. and Lord, E. M. (1987). Specific (EMT6) and non-specific (WEHI-164) cytolytic activity by host cells infiltrating tumour spheroids. *British Journal of Cancer*, **55**, 141–146.

FURTHER READING

Abbas, A. K., *et al*. (eds) (2000). *Cellular and Molecular Immunology*, 4th edn. (W.B. Saunders, Philadelphia).

Adam, J. A. and Bellomo, N. (eds) (1996). *A Survey of Models for Tumor-immune system Dynamics*. (Birkhauser, Boston).

Edelstein-Keshet, L. (1988). *Mathematical Models in Biology*. (Random House, New York).

Little, C. D., *et al*. (eds) (1998). *Vascular Morphogenesis: In Vivo, In Vitro, In Mente*. (Birkhauser, Boston).

Murray, J. D. (1989). *Mathematical Biology*. (Springer, Berlin).

Wheldon, T. E. (1988). *Mathematical Models in Cancer Research*. (Adam Hilger, Bristol).

Chapter 61

Models for Tumour Growth and Differentiation

Heimo Riedel
Wayne State University, Detroit, MI, USA

CONTENTS

- Biological Basics
- Principles of Establishment
- General Applications
- Biological and Technical Limitations
- Acknowledgements

BIOLOGICAL BASICS

From Growth Factors to Gene Expression

In prokaryotes and unicellular eukaryotes, control mechanisms have evolved to allow cell proliferation at maximum sustainable levels, often largely limited by the availability of nutrients. With the development of multicellular eukaryotes, more complex control mechanisms evolved to coordinate cell proliferation spatially and temporally, as needed for optimum function of the organism (Soto and Sonnenschein, 1993). To create functional organs, proliferation of the involved cell types had to be guided and tightly controlled by signals that induced the selective proliferation of specific cells, their alteration by differentiation or their removal by programmed cell death (apoptosis). Key coordinating signals that are exchanged between cells at a distance are transmitted via peptide growth factors (**Figure 1**). These bind to specific cell surface receptors that span the cell membrane of target cells (van der Geer *et al.*, 1994) and can involve endocrine, paracrine, juxtacrine and autocrine mechanisms. Growth factor binding frequently results in dimerization of the target receptor. This aggregated state often leads to the stimulation of an intrinsic tyrosine kinase activity of the receptor that results in autophosphorylation of the intracellular receptor domain on specific tyrosine residues. By this mechanism the growth factor signal is transmitted across the plasma membrane into the interior of the cell. Its specificity depends exclusively on the respective intracellular receptor domain and not on the growth factor (Riedel *et al.*, 1989). The phosphorylated tyrosine residues of the receptor form docking sites for a type of signalling mediator proteins that carry specific binding pockets such as SH2 (Src homology 2) domains. Assembly of these mediators on the receptor sites and their interaction with additional mediators leads to the formation of multicomponent signalling complexes. These include scaffold, anchoring and adapter proteins as well as protein kinase and phosphoprotein phosphatase enzymes. All are important members of signalling cascades that ultimately result in the activation of transcription factors in the cell nucleus by phosphorylation or dephosphorylation (Pawson and Scott, 1997). By these and other mechanisms, growth factors regulate the expression programmes of specific genes that control cell division and proliferation (Schlessinger, 2000).

Control of the Cell Cycle

The cell cycle coordinates the various processes that are involved in cell division including increase in cell mass, protein synthesis, DNA replication, separation and segregation of chromosomes and other mechanisms that lead to the formation of two daughter cells (Fantes and Brooks, 1993). The cycle consists fundamentally of a DNA synthesis phase (S) and a mitotic phase (M) which simply alternate in early embryonic cell division (**Figure 2**). If a doubling of cell mass is involved in subsequent division cycles, two gap phases (G) are introduced in the order $\rightarrow G_1 \rightarrow S \rightarrow G_2 \rightarrow M \rightarrow$ to allow the regulation of and provide the required time for the necessary synthetic cellular activities. In order to differentiate, cells typically leave the cycle in the G_1 phase and enter a G_0 or quiescent state. This state in which growth is halted can also be induced by the withdrawal of growth factors for normal and/or immortalized cells, but not for transformed cells which have lost this control mechanism. Growth factor signalling pathways can normally control progression through the cell cycle by controlling the expression or activity of the two major classes of cell cycle regulators that are deregulated in some cancers. These are represented by cyclins that vary in abundance throughout the cell cycle, and cyclin-dependent kinases which vary in activity (Hunter and Pines, 1994). An important check point or restriction

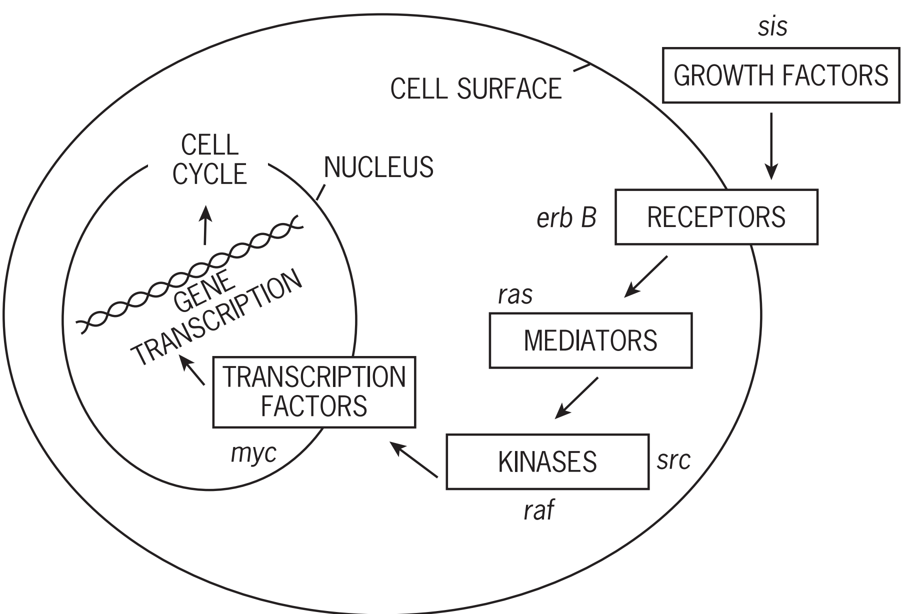


Figure 1 Major classes of signalling mediators in growth control and selected oncogene representatives. Growth factor signals are transmitted into cells via specific membrane-spanning receptors and further by various mediators and kinases to regulate transcription factors that control the expression of specific genes in the cell nucleus, and ultimately the cell cycle. For each class of mediators altered forms are expressed by oncogenes. One or two oncogene examples for each class are shown in lower-case letters (*sis*, *erbB*, *ras*, *src*, *raf*, *myc*) next to the respective mediator (shown boxed in capital letters).

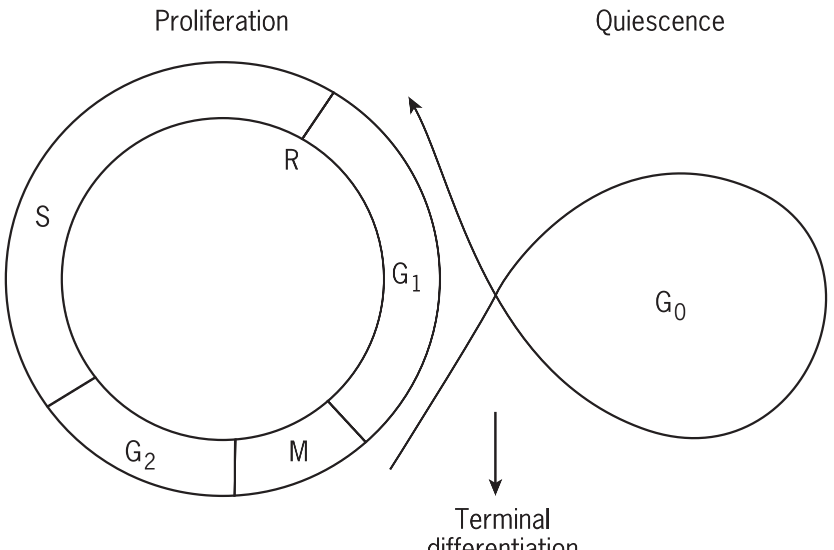


Figure 2 The proliferating and resting phases of the cell cycle. Cells are found in a gap (G) phase when they are not synthesizing DNA (S phase) or completing mitosis (M phase). Normal cells are capable of resting (quiescence) in a non-dividing state, termed G_0. One or more cycles of cell division begin once there is a need to maintain or replace tissue and stop when the required proliferation has been accomplished. Protein and RNA synthesis are active in the first gap phase (G_1). If conditions are permissive for subsequent cell division, cells pass through a restriction point (R) and quickly move into the synthetic phase (S) in which new DNA is synthesized. A second gap phase (G_2) follows when the newly duplicated chromosomes condense. In the mitotic phase (M) the chromosomes divide into two sets and the cell forms two nuclei, and divides into two daughter cells. When normal cells differentiate, typically with a gain in the properties required for organ or tissue functions, they generally lose the capacity to continue cell division. This can be irreversible in terminal differentiation.

point is found early in the G_1 phase which allows growth factor signals to control how the cycle will proceed including the transition from G_1 to S (Grana and Reddy, 1995). (See also chapter on *Overview of Oncogenesis*.)

Oncogenes and Aberrant Growth Regulation

Most human neoplasias result from abnormalities in the expression of genes that participate in the control of cell proliferation and differentiation. Since the underlying functional changes result in cancer, the altered genes are termed cancer-causing genes or oncogenes while their normal counterparts are termed proto-oncogenes. The functional alteration is typically caused either by mutational activation or by a dosage effect, which both result in an increased and/or uncontrolled function of the protein. Oncogenes have been found to encode proteins that are involved in various steps of cell growth control. Many have been originally discovered in retroviruses as nonessential genes. These have been acquired from the host cell chromosome and undergone an activating mutation during their integration into the viral genome (Fine and Haseltine, 1993). Examples of oncogenes are shown in **Figure 1** including altered forms of growth factors such as platelet-derived growth factor (v-sis), growth factor receptors such as epidermal growth factor receptor (v-erbB), nucleotide binding protein signalling mediators such as Ras (v-ras), nonreceptor membrane-bound tyrosine kinases such as Src (v-src), cytoplasmic serine/threonine kinases such as Raf (v-raf) and transcription factors such as Myc (v-myc) (Hanahan and Weinberg, 2000).

Tumour-suppressor Genes and Multi-step Tumorigenesis

Oncogenes typically act in a dominant fashion whereas a second important class of anti-oncogenes or tumour-suppressor genes generally act recessively and require the inactivation of both chromosomal alleles to result in tumorigenesis. This frequently occurs by chromosomal deletion or point mutation (see chapter on *Regulation of Cell Cycle*). The effects of tumour-suppressor genes have been discovered and best described in familial cancers. Examples of tumour-suppressor gene products include DNA repair enzymes such as BRCA1, cell adhesion molecules such as DCC (deleted in colorectal cancer), the p53 protein and the factors and major regulators of the cell cycle such as the retinoblastoma-related Rb protein. Normal cell proliferation is believed to be based on a properly regulated balance between the stimulatory function of proto-oncogenes and the constraining function of tumour-suppressor genes. Tumorigenesis is seen as a multistep process that requires a sequence of distinct alterations before it results in an invasive and metastatic cancer (Hanahan and Weinberg, 2000). A candidate sequence of six alterations in proto-oncogenes and tumour-suppressor genes is shown in **Figure 3**. The alterations involved have been particularly well described for colorectal tumour formation in the progression from normal epithelium via the stages of hyperplasia, early adenoma, late adenoma, early carcinoma and late carcinoma, to the level of metastases (Fearon and Vogelstein, 1993).

Early Tumour Growth is Difficult to Detect in Human Patients

According to the somatic mutation theory of cancer, which is favoured by most cancer researchers, arising from a single cell, a typical human tumour may have to divide about 30 times until it will become clinically detectable inside the body at a weight of about 1 g (Begg, 1993). It will take less than 10 additional divisions to reach a size that is lethal to the host before the mass approaches 1 kg (**Figure 4**). As soon as the tumour is

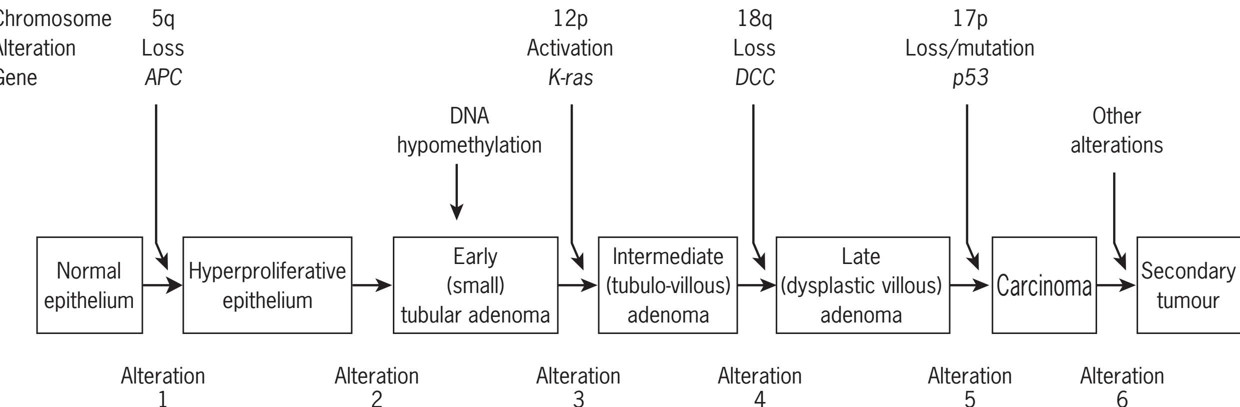


Figure 3 A model for colorectal cancer progression involving a sequence of six genetic alterations to lead from normal epithelium to a metastatic tumour. The gene involved, the type of change and its chromosomal location are indicated.

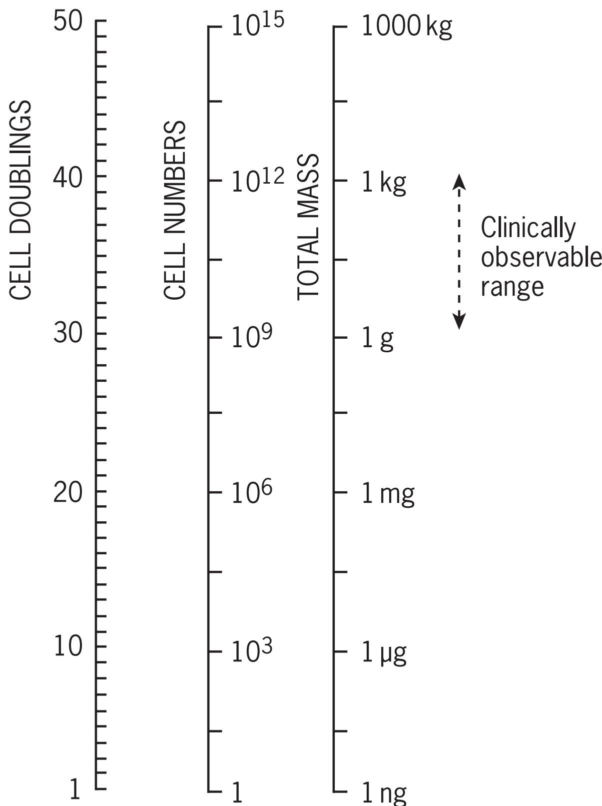


Figure 4 Correlation between the mass of a tumour, its approximate cell numbers and the numbers of required cell doublings (assuming an origin from a single cell and 10^9 cells per gram). Only the final quarter of the growth phase is clinically observable.

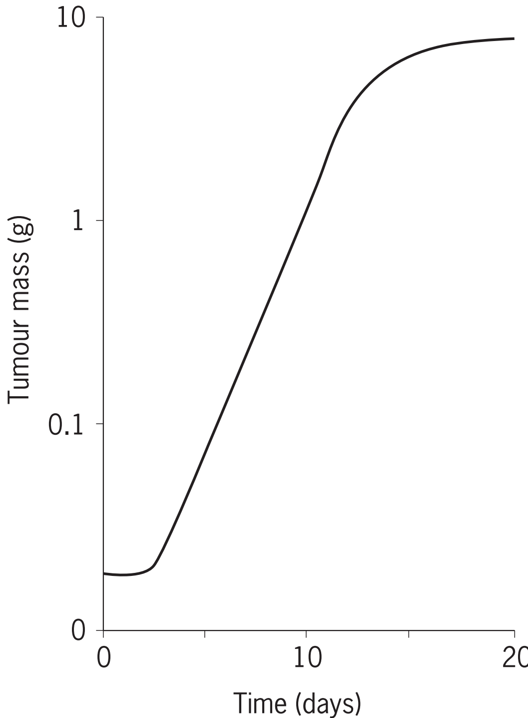


Figure 5 The phases of experimental tumour growth. Following an initial lag phase, early growth proceeds exponentially for several days before it slows and reaches a plateau. This is mainly caused by increased cell loss (such as through necrosis) that essentially balances the continuing proliferation to a static level of viable cells.

detected, therapy has to be initiated immediately. Normally tumour growth in humans can only be studied during less than the last quarter of its growth history and in the presence of inhibitory therapeutic measures. During the final growth stages of large tumours, cell loss rates by necrosis such as due to insufficient blood supply may begin to approach proliferation rates which are not representative of the earlier undetected growth phase (Alison and Sarraf, 1997a). Consequently, much of our limited understanding of tumour growth has been derived from animal models (**Figure 5**) or *in vitro* cell culture models, of which the latter will be the focus of this chapter.

Table 1 Comparison between growth fraction and cell doubling time: a correlation is shown for a number of experimental and human tumours

Tumour	Growth fraction (%)	Doubling time (days)
Experimental tumours		
L1210 (mouse)	86	0.5
B 16 (mouse)	55	1.9
LL (mouse)	38	2.9
DMBA (rat)	10	7.4
Human tumours		
Embryonal carcinoma	90	27
Lymphoma (high grade)	90	29
Squamous cell carcinoma	25	58
Adenocarcinoma	6	83

Parameters of Tumour Growth

In terms of population kinetics, the growth of any tissue including tumours is governed by (1) the rate of individual cell division, (2) the growth fraction of the cell population and (3) the rate of cell loss from the growing population such as through differentiation or cell death (**Table 1**). Healthy organs will be closely controlled with regard to their cell proliferation rate, fraction of proliferating cells and cell loss. In contrast, the cells in a tumour will not necessarily

Table 2 Comparison of growth parameters between human neoplasms and normal tissues: a correlation is shown between the labelling index and the estimated cell doubling time for a number of tissue sources

Cell type	Labelling index (%)	Estimated cell doubling time (days)
Normal bone marrow myeloblasts	32–75	0.7–1.1
Acute myeloid leukaemia	8–25	0.5–8.0
Normal B-cell lymphocytes	0–1	14–21+
High-grade lymphoma	19–29	2–3
Normal intestinal crypts	12–18	1–2
Colon adenocarcinoma	3–35	1.6–5.0
Normal epithelium/pharynx	2–3	—
Squamous cell carcinoma of the nasopharynx	5–16	2–4
Normal epithelium/bronchus	—	9–10
Epidermoid carcinoma of the lung	5–8	8–10
Normal epithelium/cervix	4–8	—
Squamous cell carcinoma of the cervix	13–40	—
Ovarian carcinoma	3–20	5–6
Benign mole of skin	0.3	—
Malignant melanoma of skin	12.8	—

proliferate faster; however, the controls will be greatly relaxed, resulting in continued unrestricted growth. Inflammatory lesions can proliferate up to 20 times faster than tumours but this rapidly ceases owing to intact control mechanisms when healing is complete (**Table 2**). Tumours can often be thought of as wounds that do not heal.

It is generally assumed that tumour cells multiply exponentially during the early phases of tumour growth (Alison, 1995). As the mass increases, the growth rate declines, which results in a sigmoid exponential (Gompertz) growth curve as shown for an experimental animal tumour in **Figure 5**. One important limiting factor is the difficulty of obtaining sufficient blood supply. This frequently results in necrosis in the centre of larger tumours whereas the cell doubling time typically changes little. Under adverse conditions, tumour cells may leave the growth fraction and enter the G_0 or a prolonged G_1 phase of the cell cycle until nutritional conditions become favourable or cell proliferation is stimulated by growth factors.

Measuring Tumour Cell Proliferation

The rate of tumour cell division may not tell us much about the actual rate of increase of tumour size. However, the proliferative index may be a prognostic indicator since high levels of cell proliferation may help identify patients who may benefit particularly from an aggressive antitumour therapy. Cell population kinetics are directed at the state and the rate parameters of proliferating tumour cells. This involves the determination of population size or cell numbers, the rates by which cells transit (fluxes) between different phases of the cell cycle (**Figure 6**) and the time cells spend in a specific phase of the cell cycle (**Figure 2**).

Radioactive Markers for DNA Synthesis

Many simple methods for measuring cell proliferation define only state parameters since they provide information about the proliferative state of cells but not about the rate at which they occur. The mitotic count is often defined as the number of mitoses per 10 high-power microscope fields (HPF) but takes no account of cell size. The typically preferred mitotic index (MI) is the percentage of the cell population that is found in the mitotic phase of the cell cycle. It depends on the time that cells spend in mitosis and also on the fraction of cells in that phase, two parameters that cannot be discriminated. Radioactively labelled [^{3}H]thymidine is specifically incorporated into DNA as opposed to RNA and has been widely used since its first description in 1957. It is typically measured as a pulse labelling index (TLI%), which represents the percentage of cells in the S phase of the cell cycle. It is always larger and presumably more reliable than the MI. The major drawbacks of this method are the cost of the reagent, the required photographic materials, darkroom and technical expertise and the time required for the autoradiographic detection of several weeks. In addition, in human patients the considerable radioactive burden must be added through the ^{3}H radioactive half-life of more than 10 years. This can be reduced by the use of [^{11}C]thymidine with a half-life of only hours to study uptake by positron emission tomography (PET). Bromodeoxyuridine represents a major alternative that is similarly specifically incorporated into DNA during DNA synthesis. It can be detected by specific monoclonal antibodies even in alcohol-based and formalin-fixed, paraffin wax-embedded tissue sections based on standard indirect immunolabelling techniques within days. It is the method of choice for identifying the S phase fraction where the substrate can be administered

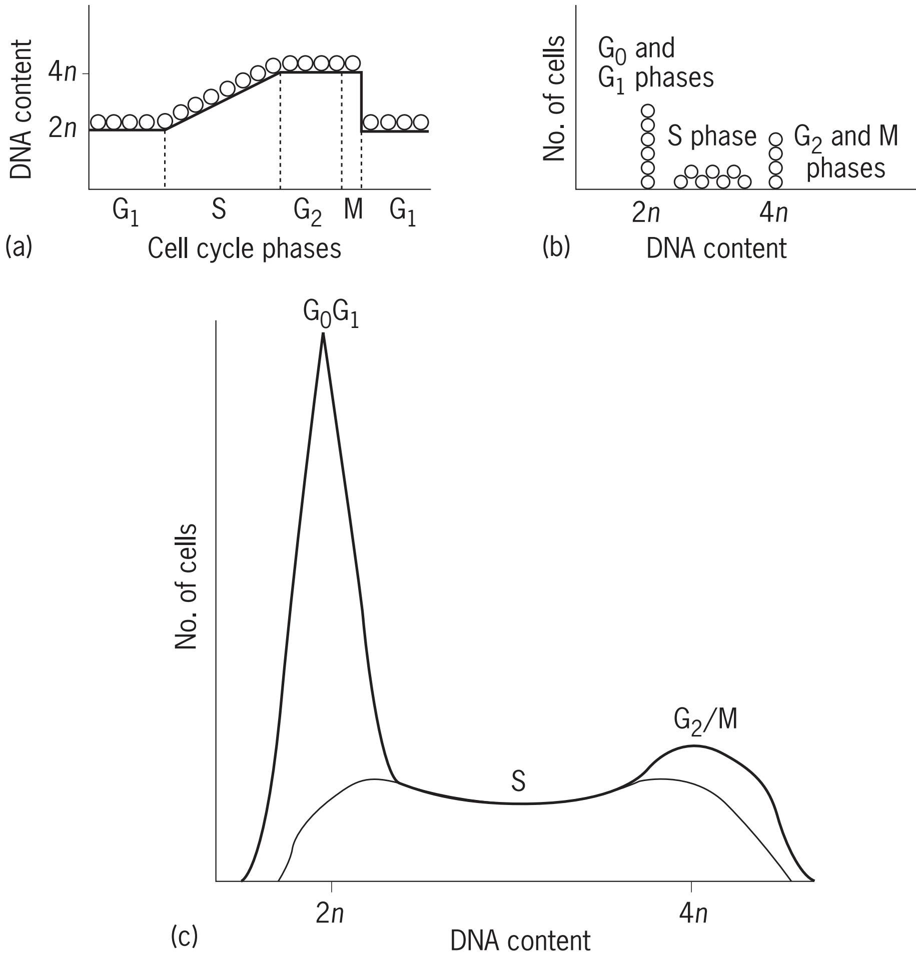


Figure 6 The principle of a DNA histogram as analysed by flow cytometry. (a) Cells double their DNA content as they pass through S phase, so cells in G_2 have twice the DNA content of G_1 cells. (b) Representation of cell segregation by flow cytometry when sorted based on DNA content. (c) DNA histogram obtained by flow cytometry; the exact number of cells in each category is derived by computer analysis. [From Alison and Sarraf, 1997a, in: *Understanding Cancer: from Basic Science to Clinical Practice*, 133–193 (Cambridge University Press, Cambridge).]

to live experimental animals and certain cancer patient groups (Alison and Sarraf, 1997a).

Antibodies for Cell Cycle Markers

Both DNA incorporation methods share as a common disadvantage the requirement for vital tissues and cannot be used on archival material in retrospective experiments (Hall *et al*., 1992). This limitation is overcome by monoclonal antibodies such as MIB1 (Molecular Immunology Borstel). MIB1 is directed against the Ki-67 protein that is expressed in all phases of the cell cycle and represents a preferred diagnostic tool with few disadvantages. An auxiliary factor to DNA polymerase delta provides an alternative: proliferating cell nuclear antigen (PCNA) which can be detected immunohistochemically in routinely processed tissues during and even shortly after the cell cycle owing to the long half-life of the antigen. As a disadvantage it is easily destroyed by prolonged fixation (Alison, 1995).

Metaphase Arrest

All proliferative indices described above reflect changes in the proliferative rate and in the duration of the cell cycle phases that cannot be discriminated. As a remedy, the metaphase arrest technique provides a firm kinetic parameter, the rate of entry into mitosis, over an experimental period as short as 2.5 h. The *Vinca* alkaloids vincristine and vinblastine have found widespread use as arresting agents that disrupt the assembly of the mitotic spindle and result in characteristic, abnormal and easily recognizable mitotic figures. However, for accuracy, this approach requires multiple samples that limit its application to experimental models (Alison, 1995).

Flow Cytometry

A considerable disadvantage of light microscopic evaluation results from having to score several thousand cells, a time-consuming procedure. This is addressed by flow

cytometry that automatically measures the DNA content of large cell numbers in suspension. For this purpose, DNA is stained by a fluorescent DNA-binding dye and individual cells are directed through an exciting laser beam. The resulting fluorescence signal is proportional to the amount of nuclear DNA that reflects the stage of the cell cycle. Data are directly available for computer and statistical analysis of cell populations (**Figure 6**). The labelling index can be estimated by measuring the BrdU–DNA profile in parallel. Patients can be injected with BrdU and its incorporation into cellular DNA can be measured by indirect immunofluorescence with a different fluorescent dye in parallel to the DNA content. Data can be displayed as a two-dimensional histogram of cell populations in different stages of the cell cycle (Alison and Sarraf, 1997a).

Control of Differentiation

In addition to abnormalities in the regulation of cell proliferation most, if not all, tumours show abnormalities in differentiation such as anaplasia. This can help elucidate their aetiology, degree of malignancy, prognosis and sensitivity to therapeutic intervention by differentiation- or maturation-inducing agents. The state of differentiation of a cell refers to its specialization with regard to expression of a specific subset such as 10 000 of its total set of up to 40 000 genes. Differentiation begins shortly after fertilization of the embryo and permits its development into specialized organs and tissues whereas proliferation is the primary mechanism by which cell mass increases early. In a particular cell the subset of expressed genes is controlled by a set of specific transcription factors. They depend upon its embryonic lineage, developmental stage, tissue and cellular environment and the specific physiological role of the cell. Embryonic cells typically have the capability to proliferate vigorously, migrate extensively, produce enzymes that degrade basement membranes and secrete factors that increase the local blood supply. Most of these characteristics are shared with tumour cells and when combined contribute to increased malignancy, but have typically been lost in differentiated adult cells. Consequently, in the adult organism loss of a highly differentiated state or activation of embryonic gene expression programmes has been found to correlate with malignant properties of tumour cells (Mintz and Fleischman, 1981).

Differentiation Versus Proliferation

Stem cells can be multi- or toti-potent and combine the capacity either to differentiate into a variety of cell types or to multiply without change in their gene expression programme by simple proliferation. Differentiation results in the expression of a specific set of genes that is paralleled by a restriction of the cell for further proliferation. Some cells reach the stage of terminal differentiation. This may still allow cells to carry out specialized functions over much of the organism's life span such as in the case of mature muscle or nerve cells. Alternatively, it may result in subsequent cell death such as in the case of keratinocytes in the outer epidermal layer. The capacity for proliferation including malignant proliferation appears to be inversely correlated with the state of differentiation. Accordingly, terminally differentiated cells have never been found to form tumours, in contrast to less differentiated myoblastic or neuronal stem cells. Conversely, when cells are induced to proliferate with a potent mitogen or by transfection of an oncogene, they can no longer undergo terminal differentiation. Tumour-cells are generally characterized by their inability to differentiate terminally under appropriate conditions either *in vivo* or in culture (Rheinwald and Beckett, 1980; Wille *et al*., 1982).

Senescence Versus Immortalization

The finite life span of normal, differentiated cells represents an obstacle for malignant transformation and senescence resembles terminal differentiation in some ways. Escape from senescence by immortalization is therefore seen as a prerequisite step for tumorigenesis. Once human cells senesce it is virtually impossible to induce proliferation by normal stimuli (McCormick and Campisi, 1991). This phenotype is genetically dominant and induces senescence in hybrids between normal and immortal cells or results in a normal phenotype in hybrids between normal and tumour cells. These findings have led to the discovery of tumour-suppressor genes such as those encoding p53 or the retinoblastoma protein Rb. The latter is required for terminal differentiation of myoblasts and for maintenance of the senescent state (Gu *et al*., 1993).

Control of Differentiation at the Cell Surface

Differentiation typically depends on external factors including the extracellular matrix, proximity and type of neighbouring cells and soluble factors. Members of each category bind to specific cell-surface receptors that initiate signal transduction pathways that control gene expression. Tumour cells often lose their ability to sense the extracellular matrix or neighbouring cells. Malignant human breast epithelial cells can often be identified by their ability to grow on extracellular matrix in a gel that inhibits the growth of normal cells *in vitro* (Lin and Bissel, 1993). An important first class of soluble factors includes fibroblast growth factors, transforming growth factors and haematopoietic growth factors such as interleukins (**Table 3**). FGF induces mesoderm differentiation in early embryos, inhibits differentiation of other cells and plays a role in the survival of endothelial cells and in angiogenesis. Like FGF, TGFβ stimulates the differentiation of some cells such as keratinocytes or intestinal epithelial cells but inhibits the differentiation of others such as myoblasts and preadipocytes.

Table 3 Cell culture models for *in vitro* differentiation: for various stem cell types the inducers of differentiation and the resulting marker phenotypes are shown[a]

Stem cell	Differentiation markers[b]	Inducers[b]
Preadipocyte	Adipocyte	Insulin, cort, cell density
Basal keratinocyte	Cornified envelope	RA deficiency, cell density
Myoblast	Myotube	GF deficiency, cell density
Squamous cell carcinoma	Cornified envelope	GF deficiency, cort
Embryonal carcinoma	Endoderm, mesoderm, ectoderm	RA, ara-C, mito, HMBA, co-culture with blastocysts
Neuroblastoma	Neuron, neurotransmitter, action potential	PI, 6TG, ara-C, MTX, dox, bleo, RA, GF deficiency
Melanoma	Dendrite, melanin, tyrosinase	PI, dox, DMSO, TPA, RA, MSH
Colon adenocarcinoma	Mucus, dome formation, CEA, columnar cell	NMF, DMSO, butyrate, low glucose, IFN, HMBA, cell density
Breast adenocarcinoma	Casein, dome formation	RA, PGE, DMSO
Bladder transitional cell carcinoma	Keratin filament, loss of surface antigen	HMBA
Erythroleukaemia	Mature erythroid cell, haemoglobin	Dox, ara-C, 6TG, mito, dact, aza, haemin, DMSO, HMBA, CSF, RA, IFN
Promyelocytic	Granulocyte, macrophage	IFN, CSF, vitD, TPA, DMSO, NMF, dact, HMBA, aza, ara-C, RA
Myelocytic leukaemia	Granulocyte, macrophage	CSF, RA, vitD, ara-C, dact, DMSO, TPA, cort, dox

[a] Data from Cheson *et al.*, 1986, Reiss *et al.*, 1986 and Waxman *et al.*, 1988.
[b] ara-C, cytarabine; aza, 5-azacytidine; bleo, bleomycin; CEA, carcinoembryonic antigen; cort, glucocorticoids; CSF, colony-stimulating factor; dact, dactinomycin; DMSO, dimethyl sulfoxide; dox, doxorubicin; GF, growth factor; HMBA, hexamethylbisacetamide; IFN, alpha- or gamma-interferon; mito, mitomycin C; MSH, melanocyte-stimulating hormone; MTX, methotrexate; NMF, *N,N*-dimethylformamide; PGE, prostaglandin E; PI, phosphodiesterase inhibitor; RA, retinoic acid; TPA, 12-*O*-tetradecanoylphorbol-13-acetate; vitD, 1,25-dihydroxyvitamin D; 6TG, 6-thioguanine.

It promotes a more differentiated phenotype of some human tumour cells and inhibits their proliferation in part by inducing expression of cyclin-dependent kinase inhibitors that result in loss of cyclin-dependent kinase activity which prevents the phosphorylation and thus the inactivation of the Rb protein (Sherr, 1994). The second class of soluble differentiation factors include membrane-permeable regulators such as retinoic acid and its derivatives that regulate morphogenesis in the early embryo via concentration gradients and stimulate or inhibit growth or differentiation dependent on the specific cell type. The sex steroids oestrogen and testosterone are believed to regulate differentiation via similar mechanisms through specific nuclear receptors that control transcription (Dolle *et al*., 1989).

Gene Expression in Differentiation Control

External differentiation factors cooperate with a small number of master regulatory genes including homeotic genes to control the expression of differentiation-specific genes. Homeotic genes carry a conserved sequence motif termed homeobox and play a fundamental role in the definition of major body structures. Mutations in homeotic genes can result in the exchange of body parts such as legs with antennae in insects; in addition, mutations may reactivate parts of an embryonic gene expression programme or suppress aspects of an adult programme. This may result in higher proliferative, invasive or angiogenic potential, features shared by embryonic and tumour cells (Deschamps and Meijlink, 1992). Well-studied master regulatory genes in a restricted lineage are the *myogenin*, *myoD* and *myd* genes which all encode transcription factors that regulate muscle differentiation. Inactivating mutations select mesenchymal stem cells at various stages of commitment to muscle differentiation but cannot terminally differentiate and possess the potential to form tumours. Typically, cells will be controlled by alternative gene expression programmes that regulate either differentiation or the cell cycle, such that they are largely mutually exclusive.

PRINCIPLES OF ESTABLISHMENT

In Vitro Cell Culture Models

Cancer in humans is seen as a slow and complex, multistep process with each step occurring at low frequency at the cellular level (**Figure 3**). The cellular origin of the tumour is often obscured by the heterogeneity and complexity of the tumour tissue and by the loss of differentiation markers (**Table 3**). This makes research on human tumours very difficult, which is further complicated by the inaccessibility of human tumours and the need to interfere therapeutically

as soon as they are detected. The elucidation of the detailed mechanism of human tumour formation and growth requires the tools of molecular and cellular biology and experimental models in which the process of malignant conversion can be predictably induced and reproducibly studied. Cell cultures have become an essential model in experimental oncology. This is based on the principle that normal, including human, cells growing in culture can be altered experimentally into tumorigenic cells that undergo morphological transformation (Hahn *et al*., 1999). These can be injected into animal hosts where they will induce tumour growth in a process termed malignant transformation.

Immortalized Cell Lines

Primary cell cultures are initially established by dissociation of a tissue into single cells such as by mechanical disruption and proteolytic enzyme digestion. In the case of mouse embryo tissue many cells will initially grow but the culture will enter crisis within days. Most cells will die as a result of their limited life span, which is characteristic for all somatic cells. From rodent tissues, but typically not from human tissues, a small number of variants are expected to survive, overgrow the culture and result in immortalized cell lines (**Figure 7**). These are normally fibroblasts that will form a monolayer on the culture plate. Cells can be dissociated by proteolytic treatment and can be subcultured (or passaged) into fresh cultures. In contrast to most primary cultures, immortalized cell lines are able to grow at low cell density and allow the cloning of single cells into homogeneous populations. However, they invariably display abnormal karyotypes, are in many respects already preneoplastic, and may undergo spontaneous transformation since they have already acquired some of the alterations required for the transition to a fully malignant phenotype (Martin, 1996). (See also chapter on *Basic Tissue Culture in Cancer Research*.)

Cell Transformation

Proliferation of cultured cells depends on many nutrients that are routinely supplied by a synthetic medium and on many other components including growth factors that are typically provided by supplementing the medium with serum. In particular, rodent fibroblast cultures have become popular as *in vitro* cancer models since immortalized variants are easily established and can readily undergo neoplastic or malignant transformation by oncogenes such as from tumour viruses or by chemical carcinogens. This is scored most stringently by the cells' ability to form tumours upon injection into experimental animals such as mice. In contrast to normal cells, transformed fibroblasts usually acquire autonomous or unregulated growth properties and lose contact inhibition. They continue to proliferate at high densities and form cell foci, and can proliferate when deprived of growth factors (**Figure 8**). They typically display anchorage independent cell-growth (**Figure 9**) in the absence of adhesion to an extracellular matrix and display a more rounded or spindle-like morphology. These changes result from a disruption of the actin cytoskeleton and its

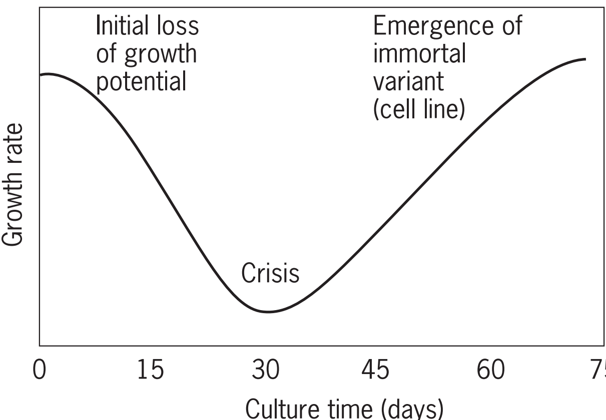


Figure 7 Principle of rodent fibroblast immortalization. In a primary culture of mouse embryo cells initial cell death is paralleled by the emergence of healthy and proliferating cells. As these cells are diluted and continue to divide the culture goes into crisis, and most cells die. Only a small number of cells survive and proliferate until their progeny finally overgrow the culture. These cells constitute an immortalized cell line, which will proliferate as long as it is appropriately diluted and supplied with nutrients.

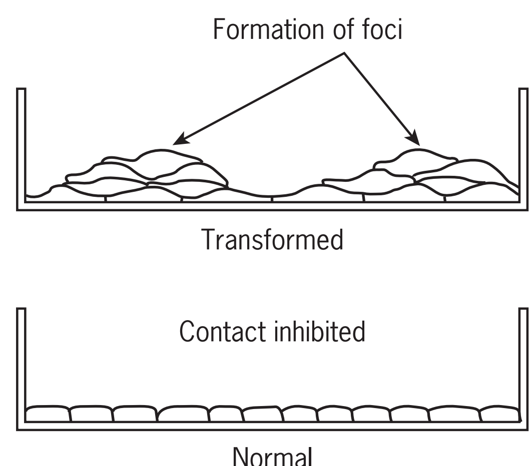


Figure 8 Morphological cell transformation. Normal, immortalized fibroblasts typically grow on a culture dish as a monolayer that is maintained by their response to contact inhibition. Transformed cells have typically lost contact inhibition and grow on top of each other in a less organized pattern. This results in domes of cells termed foci that are frequently visible with the unaided eye after many cell divisions.

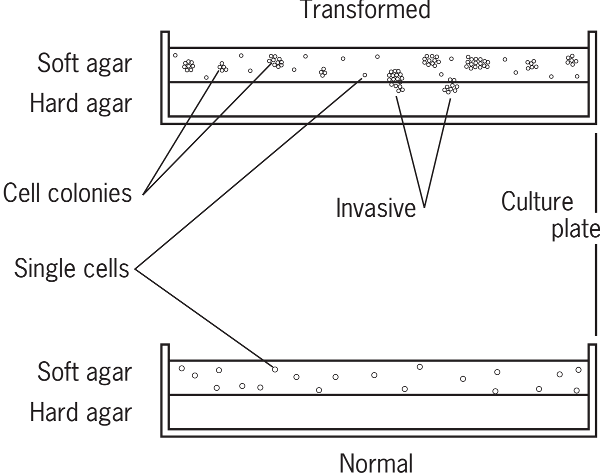


Figure 9 Anchorage-independent cell growth. Normal immortalized fibroblasts do not proliferate in a semisolid medium, such as in soft agar. The loss of anchorage dependence is a key criterion for the transformed phenotype that allows transformed cells to proliferate and form colonies in a semisolid medium. If colony formation proceeds even at increased agar concentration this frequently reflects an invasive phenotype and may correlate with metastatic potential.

connections to and changes in the extracellular matrix. They will often, but not always, correlate with the potential to form tumours *in vivo*. Since transformed cells may be rejected by the immune system, immunologically compromised nude mice are most commonly used for tests of tumorigenicity (Cifone and Fidler, 1980; Li *et al*., 1989).

Malignant Transformation by Tumour Viruses

Tumour viruses are found in many mammalian and avian species and some, in particular papilloma and certain hepatitis viruses are a significant cause of human cancer. Cultured cells can be rapidly and reproducibly transformed by many tumour viruses and the involved viral genomes are relatively simple. This has greatly facilitated the molecular dissection of the resulting cellular changes and of the underlying transforming genes. Two groups of tumour viruses have been particularly important in the development of molecular oncology. Acutely-transforming retroviruses carry transforming genes or oncogenes that have been acquired from normal cellular progenitors, termed proto-oncogenes. The other group is represented by papova viruses including monkey simian virus 40 (SV40) and mouse polyoma virus. Both viruses encode proteins that are involved in DNA replication and malignant transformation. Both virus classes result in the integration of viral DNA into the host chromosome, a critical step for the replication of viral DNA. This permits the consistent expression of viral gene products, in particular those of the viral oncogenes that are responsible for the transformed state of the cell (Topp *et al*., 1980; Klein, 1993).

Some retroviruses, including acute leukaemia viruses, which induce rapid leukaemia *in vivo*, can transform other cell types, in particular haematopoietic cells. When suspended in a gelled medium, normal target cells proliferate to a limited extent before they undergo terminal differentiation, whereas transformed cells grow into larger, less differentiated colonies. Abelson murine leukaemia virus has been shown to transform lymphoid cells, avain erythroblastosis virus to transform erythroid cells and avian myelocytomatosis or myeloblastosis viruses to transform myeloid cells (Beug *et al*., 1994; Rosenberg, 1994). With these approaches the largely opposing effects of cell differentiation and malignant transformation have been characterized. (See also chapter on *Haematological Malignancies in Cancer Research*.)

Many retroviruses transform most of the fibroblast population in a culture at high multiplicities of infection. Transformed cells become anchorage independent and will continue to grow into colonies when suspended in a semisolid medium such as soft agar or methylcellulose, in contrast to normal cells. The reproducible experimental induction of this phenotype by transforming tumour viruses and its correlation with the viral induction of sarcomas in experimental animals represents one of the foundations of experimental tumour virology (Klein, 1993).

DNA Transfection of Cultured Cell Lines

Since oncogenes had been identified as the tumour-causing principle of tumour viruses, experimental approaches were designed to test for the presence of oncogenes in human tumours based on the technique of DNA transfection (Kern, 1996). For this purpose, genomic DNA was isolated from tumour cells and treated with calcium phosphate solution to form a DNA precipitate. Such a precipitate is effectively taken up by cultured mammalian cells into the cytoplasm. Immortalized, normal NIH 3T3 mouse fibroblasts, a model system to study malignant transformation, were transfected and cultured for 2 weeks and their morphology was analysed for the characteristics of malignant transformation (**Figure 10**). Foci of transformed cells were observed, based on the presence of the human EJ bladder carcinoma oncogene which was established as a homologue of the (Harvey) *v-ras* oncogene in human tumour cells by this procedure (Parada *et al*., 1982). In the case of human primary cells, *in vitro* malignant transformation has only recently been accomplished since it requires the deregulation of at least four signalling pathways in some specific nonrandom order. The four steps involve immortalizing cells by overexpression of the enzyme telomerase, activating mitogenic signalling by overexpression of a *v-ras* oncogene, and by interfering with two tumour suppressor genes, cellular *p53* and *Rb*, which is accomplished by

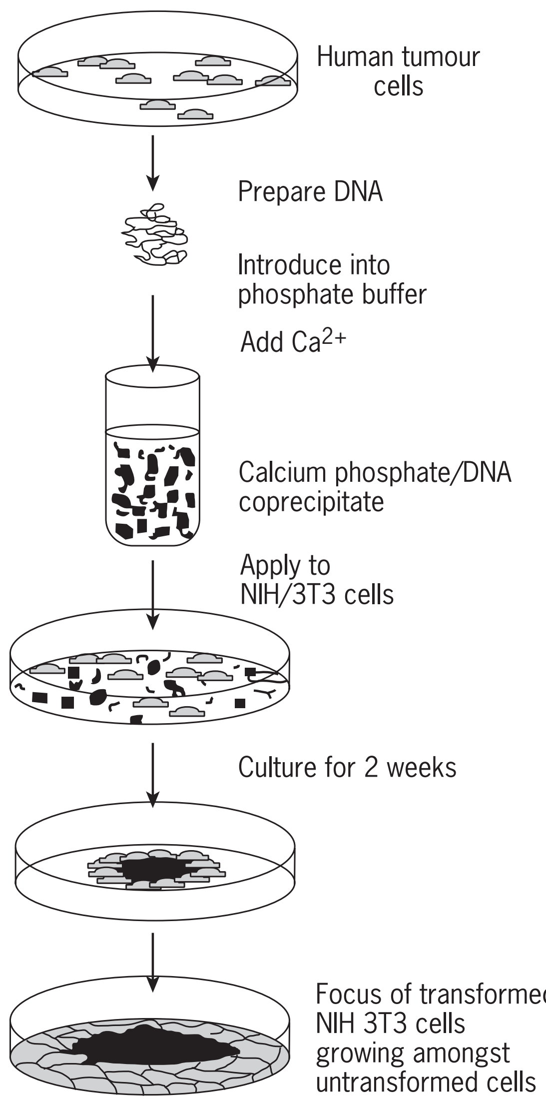


Figure 10 The prototypical transfection experiment to demonstrate a cellular oncogene. [From Alison and Sarraf, 1997b, in: *Understanding Cancer: from Basic Science to Clinical Practice*, 133–193 (Cambridge University Press, Cambridge).]

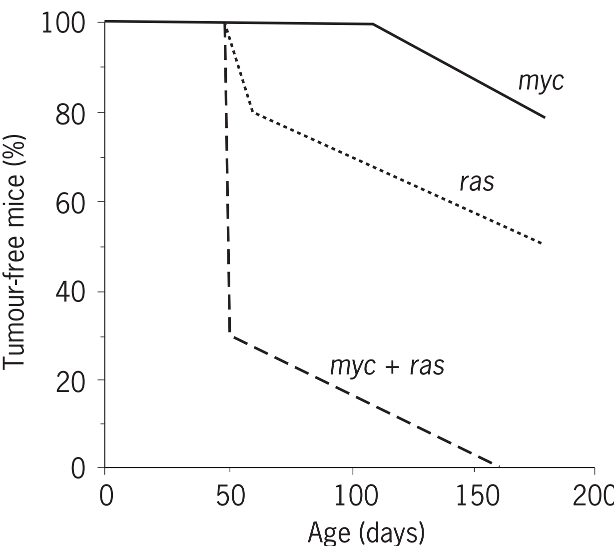


Figure 11 Oncogene cooperation in tumorigenesis. The incidence of tumorigenesis in three types of transgenic mice is shown, one carrying a *myc* oncogene, one carrying a *ras* oncogene and one carrying both oncogenes. Co-expression of *myc* and *ras* results in a dramatic and synergistic acceleration of tumour formation.

overexpression of SV40 T antigen (Hahn *et al*., 1999). For normal mouse fibroblasts, expression of at least two different oncogenes was found to be minimally required to result in cell transformation. Cooperation of different oncogenes has been shown to accelerate and enhance dramatically tumorigenesis in transgenic mice (**Figure 11**).

Alterations in the DNA Sequence of Oncogenes

Molecular cloning and sequence analysis of tumour virus oncogenes and the identification of their cellular proto-oncogene counterparts by nucleic acid hybridization analysis were instrumental techniques in the definition of the specific alterations that convert normal cellular genes into transforming oncogenes (Kern, 1996). Frequently a gene dosage effect was implicated which led to an increased gene expression level in the tumour virus, or to functional gene activation by fusion of the proto-oncogene with a viral gene that resulted in an activated and/or deregulated fusion protein. Similar gain-of-function results were also observed by mutations within the amino acid coding region of the oncogene. These were minimally based on a single nucleotide exchange that resulted in the substitution of one critical amino acid. With this knowledge, it was possible to recreate oncogenes by *in vitro* site-directed mutagenesis of proto-oncogenes that had been isolated from normal cells. These oncogenes were integrated into cDNA expression plasmids that were introduced into normal immortalized rodent fibroblasts to confirm the resulting transformed phenotype as shown in **Figures 8** and **9** (Rodrigues and Park, 1994).

Oncogene-induced Alterations in Cell Morphology

In order to study the molecular details that result in the transforming potential of oncogenes, *in vitro* assays were developed to score oncogene function (Cifone and Fidler, 1980). These assays were based on the observed deregulation of the cell cycle and on alterations in cytoskeletal and adhesion proteins. Typically, oncogene expression results in morphological alterations in rodent fibroblasts that are detectable in the phase contrast microscope, disturb organized growth and often result in a crisscrossed and chaotic growth

pattern rather than in an organized monolayer. Contact inhibition is normally lost and leads to piles of cells or foci that are detectable with the unaided eye. Foci are routinely scored by staining cells with a dye for increased contrast such crystal violet (**Figure 8**). Changes in cytoskeletal organization can be evaluated by indirect immunofluorescence with antibodies or specific dyes directed against major cytoskeletal markers such as actin which frequently display a disorganized pattern in transformed cells. Cell morphology is frequently more rounded and includes new surface projections. Some oncogenes such as v-*abl* of Abelson murine leukaemia virus result in a distinct phenotype of rounded cells that have lost their adherence to the culture plate.

In Vitro Assays for Cell Transformation

Anchorage-independent growth, a classical marker of fibroblast transformation, can be scored by colony formation in semisolid medium such as soft agar, conditions under which normal fibroblasts cannot proliferate (**Figure 9**). Combined with focus formation (**Figure 8**), the observation of anchorage-independent proliferation frequently correlates with the potential of transformed cells to form tumours in nude mice, the ultimate marker for experimental malignant cell transformation. These *in vitro* assays consequently permit the prescreening of numerous samples for their transforming potential and can limit the use of experimental animals to confirm the activity of the samples with the highest *in vitro* scores. The capacity for anchorage-independent growth can be co-evaluated with the invasive potential of transformed cells which is reflected in their ability to grow in agar of increased concentration (**Figure 9**). Cells which proliferate into colonies under these conditions may show invasive and metastatic potential in human and animal hosts (Li *et al*., 1989).

In addition, a minimal dependence of the cell cycle on growth factors is often the result of oncogene action (Aaronson, 1991). Normal fibroblasts typically require a serum supplement to their synthetic growth medium that can be partly replaced by a mixture of several factors including transferrin, insulin and epidermal growth factor. Oncogene expression frequently renders fibroblasts independent of one or more of these factors to undergo several cell divisions whereas normal fibroblasts will not repeatedly divide. Alternatively, oncogene expression may support cell proliferation at reduced serum concentrations where normal fibroblasts will not divide. Cell proliferation is routinely assayed either by cell counts in the phase contrast microscope, in a flow cytometer, or indirectly (MTT assay) by measuring mitochondrial succinate dehydrogenase activity (Denizot and Lang, 1986).

In Vitro Models for Differentiation

The elucidation of the fundametal molecular mechanisms that control cell proliferation, including growth factors and their signalling pathways, the cell cycle and the involved gene expression programmes, has to a large extent been a byproduct of the investigation of oncogene function. Many growth factors are also differentiation factors at specific stages in development and regulate similar signalling cascades in either function which activate distinct gene expression programmes dependent on the specific cell, tissue or developmental context (Cross and Dexter, 1991). With this knowledge, rapid progress was made initially in simple experimental model systems in the unravelling of the molecular mechanisms in differentiation and development. Since the molecular mechanisms and components are frequently shared between proliferation and differentiation programmes, many of experimental approaches described above can be applied interchangeably. In particular, a number of cell culture models were established based on specific stem cells such as preadipocytes or myoblasts which were known to differentiate into adipocytes and myotubes, respectively (**Table 3**). When specific mechanisms of induction were defined, including insulin and cell density or growth factor deficiency and cell density, respectively, experimental models were established to study the control of cell differentiation under defined *in vitro* conditions. These include a significant number of tumour models that permit the molecular dissection of the opposing effects of transformation and differentiation. Future antitumour differentiation therapies will benefit from these experimental strategies.

Simple Genetic Models in Development and Differentiation

Even more importantly, simple experimental models with established genetics were instrumental in the observed rapid progress in this field. They provided a unique advantage by allowing the application of genetic and biochemical experimental strategies in parallel. Even the simplest multicellular organisms undergo differentiation and development but typically malignant transformation is not easily observed. Simple eukaryotic models have not been applicable to study tumorigenesis directly. The most important genetic model systems are the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans*, the slime mould *Dictyostelium discoideum* and more recently the mouse *Mus musculus*. In these models, master genes such homeogenes have been identified which control central developmental programmes in most organisms. With the progress of genome mapping and sequencing in mammals including mice and humans, the knowledge gained with the simple genetic models is now rapidly being applied to the understanding of differentiation and development in mammals including humans.

GENERAL APPLICATIONS

Retrovirus Models

Retrovirus models have not only been been instrumental in the dissection of oncogene function but have similarly helped elucidate the pathogenic mechanism involved in other human diseases that they cause. One important example is the human immunodeficiency virus (HIV) and the resulting acquired immunodeficiency syndrome (AIDS). The investigation of the underlying molecular mechanisms and the search for effective therapies has greatly benefitted from the knowledge gained earlier with transforming retroviruses. Since retroviral RNA is integrated into the genome of the host cell as part of the viral life cycle, this mechanism has been exploited to design expression vectors for experimental models as well as for human gene therapy. Such vectors result in the stable and predictable integration of defined protein-coding sequences into the genome. This allows their stable propagation into daughter cells in parallel with genomic DNA replication and cell division. In addition, the expression level of the inserted sequences can be adjusted or externally regulated by the choice of the specific transcriptional promoter. Retrovirus models have greatly contributed to the discovery of fundamental mechanisms in gene expression and transcription. A previously unknown enzymatic function, termed reverse transcriptase, was shown to catalyse the transcription of the retroviral genomic RNA template into a DNA copy to allow integration of the sequence into the DNA genome. This enzyme has been instrumental in the cloning and sequencing of thousands of protein-coding sequences starting with their specific mRNA (Gallo, 1995).

DNA Tumour Virus Models

Other tumour virus models such as polyoma virus and, in particular, simian virus 40 (SV40) have greatly helped elucidate gene structure and function based on their small DNA genome size. They represent one of the smallest self-replicating eukaryotic systems and have resulted in the first completely sequenced viral genome (SV40). In both viruses many regulatory functions are controlled by a transforming early viral protein family termed T or tumour antigen that regulates viral transcription. In contrast, the oncogenes of retroviruses encode cellular growth-regulating proteins that have become activated or deregulated by integration into the retroviral genome. The investigation of retroviruses has helped identify many key mediators of cell proliferation and differentiation, typically one per virus. In contrast, the transforming proteins of DNA tumour viruses had evolved to interact with many key mediators of mitogenic signalling pathways at the same time to maximize the resulting aberrant cell proliferation, and have significantly aided in the identification of the underlying key regulatory mechanism with which they interfere. The early SV40 gene region including the origin of replication has served as a prototype in parallel with retroviruses for the design of mammalian expression vectors which can randomly integrate into the genome of mammalian cells and allow for high levels of cDNA expression when combined with effective selectable markers. Alternatively, vectors based on the early or late SV40 gene regions have been used for effective transient expression of foreign cDNAs (Topp *et al*., 1980).

In Vitro Cell Lines, DNA Transfection and Cell-permeable Peptides

In vitro cell culture models have found numerous applications in every area of modern molecular medicine since they allow the specific and reproducible investigation of molecular processes in the absence of other complicating factors that are found in more complex *in vivo* systems. In combination with the technique of DNA transfection, this approach allows the transient or stable overexpression of essentially any protein or protein fragment (for which a coding cDNA is available) in a large number of immortalized cell lines. This approach can also be used to interfere with the function of a cellular protein in a dominant-negative approach. This involves an antisense expression plasmid that encodes a single-stranded sequence complementary to the cellular message and results in inhibition of the expression of the respective cellular protein (Baserga and Rubin, 1993). Alternatively, protein domains or smaller proteins can be expressed with an attached cell-permeable peptide (such as of the *Drosophila antennapedia* homeodomain or the HIV TAT protein) that will direct effective uptake across the cell membrane into the cytoplasm and nucleus of cultured cells and even spread the protein throughout a whole organism. With this new experimental strategy overexpression of or interference by proteins and their consequences on host regulatory mechanisms can be studied rapidly in a wide variety of cell lines or tissues (Schwarze *et al*., 1999; Riedel *et al*., 2000).

Insulin-like Growth Factor I Receptor as an Anticancer Target

The combination of these experimental models has led to the new concept of transformation competence genes in cancer research. The insulin-like growth factor I (IGF-I) receptor (IGF-IR) has been identified as a candidate mediator which participates in but is not essential for normal cell proliferation whereas normal IGF-IR levels have been found to be essential for malignant cell transformation. This is suggested by the finding that normal immortalized mouse fibroblasts carrying a disruption in the IGF-IR gene generally cannot be transformed by a

number of different oncogenes. Alternatively, malignant transformation in cell culture, in rodents or in human patients can be reversed with a variety of anti-IGF-IR strategies including antisense constructs, specific antibodies or dominant-negative forms. These approaches only result in limited growth inhibition of normal cells but cause massive apoptosis of tumour cells in particular *in vivo* which should result in a favourable therapeutic index. In addition, they elicit an unidentified novel type of host response that eliminates surviving tumour cells and abrogates tumour growth and metastasis. A better understanding of the involved mechanisms will help in the design of rational therapeutic strategies that may also be favourably directed at downstream targets in the IGF-IR signalling pathway (Baserga, 1998).

erbB-2/neu as a Diagnostic Tumour Marker

The improved understanding of the molecular mechanisms that underlie oncogene action *in vitro* has resulted in diagnostics and the development of therapeutic tools that can be applied to human cancer. Important alterations have been identified that can be consistently demonstrated in a significant fraction of human tumours. These include amplification of the *erbB-2* or *neu* gene as an early example and inactivation of the tumour-suppressor gene *p53*. A novel oncogene termed *neu* was identified as a result of chemical carcinogenesis and was later found to represent an altered form of *erbB-2*, a member of the epidermal growth factor receptor gene family. The alteration was limited to a single nucleotide change in the codon for a critical amino acid in the transmembrane domain. Based on these molecular studies, hybridization probes became available that demonstrated increased expression of normal *erbB-2* in 20–60% of various human tumours. These included breast, lung, ovarian, oesophageal, colorectal and prostate cancer, and *erbB-2* amplification in primary tumours was linked to metastatic lesions (Salomon *et al*., 1995).

erbB-2/neu as a Therapeutic Target

The substantially increased expression levels and the location of *erbB-2* at the cell surface as a membrane-spanning protein, identified it as a potential target for a variety of interfering therapeutic strategies and promised a favourable therapeutic index. These strategies predominantly target the extracellular domain of *erbB-2* with specific monoclonal antibodies that alter *erbB-2* activity, present it to the immune system or deliver toxins or enzymes that can activate prodrugs in the transformed cells. Other classes of strategies target the signalling pathway beginning at the receptor level with specific tyrosine kinase inhibitors, downstream signalling mediators, target expression of *erbB-2* via antisense approaches, or focus on potential vaccines that activate an immune response towards the overexpressed *erbB-2* (Esteva-Lorenzo *et al*., 1998). With the rapid progress in the elucidation of the molecular alterations underlying human tumours, in molecular modelling and rational drug design, exciting new diagnostic and therapeutic strategies are being developed at a rapidly increasing pace (Levitzki, 1994).

Differentiation and Cancer Diagnosis

The analysis of tumour cell differentiation can assist in the diagnosis and therapy of human cancers. Tumour cells may release products into body fluids that allow conclusions about their state of differentiation. These are typically quantified with sensitive immunoassays. Serum levels of human chorionic gonadotrophin or α-fetoprotein greatly assist in the diagnosis of male testicular carcinomas and female gestational neoplasms. Similarly the origin of metastatic neoplasms may be identified through the presence of oestrogen receptors and α-lactalbumin in breast cancer, prostate-specific antigen and prostate acid phosphatase in prostate cancers and myoglobin and desmin in sarcomas. Elevation of these markers in serum is often the earliest sign for the relapse of a given neoplasm but specificity is typically poor. Tissue markers for neuroendocrine differentiation help identify carcinoma patients with improved response to chemotherapy (Gorstein and Thor, 1990).

Differentiation as Cancer Therapy

Terminal differentiation has been induced in some tumour cells in culture and in animals (Reiss *et al*., 1986) which often stably suppresses tumour growth. The drugs involved typically exert only minimal toxic effects on normal cells in contrast to most anticancer drugs (Waxman *et al*., 1988). Several differentiation-inducing drugs have been used clinically including oestrogens and androgens to treat specific breast, prostate and gynaecological tumours that express the appropriate nuclear receptor. Retinoic acid, hexamethylenebisacetamide and 5-azacytidine induce differentiation and inhibit the growth of several types of tumours in laboratory models and myelodysplastic syndrome in humans. All-*trans*-retinoic acid has shown promising initial results in the clinical therapy of promyelocytic leukaemia, while retinoids in combination with interferon-alpha have helped in the treatment of patients with squamous cell cancer of the skin or cervix (Chabner, 1993) (see chapter on *Haematological Malignancies in Cancer Research*). Differentiation by mithramycin and hydroxyurea has been linked to remissions in patients diagnosed with the accelerated phase of chronic myelogenous leukaemia. Retinoids have been used against premalignant lesions or the development of cancer of the breast, cervix, colon, skin, lung or oral cavity with encouraging early results. It is difficult to predict the success of treatment to induce terminal differentiation of specific tumour cells and retinoids have the potential

to promote the formation of new tumours in some experimental models. A better understanding of the molecular controls of differentiation should lead to more accurate predictions and to the rational design of more effective and specific anti-tumour therapies (Rowley *et al*., 1993).

BIOLOGICAL AND TECHNICAL LIMITATIONS

Cell Cultures Versus Nude Mice as Models for Tumorigenesis

In vitro experimental models for tumour growth and differentiation allow the reproducible and rapid dissection of specific molecular mechanisms that participate in the process of malignant cell transformation under tightly defined experimental conditions. This represents only the beginning of the chain of successive steps that will be necessary for tumorigenesis to occur in humans. It allows single experimental parameters to be varied and critically evaluated for their contribution to the outcome of the experiment. This is impossible in humans *in vivo* where tumours are found often difficult to access and in a complex tissue environment that introduces unknown factors that contribute to tumour development and growth. Malignant transformation is typically scored *in vitro* by anchorage-independent growth, the formation of cell foci and the underlying changes in cell morphology. This is frequently correlated with the potential of the affected cells to form tumours in immune-compromised nude mice, which represents a generally accepted standard that defines experimental tumorigenesis such as the activity of the underlying oncogenes. Anchorage-independent growth will, however, not necessarily correlate with tumour formation in nude mice and needs to be interpreted with caution. While nude mice represent an advantageous model to score experimental tumour formation in the absence of a host immune response, they certainly do not realistically reflect the tumorigenic potential when compared with wild-type animals. (See also chapter on *Human Tumours in Animal Hosts*.)

Differences Between Human and Rodent Cells

Most importantly, experimental rodents such as mice and rats differ from humans in many ways that are not well understood with regard to their potential to develop cancer. This is reflected in the well-established fact that experimental antitumour therapies will frequently show remarkable efficacy when they are first tested in rodents whereas subsequent human clinical trials may show little or no effects. Consequently, factors or strategies that cause or prevent tumorigenesis *in vitro* or in experimental animals will not necessarily show the same result in humans. Some fundamental differences have been described at the cellular level. Immortalized cell lines, typically fibroblasts, are routinely isolated from rodent but not from human tissues (**Figure 7**). This explains their popularity as *in vitro* model systems but also exposes their limitations. Immortalized rodent fibroblast cells are preneoplastic and malignant transformation will routinely result from the introduction of a single oncogenic alteration. In contrast, primary human fibroblasts or epithelial cells have only recently been transformed since this requires at least four alterations in critical regulatory mechanisms. One important difference from rodent cells is reflected in the absence of telomerase activity in human cells, which results in telomeric erosion with each cell doubling and is thought to cause senescence and limit the cell life span. (See also chapter on *Genomic Instability and DNA Repair*.)

Specific Genetic Alterations Required for the Malignant Transformation of Normal Human Cells

Since normal murine somatic cells express telomerase activity which is also acquired by human tumour cells, this function is seen as a first critical and necessary change to allow human cells to overcome senescence and gain immortality. In addition to the introduction of telomerase, mitogenic signalling pathways were experimentally deregulated and activated by overexpression of a v-*ras* oncogene, and by interfering with two tumour-suppressor genes, cellular *p53* and *Rb*, which were both accomplished by overexpression of SV40 T antigen. The combination of these changes was minimally required to transform normal human primary cells to cause tumorigenesis in nude mice (Hahn *et al*., 1999). Comparable studies in normal human fibroblasts in which SV40 T antigen was replaced by the functionally similar HPV16 E6/E7 transforming proteins did not result in tumorigenesis (Morales *et al*., 1999). This suggests that either SV40 T antigen contributes an unknown fifth alteration in a mitogenic signalling pathway or that successful tumorigenesis depends on the exact order in which the four alterations are introduced into the cell. These recent landmark findings now offer the first *in vitro* experimental model to identify combinations of components and the required order of their introduction to minimally result in the malignant transformation of normal human epithelial and fibroblast cells. In contrast, normal mouse cells can be effectively transformed by the cooperation of only two oncogenes (**Figure 11**).

Somatic Mutation Theory Versus Tissue Organisation Field Theory of Carcinogenesis

Since 1914, the prevalent carcinogenesis model has been the somatic mutation theory. This postulates the initiation

of malignant transformation in a single somatic cell by a sequence of genetic alterations that are propagated to cell progeny and result in tumorigenesis. While this concept is supported by a large body of data including the successful experiment described above to transform normal human cells by a sequence of defined genetic changes in the same cell, it has not been feasible to delineate the formation of invasive tumours from a single mutated cell. Since tumours are complex, often genetically heterogeneous, and closely interact with their surrounding tissue, the somatic mutation theory cannot be easily validated at the level of the tumour. The organisation field theory of carcinogenesis is based on the postulate that proliferation is the default state of all cells (which is not in dispute for unicellular organisms) even in multicellular organisms. This is taken to imply that quiescence is caused by inhibitory factors or lack of serum nutrients rather than by the lack of growth factor stimuli. Carcinogenesis is seen as a phenomenon at the tissue and organ organisational level, not limited to a single cell. *In vitro* models have not been able to contribute much to clarify this dispute given the complexity of tumours *in vivo*. Both conflicting theories can serve as a reminder that tumorigenesis is a multistage process (**Figure 3**), of which only small segments can be studied in *in vitro* models one at a time (**Figure 11**). *In vitro* models can elucidate the mechanism by which steps in tumorigenesis can occur in a test-tube, but they cannot prove the accurate sequence of events that have occurred in a specific tumour (Sonnenschein and Soto, 1999).

ACKNOWLEDGEMENTS

Part of this work was made possible through support by the National Science Foundation under Grant Nos MCB-9316997, MCB-9696090 and MCB-9808795, by the Juvenile Diabetes Foundation International under Grant No. 197048, by the National Institutes of Health under Grant R01 CA77873 and by the Howard Hughes Medical Institute.

REFERENCES

Aaronson, S. A. (1991). Growth factors and cancer. *Science*, **254**, 1146–1153.

Alison, M. R. (1995). Assessing cellular proliferation: what's worth measuring? *Human and Experimental Toxicology*, **14**, 935–944.

Alison, M. R. and Sarraf, C. E. (1997a). Cell proliferation and cell death. In: *Understanding Cancer: from Basic Science to Clinical Practice*. 133–193 (Cambridge University Press, Cambridge).

Alison, M. R. and Sarraf, C. E. (1997b). The genetic basis of cancer. In: *Understanding Cancer: from Basic Science to Clinical Practice*. 58–96 (Cambridge University Press, Cambridge).

Baserga, R. (1998). The IGF-I receptor in normal and abnormal growth. In: Dickson, R. B. and Salomon, D. S. (eds). *Hormones and Growth Factors in Development and Neoplasia*. 269–287 (Wiley-Liss, New York).

Baserga, R. and Rubin, R. (1993). Cell cycle and growth control. *Critical Reviews in Eukaryotic Gene Expression*, **3**, 47–61.

Begg, A. C. (1993). Cell proliferation in tumors. In: Steel, G. G. (ed.), *Basic Clinical Radiobiology*, 14–22 (Edward Arnold, London).

Beug, H., *et al*. (1994). Insights into erythroid differentiation obtained from studies on avian erythroblastosis virus. *Current Opinions in Cell Biology*, **6**, 816–824.

Chabner, B. A. (1993). Biologic basis for cancer treatment. *Annals of Internal Medicine*, **118**, 633–637.

Cheson, B. D. *et al*. (1986). Differentiating agents in the treatment of human malignancies. *Cancer Treatment Reviews*, **13**, 129–145.

Cifone, M. A., and Fidler, I. J. (1980). Correlation of patterns of anchorage-independent growth with *in vivo* behavior of cells from a murine fibrosarcoma. *Proceedings of the National Academy of Sciences of the USA*, **77**, 1039–1043.

Cross, M. and Dexter, T. M. (1991). Growth factors in development, transformation, and tumorigenesis. *Cell*, **64**, 271–280.

Denizot, F. and Lang, R. (1986). Rapid colorimetric assay for cell growth and survival. Modifications to the tetrazolium dye procedure giving improved sensitivity and reliability. *Journal of Immunological Methods*, **89**, 271–277.

Deschamps, J. and Meijlink, F.(1992). Mammalian homeobox genes in normal development and neoplasia. *Critical Reviews in Oncogenesis*, **3**, 117–173.

Dolle, P., *et al*. (1989). Differential expression of genes encoding alpha, beta and gamma retinoic acid receptors and CRABP in the developing limbs of the mouse. *Nature*, **342**, 702–705.

Esteva-Lorenzo, F. J., *et al*. (1998). The erbB-2 gene in human cancer: translation from research to application. In: Dickson, R. B. and Salomon, D. S. (eds), *Hormones and Growth Factors in Development and Neoplasia*. 421–444 (Wiley-Liss, New York).

Fantes, P. and Brooks, R. (eds) (1993). *The Cell Cycle: A Practical Approach*. (IRL Press, Oxford).

Fearon, E. R. and Vogelstein, B. (1993). Tumor suppressor genes and cancer. In: Holland, J. F., *et al*. (eds), *Cancer Medicine*. 77–90 (Lea and Febiger, Philadelphia).

Fine, H. A. and Haseltine, W. A. (1993). RNA tumor viruses. In: Holland, J. F., *et al*. (eds), *Cancer Medicine*. 265–282 (Lea and Febiger, Philadelphia).

Gorstein, F. and Thor, A. (guest eds) (1990). Tumor markers in diagnostic pathology. *Clinics in Laboratory Medicine*, **10** (1) (Saunders, Philadelphia).

Grana, X. and Reddy, E. P. (1995). Cell cycle control in mammalian cells: role of cyclins, cyclin dependent kinases

(CDKs), growth suppressor genes and cyclin-dependent kinase inhibitors (CKIs). *Oncogene*, **11**, 211–219.

Gu, W., *et al*. (1993). Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation. *Cell*, **72**, 309–324.

Hahn, W. C., *et al*. (1999). Creation of human tumor cells with defined genetic elements. *Nature*, **400**, 464–468.

Hall, P. A., *et al*. (1992). *Assessment of Cell Proliferation in Clinical Practice*. (Springer, London).

Hunter, T. and Pines, J. (1994). Cyclins and cancer. II: cyclin D and CDK inhibitors come of age. *Cell*, **79**, 573–582.

Kern, S. E. (1996) Oncogenes/proto-oncogenes, tumor-suppressor genes, and DNA-repair genes in human neoplasia. In: Sirica, A.E. (ed.), *Cell and Molecular Pathogenesis*. 321–340 (Lippincott-Raven, Philadelphia).

Klein, G. (1993). RNA tumor viruses. In: Holland, J. F., *et al*. (eds), *Cancer Medicine*. 65–77 (Lea and Febiger, Philadelphia).

Levitzki, A. (1994). Signal-transduction therapy. A novel approach to disease management. *European Journal of Biochemistry*, **226**, 1–13.

Li, L., *et al*. (1989). Correlation of growth capacity of human tumor cells in hard agarose with their *in vivo* proliferative capacity at specific metastatic sites. *Journal of the National Cancer Institute*, **81**, 1406–1412.

Lin, C. Q. and Bissell, M. J. (1993). Multi-faceted regulation of cell differentiation by extracellular matrix. *FASEB Journal*, **7**, 737–743.

Martin, S. G. (1996). Normal cells and cancer cells. In: Bishop, J. M. and Weinberg, R. A. (eds), *Scientific American Molecular Oncology*. 13–40 (Scientific American, New York).

McCormick, A. and Campisi, J. (1991). Cellular aging and senescence. *Current Opinions in Cell Biology*, **3**, 230–234.

Mintz, B. and Fleischman, R. A. (1981). Teratocarcinomas and other neoplasms as developmental defects in gene expression. *Advances in Cancer Research*, **34**, 211–278.

Morales, C. P., *et al*. (1999). Absence of cancer-associated changes in human fibroblasts immortalized with telomerase. *Nature Genetics*, **21**, 115–118

Parada, L. F., *et al*. (1982). Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature*, **297**, 474–478.

Pawson, T. and Scott, J. D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. *Science*, **278**, 2075–2080.

Reiss, M., *et al*. (1986). Induction of tumor cell differentiation as a therapeutic approach: preclinical models for hematopoietic and solid neoplasms. *Cancer Treatment Reports*, **70**, 201–218.

Rheinwald, J. G. and Beckett, M. A. (1980). Defective terminal differentiation in culture as a consistent and selectable character of malignant human keratinocytes. *Cell*, **22**, 629–632.

Riedel, H., *et al*. (1989). Cytoplasmic domains determine signal specificity, cellular routing characteristics and influence ligand binding of epidermal growth factor and insulin receptors. *EMBO Journal*, **8**, 2943–2954.

Riedel, H., *et al*. (2000). PSM, a mediator of PDGF-BB-, IGF-I-, and insulin-stimulated mitogenesis. *Oncogene*, **19**, 39–50.

Rodrigues, G. A. and Park, M. (1994). Oncogenic activation of tyrosine kinases. *Current Opinions in Genetics and Development*, **4**, 15–24.

Rosenberg, N. (1994). Abl-mediated transformation, immunoglobulin gene rearrangements and arrest of B lymphocyte differentiation. *Seminars in Cancer Biology*, **5**, 95–102.

Rowley, J. D., *et al*. (1993). The clinical applications of new DNA diagnostic technology for the management of cancer patients. *Journal of the American Medical Association*, **270**, 2331–2337.

Salomon, D. S., *et al*. (1995). Epidermal growth factor-related peptides and their receptors in human malignancies. *Critical Reviews in Oncology and Hematology*, **19**, 183–232.

Schwarze, S. R., *et al*. (1999). *In vivo* protein transduction: delivery of a biologically active protein into the mouse. *Science*, **285**, 1569–1572.

Sherr, C. J. (1994). GI phase progression: cycling on cue. *Cell*, **79**, 551–555.

Sonnenschein, C. and Soto, A. M. (1999). *The Society of Cells: Cancer and Control of Cell Proliferation*. (BIOS Scientific Publishers, Oxford and Springer, New York).

Soto, A. M. and Sonnenschein, C. (1993). Regulation of cell proliferation: is the ultimate control positive or negative. In: Iversen, O. H. (ed.), *New Frontiers in Cancer Causation*. 109–123 (Taylor and Francis, Washington, DC).

Topp, W. C., *et al*. (1980). Transformation by SV40 and polyoma virus. In: Tooze, J. (ed.), *DNA Tumor Viruses*. 205–296 (Cold Spring Harbor Laboratory Press, Plainview, NY).

van der Geer, P., *et al*. (1994). Receptor protein–tyrosine kinases and their signal transduction pathways. *Annual Reviews in Cell Biology*, **10**, 251–337.

Waxman, S., *et al*. (1988). *The Status of Differentiation Therapy of Cancer*. (Raven Press, New York).

Wille, J. J., *et al*. (1982). Neoplastic transformation and defective control of cell proliferation and differentiation. *Cancer Research*, **42**, 5139–5146.

FURTHER READING

Alison, M. R. and Sarraf, C. E. (1997). *Understanding Cancer: From Basic Science to Clinical Practice*. (Cambridge University Press, Cambridge).

Baserga, R. (1995). Measuring parameters of growth. In: Studzinski, G. R. (ed.), *Cell Growth and Apoptosis: A Practical Approach*. 1–19. (IRL Press, Oxford).

Dickson, R. B. and Salomon, D. S. (1998). *Hormones and Growth Factors in Development and Neoplasia*. (Wiley-Liss, New York).

Fingert, H. J., *et al*. (1997). Cell proliferation and differentiation. In: Holland, J. F., *et al*. (eds), *Cancer Medicine*. 3–18 (Williams & Wilkins, Baltimore).

Gallo, R. C. (1995). Human retroviruses in the second decade: a personal perspective. *Nature Medicine*, **1**, 753–759.

Hanahan, D. and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell*, **100**, 57–70.

Hollywood, D., *et al*. (1995). Oncogenes. In: Peckham, M., *et al*. (eds), *Oxford Textbook of Oncology*. 54–73 (Oxford Medical Publications, Oxford).

Heldin, C.-H. (1996). Protein tyrosine kinase receptors. In: Tooze, J. and Parker, P. J. (eds), *Cancer Surveys: Cell Signalling*. 7–24. (Cold Spring Harbor Laboratory Press, Plainview, NY).

Hunter, T. and Karin M. (1992). The regulation of transcription by phosphorylation. *Cell*, **70**, 375–387.

Lemmon, M. A. and Schlessinger, J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerizaton. *Trends in Biochemical Science*, **19**, 459–463.

Murray, A. W. and Hunt, T. (1993). *The Cell Cycle: An Introduction*. (Freeman, New York).

Steel, G. G. (1993). The growth rate of tumours. In: Steel, G. G. (ed.), *Basic Clinical Radiobiology*. 8–13 (Edward Arnold, London).

Wilkins, A. S. (1993). *Genetic Analysis of Animal Development*. (Wiley-Liss, New York).

Internet Sites

http://www.nci.nih.gov

http://www.atcc.org/

http://www.stke.org/

http://ditzel.rad.jhu.edu/

http://www.math.uic.edu/~tier/M494/Tumor/tumor.html

http://cancer.med.upenn.edu/disease/

http://www.bmb.psu.edu/courses/biotc489/

Chapter 62

Angiogenesis Models

Andrew D. Grove, Karen A. Lapidos, Howard Doong and Elise C. Kohn
National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

CONTENTS

- Angiogenesis
- Applications of Angiogenesis Models
- *In Vivo* Models
- *In Vitro* Models
- Utilizing Angiogenesis Models to Study Signalling Events Related to Angiogenesis
- Conclusion

ANGIOGENESIS

Angiogenesis, the formation of new blood vessels, occurs during cancer progression, wound healing and embryogenesis. Like normal cells, tumour cells obtain their metabolic and nutrient requirements such as oxygen, glucose and amino acids from the blood supply. Without this rich nutrient broth, a tumour is unable to grow beyond 2 mm in size. The secretion of angiogenic factors by tumour and stromal cells may induce neovascularization allowing exponential growth of the neoplasm (Folkman, 1970). In addition to providing the means for massive tumour growth, angiogenesis has been shown to be necessary although not sufficient for successful metastases (Liotta *et al*., 1974). The new blood vessels provide both an outlet and a pathway for invasive tumour cells to enter the bloodstream, circulate and extravasate at a distant site to initiate a secondary tumour (Kohn and Liotta, 1995). Mortality from cancer occurs most commonly as a result of complications from the primary tumour and/or as a result of distant metastases. Experimental methods and models have been developed to study the complex process of angiogenesis in malignancy to identify mechanisms and dissect the process. This has led to the discovery and characterization of antiangiogenic therapies that hold promise to stabilize or reduce many types of cancers.

The induction of neovascularization has been shown to be a discrete event during the early stages of tumour development. This angiogenic switch is the rate-limiting step in tumour progression and is hypothesized to be caused by an imbalance of anti- and proangiogenic signals from either the tumour, stromal, endothelial or host inflammatory cells (Hanahan and Folkman, 1996). Growth factors and cytokines such as vascular endothelial cell growth factor (VEGF), transforming growth factor-β (TGFβ), tumour necrosis factor-α (TNF-α), basic fibroblast growth factor (bFGF), angiogenin and interleukin-8 (IL-8) are examples of proangiogenic molecules (Cockerill *et al*., 1995; Hanahan and Folkman, 1996). Promotion of capillary growth also may be induced by insoluble factors such as the extracellular matrix (ECM) components thrombospondin-1 (TSP-1), laminin, fibronectin and types I, III, IV and V collagen, which promote endothelial cell migration, proliferation or differentiation into vascular tubules (Ingber and Folkman, 1989). Potent endogenous antiangiogenic molecules include TGF-β and also small segments of larger, often inactive, parent polypeptides such as angiostatin from plasminogen, endostatin from type XVIII collagen or fibronectin and prolactin fragments (Hanahan and Folkman, 1996). Fumagillin and cyclosporin A are natural antiangiogenic products, and platelet factor 4 is an endogenous antiangiogenic agent (Cockerill *et al*., 1995). The result of a net imbalance of pro- and antiangiogenic factors results in induction or inhibition of capillary sprouting from pre-existing blood vessels. Activated endothelial cells secrete proteases to degrade the local basement membrane after which the cells may be stimulated by matrix components and growth factors elaborated during proteolysis. Endothelial cells migrate towards the tumour, and a capillary sprout is formed. A lumen develops within the newly formed tubes and with anastomosis, the vessel is complete. Maturation of the new capillary bed requires recruitment of pericytes and deposition of new basement membrane (Ausprunk and Folkman, 1977). (See also chapter on *Angiogenesis*.)

APPLICATIONS OF ANGIOGENESIS MODELS

Observations that a cancerous mass requires neovascularization to grow (Folkman, 1970) and that metastasis cannot occur unless the neoplastic cells have access to the

vasculature (Liotta *et al*., 1974) have led to increased interest in the field of angiogenesis. There is an earnest search under way for antiangiogenic agents in response to the hypothesis that inhibition of angiogenesis may lead to the arrest of the growth and metastasis of malignancy. This concerted effort has led to a better understanding of the steps involved in new vessel formation and identification of factors that influence this process. This progress could not have been made without the extensive use of a variety of models of angiogenesis. **Table 1** presents the various uses of models of angiogenesis. The major advantages and disadvantages of each of the model systems that we discuss in this chapter are presented in **Table 2**.

Both *in vivo* and *in vitro* models have been developed in order to provide an environment to understand the components that are involved in angiogenesis. *In vivo* models may more accurately depict the natural events occurring in angiogenesis, but are harder to scrutinize experimentally. Although *in vitro* angiogenesis models may not be true replicates of *in vivo* vasculogenesis, they are invaluable tools with which to begin dissecting the components involved in this process. These models can be used for the testing of various experimental parameters of interest in order to investigate their selective effects on the angiogenic process.

Another important use of both *in vivo* and *in vitro* angiogenesis models is the testing of potential angiogenic modulators. Candidate pro- and antiangiogenic compounds

Table 1 Applications of angiogenesis models

Allow testing of putative stimulatory and inhibitory angiogenic modulators
Allow testing of experimental parameters to investigate their selective effects on the angiogenic process
Allow identification of genes and gene products involved in angiogenesis
Allow identification of potential new antiangiogenic targets

Table 2 Major advantages and disadvantages of various models for angiogenesis studies

Model	Advantage	Disadvantage
In vivo		
Cranial and dorsal windows	Allows non-invasive, multiple time-point measurement of several microhaemodynamic parameters such as blood vessel permeability and blood flow velocity	Requires surgical techniques and specialized equipment for intravital microscopy
Exteriorized tissue models		
1. Intradermal assay	Relatively easy to prepare compared with using other *in vivo* models	Traumatizes the tissue; difficult to distinguish the newly formed vessels
2. Chick chorioallantoic membrane assay	Allows quantification of the effects of putative angiogenic and antiangiogenic molecules; lack of rejection of test substances due to underdevelopment of the chick immune system	Discerning neovascularization from alterations in blood vessel density or reorganisation may be difficult
Corneal micropocket assay	Reliable qualitative and quantitative analyses can be performed using image analysis programme	Expensive; good surgical technique required
Matrix implants		
1. Cannulated sponge model	Ability to quantify blood flow sequentially over a period of time	Gross inspection of neovascularization is not possible
2. Disc angiogenesis sponge model	Multiple discs can be implanted to increase statistical accuracy	Angiogenic responses are determined only at defined time points
3. Matrigel plug	Reproducible and quantifiable data with image analysis or manual blood vessel counts; allows histo- and immunohistochemistry, *in situ* hybridization	Only one time point per Matrigel plug so requires more animals
Retinal assay	Ideal for studying the mechanisms and process of hypoxia-induced angiogenesis	Only one time point per animal so requires more animals; requires special equipment for maintaining oxygen levels
In vitro		
Tubulogenesis in culture	Ability to dissect out the component parts of the angiogenic process	Results may not hold true in an *in vivo* setting
Co-culture model	Ability to determine effects of nonendothelial cell types on endothelial cell tubulogenesis	Variation of results due to endothelial cell type or source
Immortalized endothelial lines	Avoids the limitations of primary endothelial culture	Immortalized endothelial cells may not behave as true endothelial cells would in a given condition

can be distinguished from compounds having mixed or no therapeutic potential through the screening provided by model systems. Different models may isolate individual steps, allowing some functional mapping of inducers and inhibitors of angiogenesis if several models are used. This provides a mechanism whereby candidate compounds can be identified quickly and tested to better understand their potential for the clinical setting.

Angiogenesis models are also critical to our understanding of the molecular events that are involved in the formation of new vessels. Although the biological steps involved in the formation of new vessels have been determined, relatively little is known about the molecular genetic events underlying angiogenesis. New models or modifications of existing ones will allow researchers to isolate RNA, DNA or protein of endothelial cells in specific stages of the angiogenesis process. These materials can then be used to identify genes or gene products that play a role in angiogenesis and therefore potential new targets for the development of antiangiogenic therapy.

IN VIVO MODELS

In vivo models used to quantify neovascularization can be classified briefly into five categories: (1) angiogenesis windows or chambers created surgically for the observation of neovascularization in live animals; (2) exteriorized tissue models, such as the chick chorioallantoic membrane (CAM) and intradermal assays; (3) vascular induction on normally avascular tissues such as cornea; (4) implanted materials to induce angiogenesis, such as the Matrigel plug; and (5) physiopathological induction of angiogenesis, such as retinopathy of prematurity. In the following section, these frequently used *in vivo* models will be discussed. A review of the major advantage and disadvantage of each of the *in vivo* models is presented in **Table 2**.

Angiogenesis Windows

There are several types of angiogenic windows such as the cranial window, dorsal skin fold chamber, rabbit ear chambers and the hamster cheek pouch window. The cranial window and dorsal skin fold chamber assays will be the focus of the current discussion. A cranial window is accomplished by drilling a hole on the apex of the skull in an anaesthetized animal, allowing access to the subdural space for placement of the proangiogenic materials, such as a tumour xenograft. A protective closure is applied that allows intravital microscopy for real time observation of outcome (Dellian *et al*., 1996). Quantification of angiogenesis in window assays requires the placement of the anaesthetized animal under a stereomicroscope and images taken for analysis. Many quantitation approaches are used after the initial image analysis including point counting and vascular area.

The paracrine antiangiogenic effects of a prostatic tumour xenograft was observed using the angiogenic window assay in an example of the use of this model. In this variation of the assay, a gel enriched with a proangiogenic molecule(s) such as bFGF was embedded between two pieces of nylon mesh, and implanted inside the cranial window (Sckell *et al*., 1998). The nylon mesh facilitated quantification of blood vessel growth as it provided a grid with which to count newly formed vessels. Tumour tissue was then implanted into the flank of severe combined immunodeficient (SCID) mice and the angiogenic response at the cranial window observed. In this study, the authors demonstrated that the initial implanted tumour volume correlated positively with the inhibition of angiogenesis at the cranial window. These data suggest that the tumour cells in the flank of the mice secreted antiangiogenic factors that suppressed angiogenesis at the cranial window. It is advantageous that *in vivo* quantification of angiogenesis and tumour size can occur concurrently, and that these two biological endpoints can be evaluated at a site other than the primary tumour (Sckell *et al*., 1998).

The dorsal skinfold chamber assay is similar to the cranial window except that Teflon S-treated aluminium frames with spacers between the frames are used to create the angiogenic window. The window-containing frame is used to push out a double layer of skin. A circular area of skin is excised from the upper layer of skin to make the window while the lower layer of subcutaneous tissue remains intact. The remaining frame of the chamber is inserted on the upper layer of skin to complete the window. Catheters are surgically implanted into the external jugular vein of the animal to allow blood flow velocity studies in the conscious animal (Endrich *et al*., 1980). A glioma model of neovascularization required the implantation of a rat glioma cell line on the lower layer of skin of the dorsal skinfold. A fluorescent neuronal tracer was used to differentiate tumour cells from the alternatively stained vasculature such that both tumour and blood vessels could be quantified.

Intravital microscopy (**Figure 1**) makes the real-time quantification of angiogenic window assays possible. Fluorescein isothiocyanate (FITC)–dextran stain of the blood vessels provides contrast against surrounding tissue allowing digitized video imaging analysis of blood vessel length and density (Endrich *et al*., 1980). Quantitation of angiogenesis in this assay is non-invasive and may occur at multiple time points during the life of the animal rather than only post mortem. More recently, multiple microhemodynanic parameters have been quantified such as blood flow velocities, perfusion, patency, microvessel length and diameter and vascular permeability. In addition, serial quantification for up to 4 weeks has been shown to be possible (Sckell *et al*., 1998). Both blood vessel growth and the functionality of the neovessels can be investigated through intravital microscopy of the angiogenic windows.

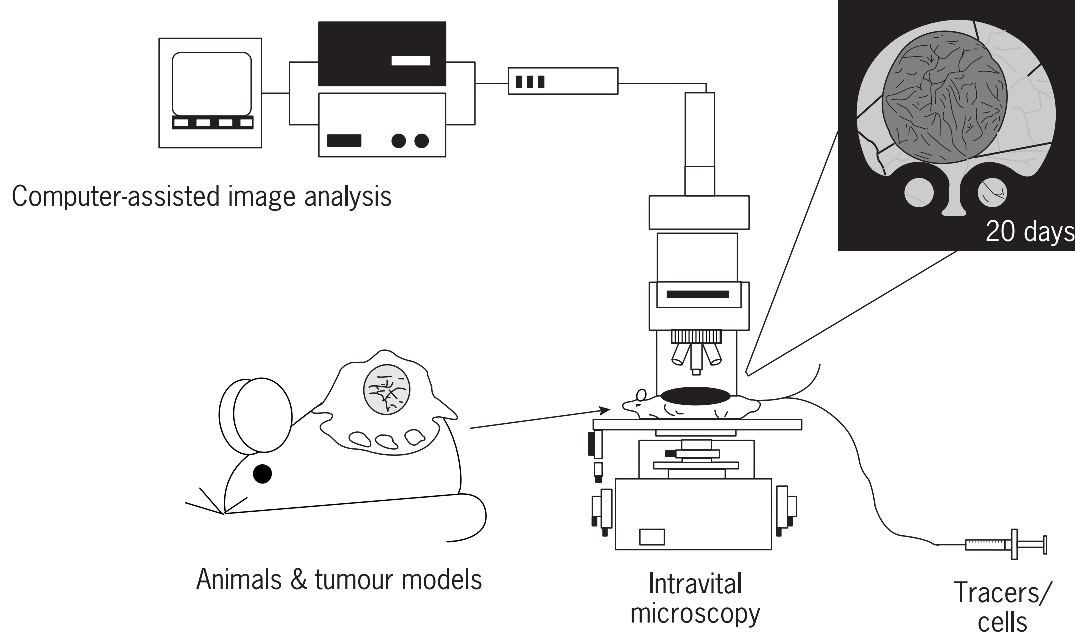


Figure 1 Diagram of the use of angiogenesis windows for intravital microscopy. In this example, an animal with a dorsal fold angiogenic window is anaesthetized and placed under a stereomicroscope connected to video imaging equipment. Insertion of a catheter and administration of stains or tracers allows distinction of the vasculature from other cell types. Using this method, real-time quantification of angiogenesis assays is possible. (Adapted from Jain *et al.*, 1998.)

Hence these assays are particularly powerful *in vivo* angiogenesis models.

Exteriorized Tissue Models

Intradermal Assay

This is a method of innoculating tumour cells intradermally to quantify tumour-induced angiogenesis. Animal skin first has to be separated from its underlying fascia. A triangular skin flap is created in anaesthetized mice by making a skin incision along the midline of the abdomen and extending it to the groin. A defined number of tumour cells or a pellet containing the candidate angiogenic agent is then injected intradermally. At certain time points, the skin flap is everted to measure the angiogenic response, and then replaced. Angiogenesis can be monitored every 2 days for a period of 2 weeks (Runkel *et al.*, 1991). The major concern for this assay is the trauma caused by periodic disruption of the flap. Also, it is difficult to identify all the newly formed blood vessels.

Chick Chorioallantoic Membrane Assay

The chick chorioallantoic membrane (CAM) assay allows the investigation of neovascularization in a developmental setting. This assay lends itself well to the study of pro-angiogenic or inhibitory macromolecules. Three days postfertilization, a window is placed in the shell or the egg contents are removed to a petri dish. Primitive blood vessels developed by day 3 in the extraembryonic membrane are now part of the chorioallantois and can be either stimulated or blocked experimentally. There are three distinct phases of development in the CAM assay. Days 5–7 are characterized by capillary sprouting, days 8–12 by the formation of intercapillary posts that grow to new capillary mesh formations and during days 13 and 14 blood vessels have begun to mature. Final maturation of this *de novo* CAM angiogenesis occurs by about day 17. Test compounds can be applied in methylcellulose discs or filter-paper or dried on a plastic cover-slips and placed on the CAM. Alternatively, inert synthetic polymers allowing sustained release of molecules may be used. Windowed eggs are then sealed or externalized contents moved to an incubator. CAMs are observed for changes in angiogenic development. Tumour cells can be used as angiogenic stimulants for studying the effects on the angiogenic response. An example of angiogenesis inhibition is shown in **Figure 2; see colour plate section**.

A recent advance of the CAM assay has involved the use of a collagen gel on the CAM into which capillaries are induced to grow vertically, thus allowing discrimination of induced capillaries from existing CAM vessels. In one system, a suspension of stimulant or growth factor such as bFGF- and sucralfate-enriched

type I collagen was placed over a mesh that is implanted on day 2 about 2–3 mm away from primary blood vessels in the CAM. A second, smaller layer of mesh is then added to the gel. On days 3–9 the number of squares in the upper grid that contain blood vessels is determined and then expressed as a percentage of the total number of grid squares. Inhibitors of stimulant-induced angiogenesis may be tested by incorporation into the upper layer of mesh, above the collagen gel. Tumour cells in suspension can also be added to the upper mesh to determine the capacity of these tumour cell lines to induce angiogenesis. This preparation is ideal for testing antiangiogenic agents. One advantage of this modification to the CAM assay is that quantitation is rapid. Up to 60 embryos may be quantitated per hour. In addition, the use of image analysis or photography, dyes or histological sectioning is not necessary for quantification (Nguyen *et al*., 1994).

When evaluating neovascularization of the CAM assay, one must be aware of several factors that may effect the angiogenesis. For example, a nonspecific inflammatory response may be induced during preparation of the CAM which results in false positives (Nguyen *et al*., 1994). The incubation environment must be tightly controlled because of environmental factors such as local hypoxia or hypoglycaemia that effect capillary formation (Cockerill *et al*., 1995). Discerning true neovascularization from changes in vascular density or blood vessel rearrangement may be difficult. Interpretation of the results may be dependent on what day(s) of development is used for the assay as a change in pattern of growth may be misinterpreted as an experimental result when it may be physiological. Lastly, some carrier molecules have been reported to induce their own angiogenic response. Despite these apparent disadvantages, the CAM assay is a frequently used *in vivo* angiogenesis model because of its several advantages such as the ability to quantify angiogenesis, the lack of rejection of implanted discs, sponges or cells because of the underdevelopment of the chick immune system and the ability to use the test materials for immunohistochemistry, electron and light microscopy.

Models of Induction on Normally Avascular Tissues

The classical assay of this category is the rabbit corneal micropocket assay, introduced in 1974 (Gimbrone *et al*., 1974). Because the cornea is an avascular tissue, it provides a natural site for continuously monitoring new blood vessel formation. Mouse and rat corneas were later also used for angiogenesis studies. The corneal bioassay requires implantation of tumour cells or angiogenic factors in the cornea. A surgical pocket is made into which the angiogenic factor, tissue to be tested or control is placed. Most commonly, an inert material pocket is implanted. The two most often used polymers are poly (2-hydroxylethyl methacrylate) (Hydron) and ethylene–vinyl acetate copolymer (Elvax). The polymer pellet will release its agent slowly without causing inflammation. Angiogenesis is allowed to proceed for 5–7 days after implantation, during which time animals can be studied by lamp examination or killed for histological examination (**Figure 3; see colour plate section**). A positive response is recorded if ingrowth of capillary loops or sprouts is detected. The vessel length and number can be determined when the vessel development is in its early stages. However, it becomes more difficult to quantify as the vessels grow more three-dimensionally with time. Some researchers have used the methods of perfusing the cornea with colloidal carbon and computer-assisted imaging to quantify vascular length.

This model has been used to test the hypothesis that tumours are angiogenesis dependent. Gimbrone and co-workers demonstrated that when tumour spheroids were free floating in the anterior chamber of the eye, they remained dormant (Gimbrone *et al*., 1974). When tumour spheroids contacted the iris, they became vascularized and grew as a three-dimensional tumour mass. These authors further demonstrated that the tumour-induced angiogenesis required the production of angiogenic factors using the rabbit corneal model.

Because the cornea is avascular, any vessels that penetrate into the corneal stroma can be easily identified as induced vessels. The advantage of using the cornea as a model is that both qualitative and quantitative analysis can be performed and the results are reliable, reproducible and interpretable. Using this model, one can also compare the effects of different angiogenic factors in the same animal. The two major disadvantages of using animal corneal models are that they are expensive and good surgical technique is required. Although it is cheaper to use rats or mice than rabbits, the surgery is more difficult. Because the rat or mouse cornea is thinner than that of the rabbit, it is more difficult to study three-dimensional growth of vessels in the former.

Matrix Implants

Matrix implants consisting of biocompatible polymers have become a popular method to study angiogenesis. These implants are porous and initially do not contain blood vessels. The materials used to make these implants include polyesters, polyurethane, poly(vinyl) alcohols (Ivalon), gelatin (Gelfoam), Matrigel and collagen. After the matrix implants have been placed subcutaneously into the animals for a period of time, neovascularization can be accessed by vascular filtration into the implant (histologically), by vascular density (morphologically) or by haemoglobin or DNA contents (biochemically). The most commonly used implant models are the cannulated sponge, the disc angiogenesis system (DAS) and the Matrigel plug.

Cannulated Sponge Model

Andrade and colleagues were the first to use the cannulated sponge implants (Andrade *et al*., 1987). The sponge model resembles the corneal model in the ability to introduce modulators of angiogenesis. The cannulated sponge model uses a ^{133}Xe clearance technique to quantify blood flow through the implanted sponges over time. The loss of ^{133}Xe from the sponge occurs as a result of neovascularization and outflow of ^{133}Xe from the source. This model has been used to demonstrate that inflammatory polypeptides, such as substance P and bradykinin, could induce neovascularization. Their neovascular response was inhibited by antagonists targeted to the particular receptor subtypes. When tumour cells were put into the sponge implants, nitric oxide (NO) synthesis increased. The increase in NO synthesis is related to neovascularization in the sponge and is important in maintaining tumour blood flow. There are several advantages of using the cannulated sponge model. It can be used to measure blood flow repetitively and sequentially over time (2–3 weeks), because the injected ^{133}Xe is quickly cleared out of the implant within 2 h. Cytokines, growth factors and inflammatory agents can be retrieved from the sponges, thus facilitating the elucidation of the roles of these agents in angiogenesis. The differences of granulation neovasculature and tumour neovasculature may be studied in detail, because either cytokines or tumour cells can be incorporated into the sponges. The model allows the testing of compounds administered locally into the sponges or systemically. The major disadvantage of using this model system is that, unlike the corneal micropocket assay, gross inspection of sponge neovacularization is not possible. Supporting histological evidence of neovascularization in the sponge is required.

Disc Angiogenesis System

The DAS was first developed by Fajardo and co-workers after their initial observation of vascular growth in subcutaneous implanted sponges (Fajardo *et al*., 1988). The implanted sponge is covered with cell-impermeable filters (45 μm) leaving only the sandwiched sponge available for cell penetration. Test agents are incorporated into the sponge in a slow release polymer through a hole in the sponge. A 7–21-day incubation is used for a 13-mm implanted disc. Cell proliferation can be determined by [^{3}H]thymidine incorporation into the disc using intraperitoneal injection prior to killing the animal. The disc system has been modified to allow the introduction of viable cells. Tumour cells or inflammatory cells can be put into the central hole of the disc at various times after implantation. Cytokine stimulation of vascularization was demonstrated in this model system. Epidermal growth factor (EGF), bFGF and prostaglandin E1 were shown to stimulate endothelial cell proliferation and increase microvessel formation. The major advantage of the DAS is that multiple discs can be used for different time and dose points, which allows reproducible measurements of vascular growth and increases statistical accuracy. The disc system also allows the independent measurement of proliferation and migration of endothelial cells and fibroblasts. The DAS is relatively inexpensive and angiogenesis can be easily quantified using histological, autoradiographic and morphometric methods of analysis. Its disadvantage is that the kinetic process of angiogenesis is difficult to determine using this model system. The DAS is more restricted to evaluate responses at a defined time point.

Matrigel Plug Assay

Matrigel is a crude basement membrane extract containing multiple extracellular matrix proteins and growth factors such as laminin, type IV collagen, fibronectin, entactin, TGFβ, bFGF, EGF and platelet-derived growth factor (PDGF). Liquid Matrigel is injected into the animal at a subcutaneous location or into the mammary fat pad, where it solidifies at physiological temperature. Putative angio-stimulatory or angio-inhibitory molecules may be added to the Matrigel, or tumour cells can be injected concomitantly with the Matrigel mixture. It has been shown that Matrigel alone is unable to induce vascularization although the presence of inflammatory cells may be detected in the Matrigel. Matrigel injected with acidic fibroblast growth factor (aFGF) or basic fibroblast growth factor (bFGF) results in an intense stimulated vascularization, optimized by the addition of heparin. Heparin binds the FGFs to present a more favourable conformation to the FGF receptor (Passaniti *et al*., 1992). The progression of neovascularization of the Matrigel plug initiates with endothelial cells invasion into the Matrigel by 24 h. The first evidence of neovascularization occurs by 48 h with capillary-sized vessels detectible by 72 h. Blood vessels continue to grow until day 3–4, after which there is a plateau, and vessels are maintained until the eighth day (Passaniti *et al*., 1992). Excision of the Matrigel plug from the animal allows for histological analyses, immunohistochemical staining and quantitation of results. The presence of endothelial cells can be confirmed by indirect immunofluorescence or immunostaining for von Willebrand factor, CD31 or other angiogenesis markers. Trichrome–Masson stain will stain haemoglobin, confirming blood flow into the newly formed vessels. Quantification of the vascular tubes that have invaded into the Matrigel plug can be accomplished with digitized image analysis of immuno- or Trichrome–Masson-stained tissue. Advantages of this method include the efficiency and reproducibility of the assay and the ability to quantify the blood vessel growth. Implantation and excision of Matrigel is simple, and both angio-stimulatory and angio-inhibitory agents may be tested (Passaniti *et al*., 1992). This model is also an important tool to investigate the vascularization

induced by weakly tumorigenic cell lines. It has been found that concomitant injection of tumour cells and Matrigel augments angiogenesis and may also increase metastatic potential.

Physiopathological Induction of Angiogenesis

Vascularization of the superficial and deep layers of the retina is tightly controlled by oxygen tension and cytokine levels in the local microenvironment. When neuronal metabolism increases in these areas, local hypoxia results stimulating astrocytes on the inner surface of the retina or Müller cells in the deep layer of the retina to secrete vascular endothelial growth factor (VEGF). Capillaries that originate from the optic disc are then induced to grow and eventually branch out to cover the inner surface of the retina to form the superficial layer. The deep layer of retinal vessels is formed from expansion of nearby superficial layer blood vessels. Improved oxygen tension due to new blood vessel flow secondarily downregulates VEGF production and halts new blood vessel formation (Stone *et al.*, 1995). Retinopathy of prematurity (ROP) occurs in premature infants who have required high concentrations of oxygen due to premature lungs. This hyperoxia often leads to blindness. In the first stage of ocular injury, excessive oxygen in the environment causes endothelial cell apoptosis in the retina. This loss of retinal vasculature results in retinal ischaemia when the high administered oxygen concentration is later reduced. The lowered oxygen tension creates a relative hypoxia, even though the administered oxygen may still be supraphysiological. This hypoxia induces the formation of new blood vessels which are frequently abnormal and leaky, causing retinal detachment and, ultimately, blindness (Alon *et al.*, 1995). Consistent with the importance of VEGF in physiological retinal vascularization, it has recently been demonstrated that VEGF expression is downregulated during hyperoxia and increases during relative hypoxia that occurs during return to normal ambient oxygen tension. VEGF increases where new blood vessels begin to form and attenuates when the capillaries are complete. Studies are ongoing to determine the involvement of other growth factors.

Several animal models of ROP including mouse, rat and rabbit have been established to understand better oxygen regulation of angiogenesis. In animal models of hypoxia-induced angiogenesis, postnatal day-5 animals are exposed to hyperbaric oxygen for several days and then removed to room air, stimulating a relative hypoxia. The blood vessel patterns and integrity and growth factor production in the retina are studied (Alon *et al.*, 1995). (See chapter on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*.) Retinal vascularization has also been studied using transgenic mice expressing growth factors or hormone antagonist genes or using implantation of bFGF-impregnated microspheres in rabbit eyes. Recently, it was determined that excessive CO_2 also can stimulate neovascularization of the retina, but the mechanism of this stimulation is unknown. Human retina specimens of ROP or diabetic retinopathy patients have also been studied. These models are important to oncology because the physiology may parallel that of tumour-induced neovascularization. Both oxygen and recently VEGF have been shown to play a role in retinal and tumour-induced neovascularization. In addition, blood vessels in the retina that are produced under the conditions of normal oxygen concentrations preceded by excessive oxygen are often leaky and abnormal like tumour-induced vessels. Limitations of this model are similar to other *in vivo* angiogenic assays such as the fact that only one time point can be used for each set of retinas.

IN VITRO MODELS

In vitro models of angiogenesis offer several advantages over *in vivo* models including the ability to control experimental conditions such as the presence or absence of angiogenic inducers, the ease of observing angiogenic events, the possibility of dissecting out component parts and the lower cost. Disadvantages of the *in vitro* model systems include the variety and overlap of the reagents used such as endothelial cell type and passage number, extracellular matrix, growth media and supplements and the possibility that the results obtained in an *in vitro* setting may not hold true *in vivo* (see chapter on *Basic Tissue Culture in Cancer Research*). Thus, data gathered using *in vitro* models will need to be confirmed in an *in vivo* setting. Differences in the type and source of primary endothelial cell lines may play the major role in the variability of *in vitro* models. Factors that affect endothelial cell behaviour in a given set of culture conditions include species of origin, organ of origin, age of the organism, macrovascular versus microvascular endothelial cells and arterial versus venous endothelial cells, among others. A standardization of *in vitro* models is recommended. The following is a brief review of some of the classical and more common *in vitro* models in use today. A review of the major advantages and disadvantages of each of the *in vitro* models is presented in **Table 2**.

Tubulogenesis in Culture

The spontaneous tubulogenesis model was the first model of angiogenesis to use primary culture endothelial cells to form vasculature structures *in vitro* (Folkman *et al.*, 1979). Folkman and colleagues observed that human umbilical vein endothelial cells (HUVECs) remain as a confluent monolayer on plastic in culture. If so maintained for a prolonged period of time (10–14 days) with associated

depletion of media nutrients, capillary-like structures form spontaneously. During this process, the majority of the endothelial cells die as the nutrient-rich medium is depleted. However, a small subpopulation of endothelial cells differentiate and form quiescent tubules loosely attached to the culture flask (Cockerill *et al.*, 1995). The pertinent stimulating factor(s) in that setting has not been identified. This initial *in vitro* tubulogenesis model may be cumbersome for use in the dissection of vasculogenesis but the recognition of its inherent value led to the development of the newer easier models.

The placement of endothelial cells on a biologically relevant substrate invites tubule-promoting interactions. Several models have been developed which do this. Kubota and co-workers first demonstrated the endothelial cell-differentiating effect of Matrigel (see Baatout, 1997, for a review) after observing that endothelial cells plated on the matrix formed tubule-like structures after 16–24 h in culture (Kubota *et al.*, 1988). **Figure 4** shows a typical progression of events during human umbilical vein endothelial cell differentiation on Matrigel. Within 1 h, endothelial cells have migrated toward each other and begun to align into a reticular network. Cellular morphology changes after 2 h, after which the cells have started to flatten and elongate. The cells have formed a network of structures on the surface of the substratum after 11–16 h. Kubota and colleagues demonstrated these structures contain lumens as visualized by electron microscopy (Kubota *et al.*, 1988). The lumens are each formed by one to two endothelial cells and contain various amounts of cytoplasm, suggesting that the process of tube formation is a result of rapid cellular remodelling. The differentiated endothelial cells still retain endothelial cell markers such as Weibel–Palade bodies and secrete von Willebrand factor, take up dioctadecyltetraethylindocarbocyanine perchlorate acetylated low-density lipoprotein (di-I-LDL), and release prostacyclin (PGI_2) upon stimulation with bradykinin. The cytoskeleton can be seen to have reorganized into bundles of actin filaments oriented along the axis of the tubes. The addition of transcriptional inhibitors prevents tube formation on Matrigel, indicating that gene transcription is required for tubulogenesis. The addition of cycloheximide was found to inhibit tube formation, indicating that protein synthesis also is important in this process. It has been shown also that pertussus toxin blocks differentiation of endothelial cells on Matrigel, suggesting that the process requires receptor-linked heterotrimeric G protein signalling. Activation of protein kinase C (PKC) may also be required in this model.

Since the majority of the capillary-like structures form on top of the gel, quantitation of the structures may be relatively easy, especially with the advent of imaging technology with computer-assisted analysis. Because of the relative lack of invasion of the cells into the substratum, the Matrigel model of angiogenesis may depict events that occur at later stages of angiogenesis *in vivo* and as such represents a differentiation model rather than a proliferation or invasion model. A further benefit of this model is that the majority of the cells plated on Matrigel are stimulated to differentiate so the cellular response is much more uniform than with other *in vitro* models. A disadvantage of this model is that Matrigel can be subject to variability between batches that yields differences in endothelial cell assay results.

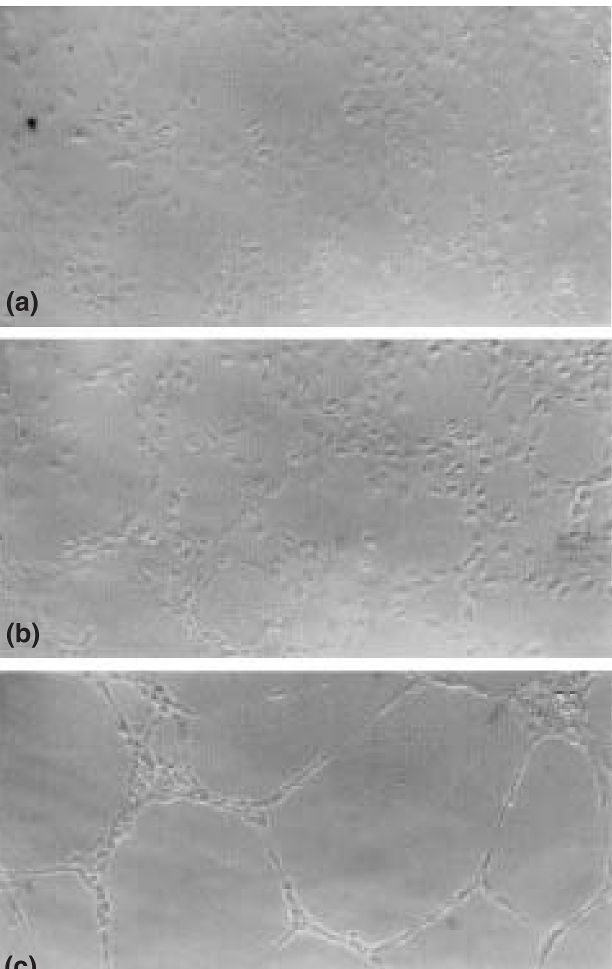


Figure 4 The formation of HUVEC tubules on Matrigel is time dependent. HUVEC cells plated on Matrigel have begun to align after 1 h of culture (a). By 2 h, the HUVECs have begun to flatten and elongate (b). A strong network of tubules is present by 12 h (c).

Tube formation also can occur when endothelial cells are plated on type I collagen gels. The process is completed within 24 h and requires both transcription and translation (Montesano and Orci, 1985). Endothelial cells are seeded on to a thick gel of type I collagen requiring the presence of phorbol myristate acetate (PMA). This induces the cells to invade into the matrix, mimicking the early events seen during angiogenesis *in vivo*. Some cells remain as a monolayer on the surface of the gel and others form

capillary-like structures throughout the gel making quantitation of these structures more difficult. Endothelial cells plated on type I collagen in the absence of PMA proliferate and form a confluent monolayer, but do not differentiate. The requirement for PMA induction of tubulogenesis implies that PKC is involved in this process. Recent evidence suggests that MAP kinase and phosphoinositide 3′-kinase, two parallel signalling pathways, are also involved in tubulogenesis in this model.

A similar model uses a fibrin substrate, a matrix also associated with the extracellular matrix compartment. Endothelial cells are seeded on to a thick fibrin gel in the presence of PMA and fibrinolytic inhibitors, since the presence of PMA alone induces the cells to lyse the fibrin matrix (Montesano *et al*., 1987). Under these conditions, endothelial cells invade the fibrin matrix and form tubules within 3–4 days, again beneath a confluent monolayer. A similar role for PKC is suggested in the process as endothelial cells cultured on fibrin in the absence of PMA fail to invade the gel and proliferate on the surface, eventually to form the confluent monolayer. Not all cells are induced to invade and differentiate and the quantification of those that do is difficult owing to invasion of the gel and its depth.

Nicosia and Ottinetti (1990) showed that 1-mm circumferential sections of rat aorta embedded into a matrix such as collagen, fibrin or Matrigel yields radial outgrowth of microvessel-like structures. This process can be monitored by phase microscopy or by an image analysis-based approach. Vessel outgrowth increases for approximately 10 days and then attenuates. This model contains all three parts of the angiogenic process–invasion, proliferation and differentiation. However, since the aortic ring is most often embedded in a matrix, this model is best suited to investigate the effects of agents on invasion. Some drawbacks of this model include variability of microvessel formation depending on the section of aorta from which the ring was cut and differences in microvessel formation in areas of the ring itself. It is not clear which cells outgrow to form the tubules, the vasa vasorum microvascular endothelial cells or the aortic endothelium.

Co-culture Models

In vitro models have been developed to determine the influence of the local cellular microenvironment on tubule formation since angiogenesis *in vivo* cannot occur in isolation of other cells. Astroglial cells have been shown to induce tubule formation of brain capillary endothelial cells (Laterra *et al*., 1990), thus providing a way to study the development of neural microvessels. Endothelial cells are introduced 24 h after the seeding of C6 astroglial cells on to fibronectin-coated wells. Formation of capillary-like structures occurs within 24 h of co-culture and requires both transcription and translation. This induction of angiogenesis required direct cell–cell contact as no differentiation occurred when the two cell types were separated into adjacent chambers where they were not allowed to touch, but were bathed in the same growth medium. Other co-culture cells have been used successfully in this paradigm to induce endothelial cells to form capillary-like structures without the requirement for direct cell–cell contact including oesophageal tumour cells, keratinocytes and fibroblasts. The oesophageal tumour cell and keratinocyte studies identified TGF-α as an important angiogenic stimulator acting through a paracrine pathway.

Immortalized Endothelial Lines

Some investigators use immortalized endothelial cell lines for *in vitro* work to circumvent the limitations of primary endothelial cell culture. These limitations include a defined life span, strict culture requirements and the presence of functional and structural endothelial cell heterogeneity between and within species and organ sources. Differences in endothelial cells within species occur according to the vessel subtype and organ location. Clonal cell lines offer the advantage of uniformity and reproducibility of cultures, suitability for large-scale experiments and the ability to use recombinant DNA techniques including stable transfections. A major disadvantage of using immortalized endothelial cells is the fact that they may lose characteristics associated with primary culture endothelial cells and therefore may not behave as true endothelial cells would in a given condition.

The EA hy926 cell line is a hybrid line formed from a HUVEC primary culture fused with the A549/8 human lung carcinoma cell line (Edgell *et al*., 1983). This clonal cell line expresses a wide range of differentiated endothelial cell properties including expression of von Willebrand factor protein, Weibel–Palade bodies, prostacyclin production, tissue plasminogen activator and plasminogen activator inhibitor type 1 secretion and uptake of di-I-LDL. Bauer and colleagues in 1992 used the EA hy926 line in the *in vitro* Matrigel model demonstrating that cells underwent differentiation and formed tube-like structures over the course of 16 h in culture. They showed that the tube-like structures contained lumens.

The ECV304 cell line was established by the spontaneous transformation of a HUVEC subline (Takashaki *et al*., 1990). It is characterized by a high proliferative potential without specific growth factor requirements but still retains anchorage dependence with contact inhibition. The original ECV304 cells express endothelial-specific Weibel–Palade bodies and also antigens recognized by human anti-endothelium antibodies including PHM5 and HAM56. The cells do not express several endothelial-specific proteins including von Willebrand factor, CD-31 and CD-34, nor do they efficiently uptake di-I-LDL. When seeded on Matrigel, ECV304 cells undergo differentiation and organize rapidly into an extensive network of tube-like structures within the

course of 1 h. These structures were shown by electron microscopy to contain patent lumens. In mid-1999, an alert was broadcast by the American Type Culture Collection (Rockville, MD, USA) indicating that cells available from at least one ECV304 source were actually the T24/EJ human transitional cell bladder cancer line and that ECV304 should be considered a variant of this line derived by cross-contamination. Users of ECV304 should validate their clone before proceeding.

The HMEC-1 cell line is an immortalized human dermal microvasculature endothelial cell (HMEC) derivative that was established by transfection with a plasmid containing the simian virus 40 large T antigen (Ades *et al*., 1992). The cells are able to proliferate in the absence of serum and still retain anchorage dependence with contact inhibition. HMEC-1 cells retain endothelial cell characteristics including von Willebrand factor expression, uptake of di-I-LDL, CD-31 expression and expression of CD-36, a microvasculature-specific marker. HMEC-1 cells differentiate on Matrigel and form a well-developed network of tube-like structures after 18 h.

UTILIZING ANGIOGENESIS MODELS TO STUDY SIGNALLING EVENTS RELATED TO ANGIOGENESIS

Angiogenesis is not regulated solely by peptide growth factors. Extracellular matrix components such as collagen and adhesion molecules also regulate angiogenesis (Bischoff, 1995). It is therefore logical to use models of angiogenesis to incorporate these different and important signals. The main focus of this section is to show an example how these models were used to study signalling pathways that are related to VEGF-induced angiogenesis.

VEGF is a 46-kDa homodimeric protein that was found to bind to endothelial cells. It stimulates endothelial proliferation and migration and regulates vascular permeability. Hypoxia is known to induce VEGF production and hence it is regarded as an important physical regulator of the endothelium. VEGF binds with high affinity to two receptor tyrosine kinases, the Flt-1 (fms-like tyrosine kinase) and Flk-1/KDR (kinase insert domain-containing receptor). Both Flt-1 and KDR contain intracellular kinase domains and seven immunoglobulin-like domains in their extracellular N-terminal regions. Both receptors are expressed predominantly but not exclusively on endothelial cells.

Friedlander and co-workers used the CAM model to demonstrate that PKC is downstream of the VEGF signalling pathway (Friedlander *et al*., 1995). The PKC inhibitor calphostin C blocked angiogenesis induced by VEGF, PMA and TGFα in the CAM system. Using both the corneal micropocket assay and CAM models, they further demonstrated that *in vivo* angiogenesis initiated by growth factors is influenced also by different types of integrins (Friedlander *et al*., 1995). VEGF-induced rabbit corneal angiogenesis is inhibited by an antibody against $\alpha_v\beta_5$ but not against $\alpha_v\beta_3$ integrins. In contrast, bFGF-induced angiogenesis is inhibited by an antibody against $\alpha_v\beta_3$ in rabbit cornea. These data indicate that these two agonists stimulate *in vivo* angiogenesis in the same model system but through independent and parallel signalling pathways. Their results also argue that adhesion-mediated signals are important in growth of new blood vessels. (See also chapters on *Models for Tumour Cell Adhesion and Invasion*; *Models for Tumour Cell–Endothelial Cell Interactions*.)

Ziche and colleagues reported that the signal transduction pathway for VEGF-induced angiogenesis in the rabbit cornea model differed from that for bFGF-induced angiogenesis through studies of the nitric oxide (NO) system (Ziche *et al*., 1997). They administered an NO synthase inhibitor, N^{ω}-nitro-L-arginine methyl ester (L-NAME), through an implant sewn on the rabbit cornea, and found that inhibition of NO synthesis blocked VEGF-induced but not bFGF-induced corneal angiogenesis. Their studies support the evidence that VEGF and bFGF induce angiogenesis through different signalling pathways.

Hypoxia is known to induce angiogenesis as described above. Ozaki and co-workers used the ischaemic retina model to demonstrate that a hypoxic environment stimulates the production of hypoxia inducible factor-1α (HIF-1α) (Ozaki *et al*., 1999). HIF-1α is a transcription factor that plays an important role in regulation of hypoxia-responsive genes, such as VEGF. They found that the increased expression of HIF-1α in ischaemic retina is temporally and spatially correlated with the expression of VEGF, implying that the synthesis of HIF-1α is upstream of VEGF signalling pathway. The retina model has also been used to demonstrate that VEGF acts as a survival factor for newly formed retinal vessels. Alon *et al*. (1995) reported that hyperoxia induced regression of retinal capillaries by promoting apoptosis of endothelial cells and by downregulating basal production of VEGF. When eyes were injected with VEGF in the setting of experimental hyperoxia, VEGF prevented apoptotic death of endothelial cells and rescued the retinal vasculature. In summary, using angiogenesis models, researchers have begun to map out the pathway of VEGF signalling in endothelial cells. Without the flexibility that angiogenesis models afford, it is probable that little progress would have been made in understanding the process of angiogenesis.

CONCLUSION

The importance of angiogenesis cannot be underestimated because of the unequivocal requirement of tumours to induce neovascularization for cancer progression. The use of *in vitro* and *in vivo* angiogenesis models has helped researchers to understand this process better and has given

rise to the identification of new agents now in clinical trials as antiangiogenic therapies. These models hold promise for the discovery of new antiangiogenic compounds, to understand the molecular mechanisms by which angiogenesis occurs, to identify genes involved in these processes and, ultimately, to decrease the morbidity or mortality of cancer.

REFERENCES

Ades, E. W., *et al*. (1992). HMEC-1: establishment of an immortalized human microvascular endothelial cell line. *Journal of Investigative Dermatology*, **99**, 683–690.

Alon, T., *et al*. (1995). Vascular endothelial growth factor acts as a survival factor for newly formed retinal vessels and has implications for retinopathy of prematurity. *Nature Medicine*, **10**, 1024–8.

Andrade, S. P., *et al*. (1987). Quantitative *in vivo* studies on angiogenesis in a rat sponge model. *British Journal of Experimental Pathology*, **68**, 755–766.

Ausprunk, D. H. and Folkman, J. (1977). Migration and proliferation of endothelial cells in preformed and newly formed blood vessels during tumor angiogenesis. *Microvascular Research*, **14**, 175–203.

Baatout, S. (1997). Endothelial differentiation using Matrigel. *Anticancer Research*, **17**, 451–456.

Bauer, J., *et al*. (1992). *In vitro* model of angiogenesis using a human endothelial-derived permanent cell line: contributions of induced gene expression, G-proteins and integrins. *Journal of Cell Physics*, **153**, 437–449.

Bischoff, J. (1995). Approaches to studying cell adhesion molecules in angiogenesis. *Trends in Cell Biology*, **5**, 69–74.

Cockerill, G. W., *et al*. (1995). Angiogenesis: models and modulators. *International Reviews of Cytology*, **159**, 113–160.

Dellian, M., *et al*. (1996). Quantitation and physiological characterization of angiogenic blood vessels in mice: effect of basic fibroblast growth factor, vascular endothelial cell growth factor/vascular permeability factor, and host microenvironment. *American Journal of Pathology*, **149**, 59–70.

Edgell, C. J., *et al*. (1983). Permanent cell line expressing human factor VIII-related antigen established by hybridization. *Proceedings of the National Academy of Sciences of the USA*, **80**, 3734–3737.

Endrich, B., *et al*. (1980). Technical report – A new chamber technique for microvascular studies in unanesthesized hamsters. *Research in Experimental Medicine (Berlin)*, **177**, 125–134.

Fajardo, L. F., *et al*. (1988). The disc angiogenesis system. *Laboratory of Investigations*, **58**, 718–734.

Folkman, J. (1970). The intestine as an organ culture. In: Burdette, W. J. (ed.), *Cancer of the Colon and Antecedent Epithelium*. 113–127 (Thomas, Springfield, IL).

Folkman, J., *et al*. (1979). Long-term culture of capillary endothelial cells. *Proceedings of the National Academy of Sciences of the USA*, **76**, 5217–5221.

Friedlander, M., *et al*. (1995). Definition of two angiogenic pathways by distinct α_ν intergrins. *Science*, **270**, 1500–1502.

Gimbrone, M., *et al*. (1974). Tumor growth and neovascularization: an experimental model using the rabbit cornea. *Journal of the National Cancer Institute*, **52**, 413–427.

Hanahan, D. and Folkman, J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, **86**, 353–364.

Ingber, D. E. and Folkman, J. (1989). How does extracellular matrix control capillary morphogenesis? *Cell*, **58**, 803–805.

Jain, R. K., *et al*. (1998). Methods for the *in vitro* and *in vivo* quantitation of adhesion between leukocytes and vascular endothelium. In: Morgan, J. R. and Yarmuch, M. L. (eds), *Methods in Molecular Medicine Vol. 18: Tissue Engineering Methods and Protocol* (Humana Press Inc., NJ).

Kohn, E. C. and Liotta, L. A. (1995). Molecular insights into cancer invasion: strategies for prevention and intervention. *Cancer Research*, **55**, 1856–1862.

Kubota, Y., *et al*. (1988). Role of laminin and basement membrane in the morphological differentiation of human endothelial cells into capillary-like structures. *Journal of Cell Biology*, **107**, 1589–1598.

Laterra, J., *et al*. (1990). Astrocytes induce neural microvascular endothelial cells to form capillary-like structures *in vitro*. *Journal of Cell Physics*, **144**, 204–215.

Liotta, L. A., *et al*. (1974). Quantitative relationships of intravascular tumor cells, tumor vessels, and pulmonary metastases following tumor implantation. *Cancer Research*, **34**, 997–1004.

Montesano, R. and Orci, L. (1985). Tumor-promoting phorbol ester induce angiogenesis *in vitro*. *Cell*, **42**, 469–477.

Montesano, R., *et al*. (1987). Phorbol ester induces cultured endothelial cells to invade a fibrin matrix in the presence of fibrinolytic inhibitors. *Journal of Cell Physics*, **132**, 460–466.

Nguyen, M., *et al*. (1994). Quantitation of angiogenesis and antiangiogenesis in the chick embryo chorioallantoic membrane. *Microvascular Research*, **47**, 31–40.

Nicosia, R. F. and Ottinetti, A. (1990). Growth of microvessels in serum-free matrix culture of rat aorta: a quantitative assay of angiogenesis *in vitro*. *Laboratory of Investigations*, **63**, 115–122.

Ozaki, H., *et al*. (1999). Hypoxia inducible factor-1a is increased in ischemic retina: temporal and spatial correlation with vascular endothelial growth factor expression. *Investigative Ophthalmology and Vision Science*, **40**, 182–189.

Passaniti, A., *et al*. (1992). Methods in laboratory investigation: a simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin, and fibroblast growth factor. *Laboratory of Investigations*, **67**, 519–528.

Runkel, S., *et al*. (1991). An intradermal assay for quantification and kinetics studies of tumor angiogenesis in mice. *Radiation Research*, **126**, 237–243.

Sckell, A., *et al*. (1998). Primary tumor size-dependent inhibition of angiogenesis at a secondary site: an intravital microscopic study in mice. *Cancer Research*, **58**, 5866–5869.

Stone, J., *et al*. (1995). Development of retinal vasculature is mediated by hypoxia-induced vascular endothelial growth factor (VEGF) expression by neuroglia. *Journal of Neuroscience*, **15**, 4738–4747.

Takashaki, K., *et al*. (1990). Spontaneous transformation and immortalization of human endothelial cells. *In Vitro Cell Developmental Biology*, **26**, 265–274.

Ziche, M., *et al*. (1997). Nitric oxide synthase lies downstream from vascular endothelial growth factor-induced but not basic fibroblast growth factor-induced angiogenesis. *Journal of Clinical Investigation*, **99**, 2625–2634.

FURTHER READING

Auerbach, R., *et al*. (1991). Assays for angiogenesis: a review. *Pharmacology and Therapeutics*, **51**, 1–11.

Folkman, J. (1992). The role of angiogenesis in tumor growth. *Seminars in Cancer and Biology*, **3**, 65–71.

Jain, R. K., *et al*. (1997). Quantitative angiogenesis assays: progress and problems. *Nature Medicine*, **3**, 1203–1208.

Madri, J. A. and Pratt, B. M. (1986). Endothelial cell–matrix interactions: *in vitro* models of angiogenesis. *Journal of Histochemistry and Cytochemistry*, **34**, 85–91.

Chapter 63

Models for Tumour Cell Adhesion and Invasion

Jörg Haier and Garth L. Nicolson
The Institute for Molecular Medicine, Huntington Beach, CA, USA

CONTENTS

- Biological Basics
- Principles of Model Establishment
- General Applications
- Interpretation of Experimental Results
- Biological and Technical Limitations
- Perspectives

BIOLOGICAL BASICS

The prevention of death in most cancer patients is dependent upon understanding the mechanisms of cancer cell spread from the primary tumour to distant organ sites. Although the pathogenesis of metastasis is the subject of numerous studies in basic and clinical research, the complex mechanisms that make a tumour cell metastatic are not well understood. The development of metastases is a non-random process, where the implantation, invasion, survival and growth of a single cell or small numbers of cells can lead to morbidity and death (Nicolson, 1995).

The sequential model for the development of metastases involves tumour growth, neovascularization and invasion at the primary sites, followed by penetration into lymphatics and blood vessels or into body cavities. Malignant cells that loose cellular contacts with surrounding cells or extracellular matrix (ECM) can undergo phenotypic change to greater motility. Subsequently, malignant cells can be released from the primary tumour and be distributed throughout the body via the blood or lymphatic circulation. To survive the circulating tumour cells must adhere to the vessel walls of distant host organs, and eventually penetrate the wall to avoid blood shear forces and host defence mechanisms. Finally, the metastasizing cells have to survive and grow (Nicolson, 1988a,b) (**Figure 1** and see the chapters *Models for Tumour Cell–Endothelial Cell Interactions* and *Invasion and Metastasis*).

The localization of tumour metastases in distant organs is not determined solely by anatomical considerations and blood flow. During haematogenous metastasis tumour cells often show selective colonization of various organs independent of the percentage of microcirculation through the organ or whether the organ is first to be encountered by the blood-borne tumour cell (Nicolson, 1989). This multistep process is thought to be determined by the unique properties of the tumour cells and the host organs (Doerr *et al*., 1989; Fenyves *et al*., 1993). Adhesion of metastatic cells to and migration through the microvascular vessel walls of host organs are very important steps in colonization of distant organs (Nicolson, 1988a,b), and during these steps specific interactions between the tumour cells and endothelial cells, the ECM or surrounding cells are required. During adhesion cells can form temporary attachments to certain structures, e.g. during chemotactic cell movement, or they can establish more long-term adhesive interactions, resulting in the formation of structures such as desmosomes, adherens junctions, gap and tight junctions. Various tumour cell surface molecules mediate these interactions, either by direct interactions, such as integrins or other adhesion molecules, or as receptors for soluble peptides or hormones that modify cellular signalling and indirectly regulate cell adhesion. Usually the same molecules are also found on normal cells where they function in the maintenance of tissue structure and normal cellular regeneration. Maintenance of the proper expression of adhesion molecules is important in tissues (e.g. epithelial layers) for normal tissue morphology and function. Adhesive cell properties play an important role in determining the organisation of tissues, and they allow directed cell migrations found during embryonic development, inflammation, immune response and wound repair. Additionally, they participate in regulation of gene expression, cell growth and differentiation and programmed cell death (Virtanen *et al*., 1990).

Adhesion molecules can also be transmembrane proteins that can anchor cytoskeletal proteins on the cytoplasmic side of the cell membrane, such as actin and actin-binding

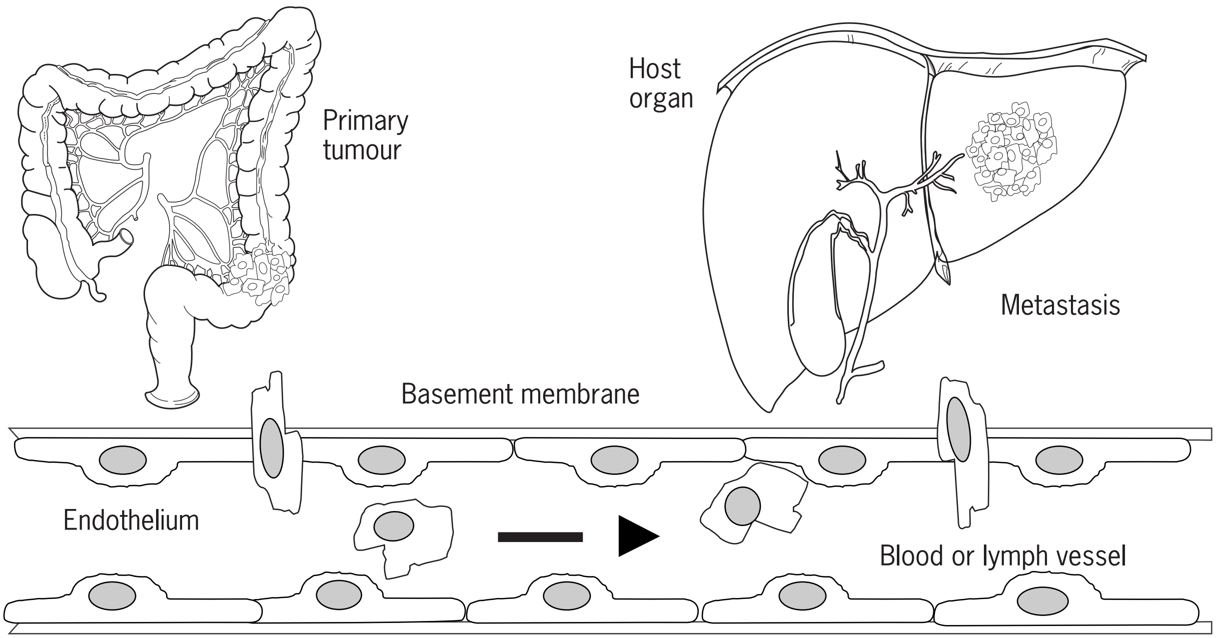


Figure 1 Sequence of events in the development of a distant cancer metastasis.

proteins, while also anchoring extracellular structures on the outer surface of the cell membrane. In addition to physical linkages between the extracellular environment (surrounding cells and ECM) and the cytoskeleton, adhesive complexes participate in important signal transduction systems as modulators or receptors. Their functions in cell signalling are probably at least as important as their cytoskeletal and cell attachment properties. An important feature of the intracellular protein complexes induced by the binding of adhesion molecules is the subsequent induction of various signal transduction molecules that can trigger further downstream signals into the cell. These signalling pathways can generate or influence gene expression, secretion of enzymes or cytokines and conformational changes of the cytoskeleton, among others. Understanding these regulatory functions appears to be of importance in determining how a transformed cell becomes invasive and metastatic. (See also chapters *Wnt Signal Transduction* and *Extracellular Matrix: the Networking Solution*.)

During neoplastic transformation, epithelial and other cells can undergo changes in differentiation status and can change their surface structures by differential expression of molecules at the cell surface or post-translational modifications, such as glycosylation. In addition, distinct organ basement membrane compositions have been found in many malignant tumours (Nicolson, 1995). Moreover, tumours, such as colon carcinomas, produce their own ECM components that can change peritumoral stroma organisation (Burtin *et al*., 1983), including the release of unusual ECM components. For example, oncofetal fibronectin can be produced by tumours and has been correlated with poor prognosis (Inufusa *et al*., 1995). Alternatively, some tumour cells fail to deposit ECM or express certain adhesion components, and subsequently their adhesive interactions are reduced compared to normal cells. Such structural alterations can involve various adhesive systems, resulting in modified or reduced cellular interactions with components in the cellular environment. For example, the loss of cell–cell and cell–ECM interactions promotes motility of tumour cells from the primary site. On the other hand, the establishment of new adhesive interactions is necessary for cell attachment to the endothelium and basement membrane in distant host organs during haematogenous metastasis formation.

Homotypic adhesion between tumour cells at the primary site, their heterotypic interactions with leucocytes and platelets during transport in the circulation and cellular binding with the endothelium and ECM in distant organs involve different cellular adhesion systems, including a broad spectrum of cell surface molecules representing different adhesion families (Nicolson, 1991). These tumour cell molecules can be classified into several main groups: (1) integrins, (2) selectins, (3) immunoglobulin-like, (4) cadherin molecules, and others, such as proteoglycans and glycoconjugates (lectin receptors) that are involved in adhesion processes. Microvascular endothelial cells (EC) and their underlying ECM are the main target structures of these adhesion molecules, and there appears to be some organ specificity in the adhesive interactions that occur (Rice *et al*., 1988; Belloni *et al*., 1992).

The detachment of tumour cells from the primary site is characterized by the loss of cell–cell adhesion, and in epithelial cells the dysfunction of the cadherin–catenin systems appears to play an essential role in this process. In poorly differentiated carcinomas, loss of epithelial cell

contacts is frequently observed, allowing the cells to break away from the primary tumour (Shiozaki *et al*., 1996). Cadherins are the major components of Ca^{2+}-dependent epithelial cell adhesion systems and are present in all epithelial cells, where they bind cells together through homophilic interactions. Members of the various subgroups of cadherins are expressed only in certain tissues; however, cells normally express several cadherin types. Catenins are the intracellular associate molecules of cadherins that provide the links to cytoskeletal components and signal transduction. Dysfunction of cadherin-mediated adhesion can be caused by loss or changes in function of any component of the complex. (See also chapters *Signalling by TGF beta* and *Wnt Signal Transduction*.)

Penetration from the primary tumour mass into the lymphatics or blood circulation requires the secretion and/or activation of degradative enzymes, some classes of which appear to be metalloproteases (MMP), cathepsins, plasminogen-activators and endoglycosidases, such as heparanase (Monsky and Chen, 1993). Under normal physiological conditions the balance between protein synthesis and degradation is well regulated by various activators and inhibitors. These control mechanisms ensure the controlled destruction and replacement of ECM at points where it is required for cell migration, immune response or wound repair (Shapiro, 1998). Degradative enzymes are usually secreted from most cell types into their extracellular environment as inactive proenzyme forms. There they can be bound to specific inhibitors, or they can be proteolytically activated by other proteases. The regulation of ECM or basement membrane degradation is organized into various cross-linked cascades to guarantee restricted activity at specific sites. Tumour cells themselves can produce various degradative enzymes, or they can induce or modify enzyme secretion by EC or stromal cells, such as fibroblasts, mast cells, and macrophages, through release of paracrine factors (Lengyel *et al*., 1995; Polette *et al*., 1997). In many tumours the expression of degradative enzymes and/or expression of their inhibitors were demonstrated to cause an imbalance of degradative activity, especially at the invasion front of malignant tumours (Liotta and Stetler-Stevenson, 1991; Conese and Blasi, 1995). This increased degradation induced by tumour cells results in higher motility and invasiveness and permits remodelling of the cellular environment within and around invasive tumours. (See also chapter *Modelling Tumour–Tissue Interactions*.)

Once tumour cells have reached the circulation or body cavities, they can be transported to near and distant organs where they must bind to host structures to establish distant metastases. The binding of circulating cells to EC and subsequently ECM occurs in discrete steps: initial cell contact, followed by adhesion stabilization and subsequent signal transduction, and finally degradation and cell migration (Nicolson, 1989; Pauli *et al*., 1990; Dorudi and Hart, 1993). Following the primary contact of malignant cells, microvascular EC are stimulated to retract with the subsequent exposure of basement membrane or subendothelial ECM (see the chapter *Models for Tumour Cell–Endothelial Cell Interactions*). Thus many tumour cells reach a relatively rigid ECM structure in blood vessel walls (Nicolson, 1995). In general, the subendothelial ECM has been found to be a better substrate for cell adhesion of tumour cells than the EC surface (Kramer *et al*., 1980; Nicolson *et al*., 1981). Additionally, the adhesive behaviour of tumour cells can be altered by the organ-specific composition of the subendothelial ECM. Moreover, in some metastatic systems differential adhesion of tumour cells to specific organ subendothelial ECM has been seen (Lichtner *et al*., 1989).

Various studies have demonstrated a correlation between the organ preference of metastasis and the *in vitro* adhesion rates of malignant cells to specific organ-derived EC (Tressler *et al*., 1989) or subendothelial ECM (Lichtner *et al*., 1989). For example, contacts between carcinoma cells and the microvascular endothelium seem to be related to the expression of selectins, sialyl-Lewis^x (sLe^x) and other carbohydrate structures, intercellular adhesion molecules (ICAM) and possibly annexins (Menger and Vollmer, 1996). However, integrin-mediated interactions with the subendothelial ECM could be among the most important determinants for organ-specific metastasis. Reflecting on this cell-specific behaviour, distinct adhesive properties of tumour cells with different metastatic potential to EC/ECM were found. In support of this in different cell systems it was demonstrated that adhesion of highly metastatic cell lines to organ EC or ECM occurred at higher rates than seen with comparable poorly metastatic cells (Tressler *et al*., 1989; Haier *et al*., 1999c).

Integrins mediate adhesion to various ECM components. Binding to their receptors and clustering of integrins at the cell surface appear to be the trigger that initiates signals that can induce interactions with cytoskeleton components. Integrins important in this process are the β_1- and β_3-integrins, which are linked to cytoskeletal proteins, such as actin, α-actinin and talin. These components are further connected to various other cytoskeleton proteins such as vinculin and zyxin. This signalling process is also linked to phosphorylation and dephosphorylation events, mainly on tyrosine residues, and this, in turn, may activate the intracellular signalling cascades (Akiyama *et al*., 1994). Different complexes may be involved in downstream signal transduction after the establishment of adhesive interactions, such as focal adhesion kinase (FAK)–paxillin or Grb2–SOS complexes (**Figure 2**). (See also chapter *Extracellular Matrix: the Networking Solution*.)

During the interactions between tumour cells and EC/ECM components, adhesion molecules function both for the primary contact, definitive adhesion stabilization and cell migration. Tumour cells adhere at the points of EC/ECM contacts by sequestering or clustering receptors at zones of adhesive interactions (e.g. focal adhesions). At this time adhesion has not yet been stabilized, and increasing the shear force under flow conditions breaks the adhesive

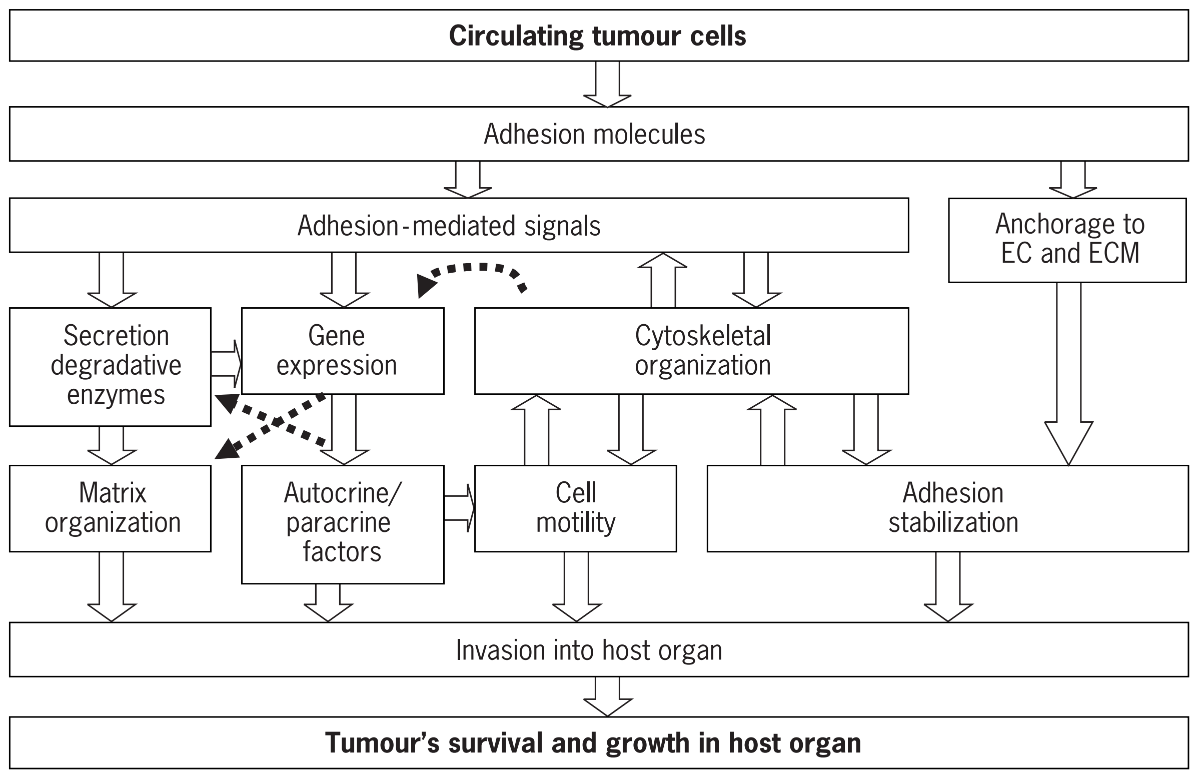


Figure 2 Role of adhesion molecules in metastasis formation at distant organ sites.

contacts and releases the tumour cells from the EC/ECM. However, if adhesion can be stabilized, tumour cells can resist such hydrodynamic shear forces (Nicolson, 1988a). The stabilization process appears to require receptor clustering and possible signalling before actual receptor cross-linking or changes in receptor affinity occur mediated by membrane-associated enzymes or signalling cascades. Following these initial adhesive interactions, the cells have to actively change their shape by spreading or flattening to increase contact areas and initiate cell migration through the vessel wall (Menter *et al.*, 1992, 1995). These processes might be accompanied and regulated by various transmembrane signal transduction events (Haier *et al.*, 1999a,b).

In contrast to static cell adhesion, biophysical factors influence adhesion processes under flow conditions (**Figure 3**). First, under dynamic conditions flow velocity determines the available time for adhesion. Since initial areas of cell contacts to the vessel wall are very small because of the more or less spherical shape of flowing tumour cells, the initial cell contacts may occur via cell protrusions or filopodia. Possible alterations in cell shape and the morphology of participating cells are important in determining the total area of cell–cell contact. Second, cell flattening enables adherent cells to avoid high shear forces within the centre of the parabolic curve of fluid flow. These factors and the character of the flow (laminar under physiological conditions of microcirculation) are responsible for the strength of shear forces acting against adhesion. Flow rate and viscosity are the main factors determining strength of cell shear stress. In addition to these biophysical factors on moving cells shear stress alone induces functional reactions of EC (Lawrence *et al.*, 1987) and possibly other cells. These cells release interleukins IL-1 and IL-6 depending on flow rates and/or changes in the phosphorylation status of proteins related to adhesion complexes.

During adhesion under flow conditions, tumour cells have two different patterns of reaction (Giavazzi *et al.*, 1993). Colon, ovary and breast cancer cells showed modest transient interactions with resting human umbilical cord vein endothelial cells (HUVEC) comparable to the behaviour of leucocytes, such as rolling on IL-1- or tumour necrosis factor-activated HUVEC. In contrast, melanoma and osteosarcoma cell lines bound to EC and demonstrated no rolling. Furthermore, differences in adhesive properties for different ECM components were noted. Initial tumour cell–ECM interactions under flow conditions appear to depend mainly on biophysical or mechanical factors, whereas specific interactions are required for further adhesion stabilization. For example, receptor cross-linking by transglutaminase was found to be an important determinant of adhesion stabilization in several cell systems (Menter *et al.*, 1992). Moreover, various parameters that describe cell adhesion stabilization under flow conditions correlated with the metastatic potential using various ECM components (Smith *et al.*, 1996).

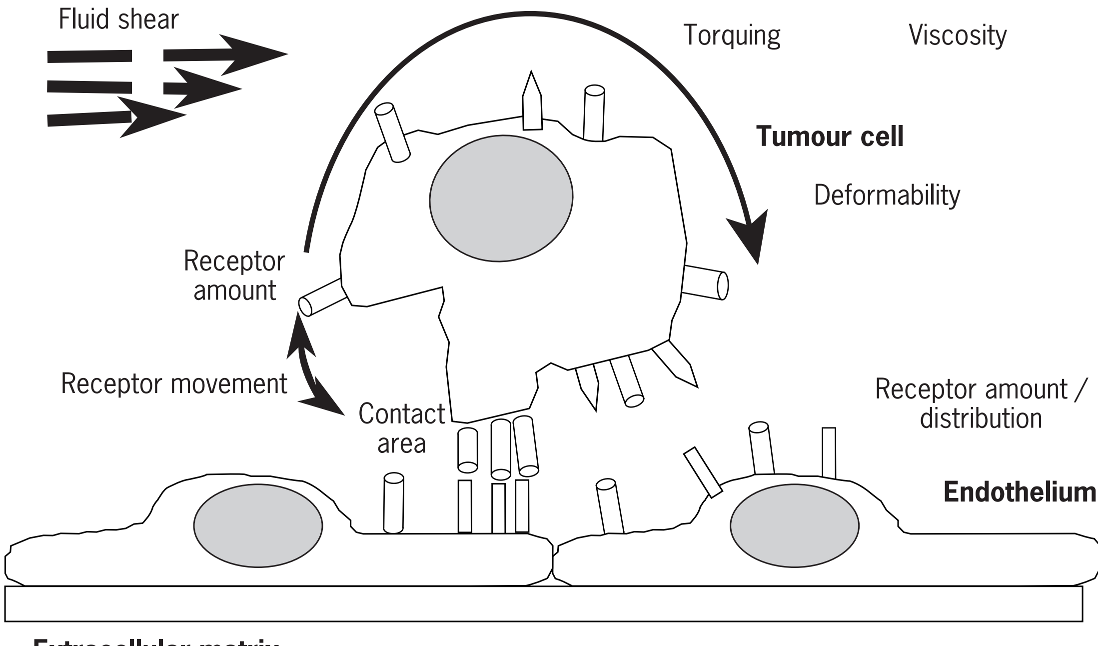


Figure 3 Factors influencing tumour cell adhesion under dynamic conditions of fluid flow. The numbers of receptors in a certain contact area determines the maximum possible number of adhesive interactions. Receptor movement and cellular deformability can affect adhesion stabilization. Fluid shear forces occur as parabolic phenomenon in the microcirculation, and flow velocity, fluid viscosity and resulting cell torque are counteracting forces that prevent cell adhesion.

Parallel to morphological alterations, tumour and host cells undergo functional responses to specific adhesive interactions. For example, they can release hydrolases and proteases, destroying the integrity of the vascular basement membrane (Marchetti *et al*., 1993). As described above, the composition of subendothelial basement membranes is organ-specific, where the main components are type IV collagen, laminin, FN and vitronectin. These matrix proteins are targeted for degradation mainly by different MMPs that have relative substrate specificity. The tumour cells or EC of the host organ can secrete MMPs, and these degradative enzymes are then activated by other membrane-bound proteases, or by inactivation of their inhibitors. Together with the induction of directed cell motility, the porous subendothelial membrane then allows tumour cells to penetrate the basement membrane by active locomotion, thereby leaving the circulation and invading the host organ.

Cancer cell spreading and motility are stimulated by specific chemical stimuli. Cellular motility consists of directed morphological alterations (such as ruffling and pseudopodia formation), establishment of new adhesive interactions at the moving cell front and controlled detachment at the rear of the cell. Complex cellular processes are required for the response to soluble motility factors (chemotaxis) or insoluble motility factors (haptotaxis) (Nicolson, 1989). Alternatively, tumour cells may produce their own autocrine motility factors (Menter *et al*., 1995). Cells that have the capacity to migrate actively towards a soluble chemoattractant must detect the direction of a chemical gradient through recognition of receptor occupancy and respond by orienting toward the gradient and initiating cytoskeletal activation and directed movement. The ability of tumour cells to respond to organ-specific or organ-associated chemoattractants may be an important determinant in the selective migration of malignant cells into specific tissues (Nicolson, 1988a).

A likely source of organ-associated tumour cell motility factors is organ microvascular EC. The EC-produced motility factors might interact with other soluble tissue factors to promote organ-specific tumour cell invasion (see the chapter *Models for Tumour Cell–Endothelial Cell Interactions*). Chemotactic factors are also present in ECM, and they can stimulate directed migration of malignant cells (Cerra and Nathanson, 1989). Indeed, tumour cell chemotaxis appears to be specifically influenced by the composition of the subendothelial basement membrane in the target organ (Aznavoorian *et al*., 1990).

Tumour cells can also synthesize and release their own motility and growth factors. The autocrine motility factor is an important example of such factors made by metastatic cells. This factor is secreted by a wide variety of tumour cells, and it stimulates chemotaxis in various cell systems (Liotta *et al*., 1986; Atnip *et al*., 1987; Siletti *et al*., 1991;

Watanabe *et al*., 1991). One of the interesting tumour cell properties associated with organ-specific metastasis is that tumour cells may respond to local concentrations of growth factors produced by or present in the target organs. In support of this it was shown for various tumour systems that tumour cells could preferentially grow in medium conditioned by the target organs for metastatic colonization. Although the production and secretion of autocrine growth factors was not found to be dependent on the metastatic potential, the general results indicate loss of paracrine growth requirements with tumour progression (Rodeck and Herlyn, 1991), and this often appears to be accompanied by the secretion of autocrine motility and growth factors. Both tendencies together can reduce the dependence of tumour cells from their host environment.

The establishment of suitable tumour cell models remains a challenging task. *In vitro* and *in vivo* models each have their advantages, limitations and specific conditions. For example, results using cells from different cell origins cannot be simply transferred to other cells, and findings obtained with one model must be confirmed for other systems.

PRINCIPLES OF MODEL ESTABLISHMENT

Cell Detachment

Adhesion molecules can be subdivided into two groups with regard to their requirements for calcium (Birchmeier *et al*., 1993). This characteristic allows a simple distinction between major groups of adhesion molecules by removal of calcium from the medium. If the involved adhesion system is calcium dependent, cells detach rapidly after trypsin treatment or they fail to undergo adhesion to a normally good substrate. After removal of calcium (e.g. by EDTA) cultured cells usually dissociate within a few minutes, round up, and gaps between the cells can be found. However, treatment with chelating agents and mechanical forces alone are not sufficient in many cell systems to obtain complete loss of cell–cell interactions. The remaining tight junctions must be destroyed by trypsin treatment. These simple procedures only provide rough information about adhesion systems and do not yield quantitative results or analysis of complex interactions.

A more specific but less sensitive way to evaluate the involvement of specific molecules, such as cadherins, is inhibition of their function by blocking antibodies against their extracellular domains. In antibody-treated cell monolayers morphological changes can be observed, and intercellular gaps become visible after a few hours. In three-dimensional cultures the morphological changes appears as the formation of spherical cell shape and loss of compact cell mass. Various microscopic techniques can be applied for the investigation of the functional consequences of cadherin receptor interference. Frequently, fluorescence or confocal microscopy is used for visualization of cytoskeletal behaviour or molecular changes in cadherin–catenin complexes along with cell morphological changes.

Molecular cloning of truncated or modified cadherins has been developed for specific studies on the role of extra- or intracellular domains of these molecules in adhesion events. cDNA constructs in dominant negative cells with truncated cadherin extracellular domains but an intact intracellular domain or the inverse have been used (**Figure 4**). These two types of molecules are functionally incomplete but compete with normal cadherin function. In these cells the remaining intercellular contacts or cytoskeletal anchorage are severed. However, technical difficulties, and the fact that cells normally express multiple forms of cadherins, limit the application of this method to cells that only express one type of cadherins. In contrast to this technique, mutations of specific protein motifs have been successfully used to modify the functional properties of cadherin–catenin complexes and their related signal transduction pathways with high specificity (**Figure 4**).

Cell Adhesion

Cell Adhesion Assays

The general purpose of cell adhesion assays is to obtain quantitative information on tumour cell–cell and cell–substrate interactions. Their basic principle is to measure reattachment of single cells that have been removed from tissue culture and resuspended as a single cell solution. These *in vitro* assays use labelled tumour cells (fluorescence, radioisotope, or staining) that are allowed to adhere to cell monolayers or substrate-coated plastic surfaces in microtitre or other plates. After a defined period of time, unattached cells are removed by washing, and the remaining cells are enumerated and compared to the total number of cells added to the cell suspension (**Figure 5**). This procedure can be used to evaluate adhesion to certain substrates, such as ECM components, and to identify mediating receptor molecules on the tumour cell surface by inhibition of adhesion with specific antibodies or mimicking synthetic peptides which specifically block the adhesive functions of target molecules (Tressler *et al*., 1989; Haier *et al*., 1999a,b).

Cell attachment to specific substrates is a dynamic process that results in functional cell responses and changes in cell morphology. Within a certain time period, a percentage of attached cells can be seen without specific adhesive interactions, whereas other cells reach stable adhesion states or even begin to spread. However, adhesion–time curves show sigmoid patterns where a plateau value can be reached (see **Figure 10**). Therefore, the adhesion assays have to consider this time-dependent behaviour. Statistical analysis of a large number of cells is necessary to allow the characterization of adhesive properties. Microtitre plate assays do not measure the simple

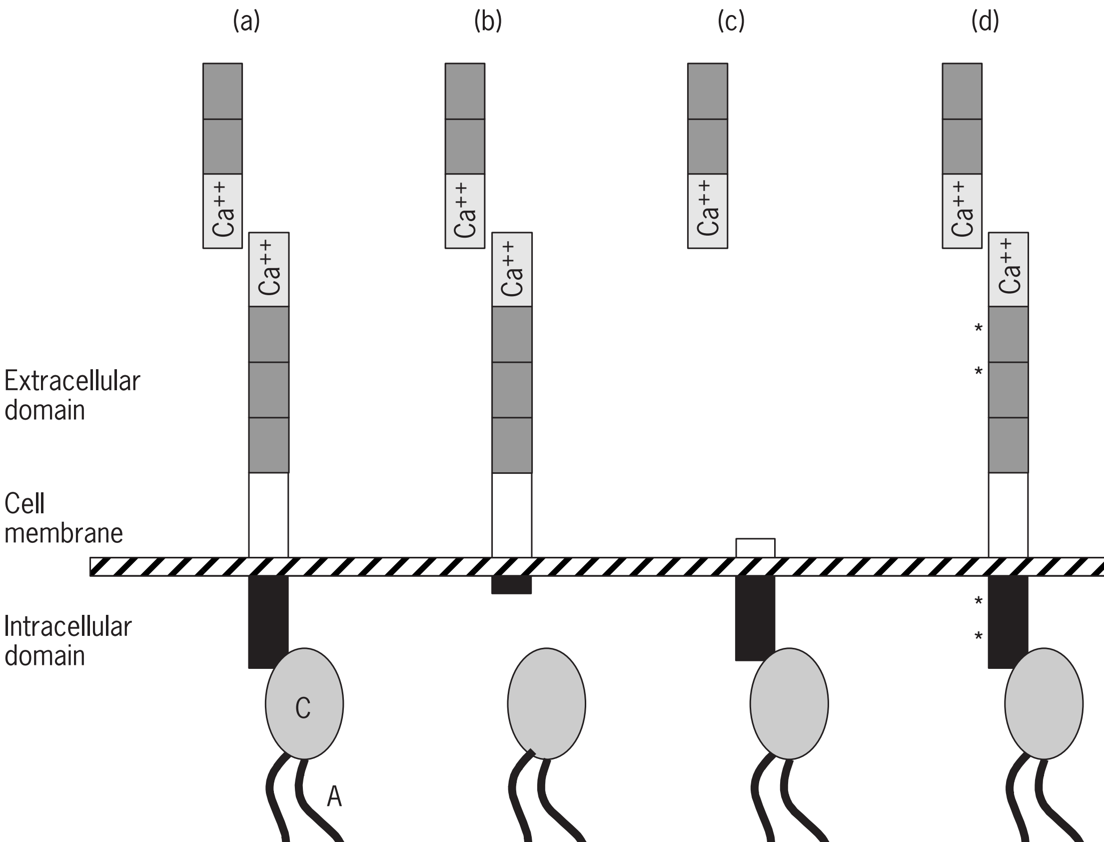


Figure 4 Truncation of cadherin domains for functional investigation of the role of various domains of the molecule. (a) Intact molecule with normal homotypic receptor interaction and cytoskeletal anchorage; (b) truncated intracellular domain with loss of cytoskeletal interactions; (c) truncated extracellular domain with altered adhesive interactions. (d) mutations (*) in different domains resulting in loss of domain functions. C, catenins; A, actin.

interactions between adhesion molecules and their ligands, but a more complex cell response to adhesion receptor binding. They are unable to discriminate between primary cell contact, adhesion stabilization and cell spreading. One of their general limitations is that the washing procedures for removal of unattached cells is essential but difficult to apply reproducibly, and weak specific interactions between adhesion molecules and their substrates may not resist these forces. Additional technical and biological limitations are discussed below.

For further analysis of involved cellular structures and signalling pathways, the tumour cells can be pharmacologically pretreated with specific activators or inhibitors. This technique can be employed, for example, for the investigation of participating cytoskeletal proteins, whose native structures or ability to aggregate are modified. Other frequently used types of pretreatment can evaluate the role of various kinases or phosphatases in the regulation of adhesive properties. Finally, genetically manipulated cells can be used where specific signalling molecules are specifically altered (e.g. modification of binding sites or exchange of amino acids for prevention of phosphorylation). In principle, cell adhesion of untreated and treated cells must be compared, and the relative differences in the number of adherent cells can reflect involvement of target structures or functions in cell adhesion or its regulation.

General cell adhesion techniques have been modified in numerous studies for the investigation of regulation or function of adhesion molecules. Instead of microtitre plates, coated microbeads have been introduced, for example, to study the involvement of cytoskeletal proteins. The use of agarose beads allows the application of confocal, laser scanning or other microscopic techniques to visualize intracellular processes. Using these methods, a hierarchical involvement of actin-binding proteins was found in areas of focal contacts after receptor binding of integrins (Miyamoto *et al*., 1995). Another technique that is frequently employed for studies on receptor clustering is binding and artificial ligation of monoclonal antibodies or the use of chemically cross-linked antibodies with high specificity to the targetted adhesion molecules (**Figure 6**).

Flow or Dynamic Cell Adhesion

Static cell adhesion assays are widely used for characterization of the adhesive properties of various normal and

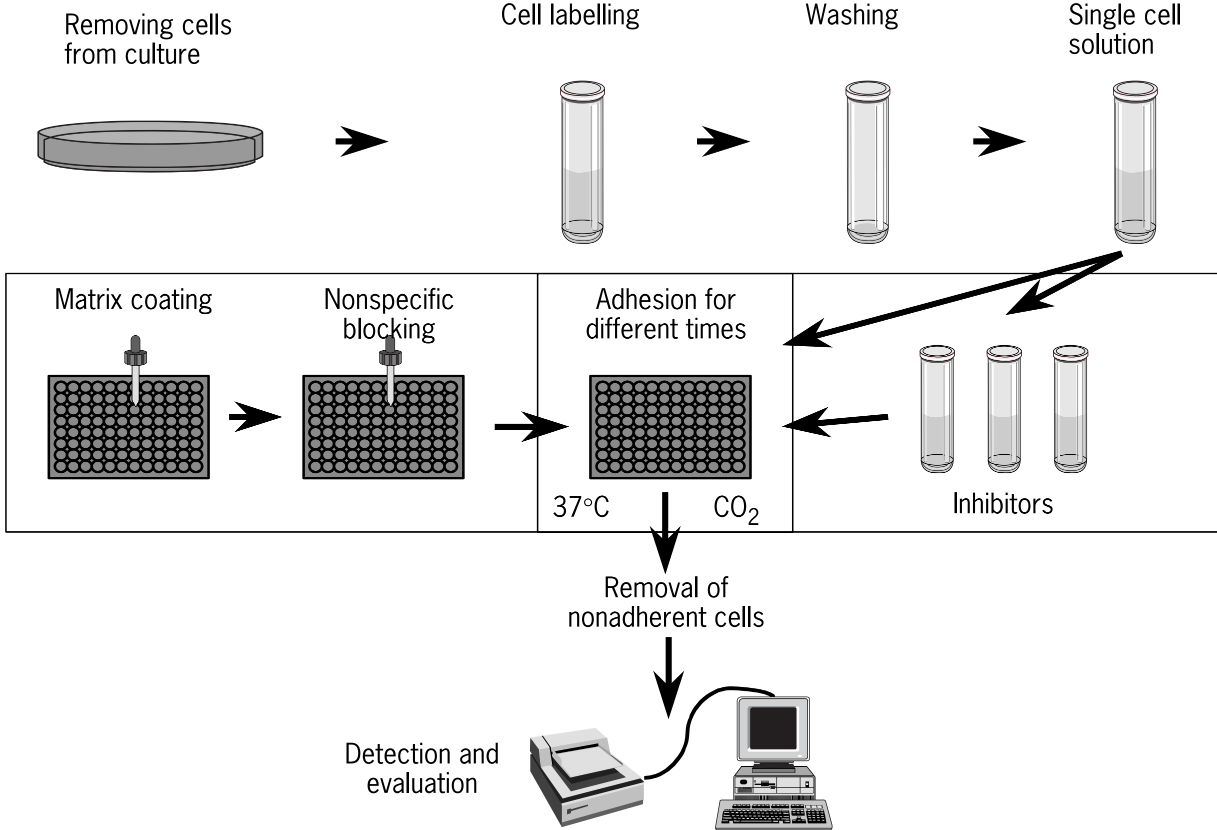


Figure 5 Principle of cell adhesion assay in microtitre plates.

tumour cell lines. Unfortunately, these assays do not consider hydrodynamic forces, such as wall shear stress or the parabolic form of laminar flow that occurs within the microvasculature. In addition, these assays measure cell adhesion as one event, instead of complex initial cell–surface interactions, adhesion stabilization and cell spreading. Hydrodynamic adhesion assays are capable of mimicking haemodynamic conditions in the microcirculation. Using specific parallel plate laminar flow chambers the two phases of adhesive interactions between circulating cells and vascular surfaces (EC or ECM) can be studied separately in haemodynamic adhesion assays (Yun *et al*., 1997; Patton *et al*., 1993). For example, Lawrence *et al*. introduced a parallel plate flow chamber to study adhesion of neutrophils flowing over activated EC (Lawrence *et al*., 1987, 1990) (**Figure 7**). They showed that this system allows controlled studies of the flow effects on cell suspensions. The conditions of hydrodynamic adhesion in these reports revealed that both biophysical and biochemical interactions were involved in cell adhesion.

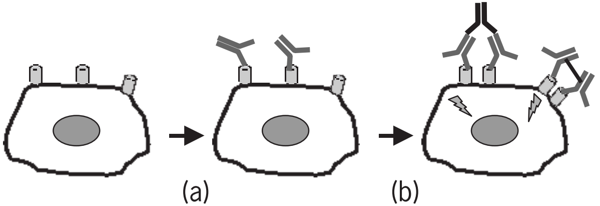


Figure 6 Artificial clustering of adhesion molecules using antibody ligation. (a) Tumour cells are treated with primary antibodies against a specific adhesion molecule which leads to receptor occupancy. (b) Receptor ligation is caused by anti-antibody antibodies that induce receptor-mediated signalling, or chemically cross-linked antibodies are used.

To study initial tumour cell arrest and adhesion stabilization under flow conditions, different parameters can be defined using various flow rates. These parameters can measure (1) the initial arrest of tumour cells to EC/ECM and (2) adhesion stabilization. A major advantage of this system is the reproducible control of flow and shear forces. Each of these steps can be investigated with regard to cell structural and regulatory requirements.

Using this system, we found that adhesive interactions of HT-29 colon carcinoma cells with ECM-coated surfaces under laminar flow conditions occurred in discrete steps: (1) cell rolling, (2) cell sticking or initial adhesion and (3) stabilization of cell adhesion. Under shear flow rolling of tumour cells on ECM surfaces appeared to be mediated mainly by physical/mechanical and nonspecific surface–cell membrane interactions. Differences in shear forces between the centre and margins of the flow curve appear to be responsible for the reduction of tumour cell velocity by

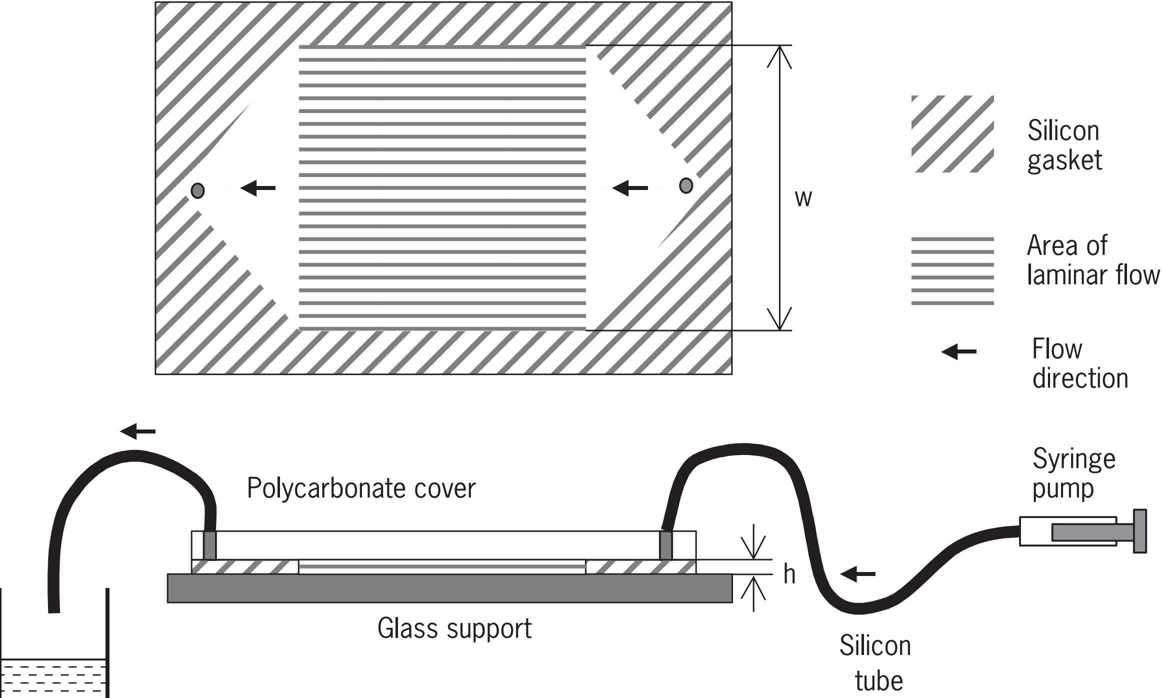


Figure 7 Principle of laminar flow parallel plate chamber. A polycarbonate shear deck supports surface-coated glass-slide to obtain a small height-to-width ratio that is required for maintenance of laminar flow. In this system, wall shear stress can be mathematically calculated as for Newtonian fluid flow. Evaluation is performed under an upright microscope (1:63.5 magnification).

rolling on the ECM surface (**Figure 8**). Once tumour cells move below a critical threshold of flow-related shear forces, specific interactions between tumour cell surface molecules and ECM molecules can be initiated. During various steps of dynamic adhesion the importance of nonspecific interactions decreases compared to static adhesion, and specific integrin-mediated binding occurs as a prerequisite for adhesion stabilization.

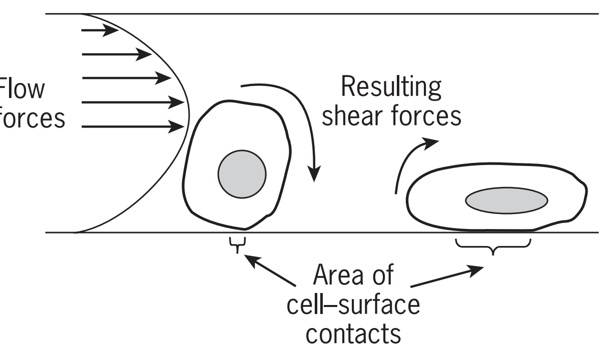


Figure 8 Influence of various parameters on cell-surface interactions. Flow velocity occurs parabolic in a laminar plate flow chamber with maximum flow in the centre of the fluid flow. Resulting shear forces are determined by the differences of fluid flow on the centrifugal and centripetal side of the tumour cell. Flattening of cells results in an increased area of cell–surface contacts and reduction of effective shear forces acting on the cell.

Spreading

Cell spreading is usually the first morphological change seen in cells after adhesion to certain substrates. It is characterized by flattening of cells that are more spherical in single cell suspensions. Spreading of cells requires specific interactions between surface molecules and substrate to initiate cytoskeletal alterations, which leads to enlargement of contact area and loss of round shape. In contrast to initial cell adhesion, active cellular functions are necessary for these shifts in cell morphology, whereas initial cell adhesion includes, in part, nonspecific surface–substrate interactions (see above).

Assays for determination of cell spreading are similar to the cell adhesion assays described above. However, after certain times unattached cells are not removed, and the cells have to be fixed for microscopic analysis. Using an inverted phase-contrast microscope, the relative percentage of bound cells that show the criteria of cell spreading can be evaluated. Specific morphological criteria for spreading are loss of sharp cell borders, development of pseudopodia, alteration in diameters and nucleus–cytoplasm relations. Often these parameters are analysed by digital image software that allows exact measurement

and comparison with average values in the given cell population.

Migration and Motility

The ability of movement is one of the essential features of tumour cell invasiveness. Most eukaryotic cells (including tumour cells) can actively migrate along a concentration gradient of chemoattractants. Three separate functions are required for this response to the cellular environment. First, cells have to recognize the gradient; second, they must establish and break adhesive interactions to their underlying extracellular substrate; and finally, cytoskeletal response is necessary for the movement and resistance to traction forces.

The basic principle of these experiments is very simple. Cells are allowed to adhere to certain substrates, and their movement along a chemical gradient is observed by direct visualization (e.g. time-lapse video-microscopy) or by counting of cells that passed a defined distance after a certain period of time (e.g. filter assays). Cell migration is detected using radioactive, fluorescent or enzymatic cell markers, or cytochemistry. The classical motility assay was described by Boyden and measured migration of cells through a porous membrane. A modified form of Boyden chamber assays is useful for invasion analysis (see below). The assays can identify new chemoattractants and investigate cellular processes required for cell motility. Selection of suitable conditions for the assays is important because cell motility can be found at very different rates and time dependence in response to chemotactic signals.

Various factors influence the experimental design of motility assays. (1) A stable gradient has to be maintained over the observed time period. This requires technical modifications if the cells to be investigated move slowly, which is a characteristic of many types of tumour cells (compared with fast-moving neutrophils). Tumour cells typically move only $0.2–2\,\mu m\,min^{-1}$, possibly because of slow detachment rates from ECM-coated surfaces. (2) Depending on the type of investigation, such as intracellular signalling or identification of chemotactic substances, the assay conditions have to be adapted to the experiment. For example, chemotactic gradients can be applied gradually or suddenly or they can be step or continuous. (3) Chemotactic cell responses are highly concentration dependent, and increasing concentrations of chemoattractants normally results in increased cell motility. However, very high concentrations of chemotactic substances may actually inhibit cellular movement and there is often a concentration curve where maximal response occurs.

Chemotaxis assays are designed to determine if a biological compound can induce directed motility of target cells. For verification of chemotactic responses, a gradient of the factor must stimulate directed locomotion rather than enhance chemokinetic cell movement or movement in random directions (see also Biological and Technical Limitations). Since differences in the motility and growth responsiveness of malignant cells to soluble factors derived from mouse liver, brain and lung microvessel EC have been found, conditioned media can be used to examine the effects of secreted factors on various tumour cell systems. Chemotactic responses of various tumour cells to these factors have been characterized using filter assays. Identified chemoattractants for these cells include, for example, autocrine and paracrine motility factors and growth factors, such as fibroblast growth factor, epidermal growth factor, insulin, platelet-derived growth factor and hepatocyte growth factor/scatter factor.

Invasion

In vitro evaluation of tumour cell invasiveness includes the determination of three combined factors that lead to movement of cells through a ECM layer: (1) cell adhesion, (2) destruction of ECM integrity and (3) chemotactic motility. Models for tumour cell invasion should simulate the extracellular environment or microvessel walls in host organs. Tumour cells that are invasive or metastatic can be studied with regard to the mechanisms that are responsible for these characteristic features of malignant cells. The major barriers for tumour cell invasion into underlying stroma are basement membranes in the primary organ or in blood vessels at secondary organ sites. These basement membranes consist of thin complex layers of various ECM components that can be imitated in invasion experiments.

The main difference between invasion assays and measurement of motility is that invasion assays are performed using thin gel layers of ECM components (as single components, such as collagen gels, or in combination, such as Matrigel) overlaying the invasion filters, whereas motility assays just use coated filters. To invade through an ECM gel layer, tumour cells have to release degradative enzymes. Commonly, in these modified Boyden chamber assays two chambers are separated by a microporous filter (usually polycarbonate). The pore size is variable ($1–8\,\mu m$) depending on the diameter and deformability of the cells to be used. For invasion experiments, a thin layer of ECM gel is put on the upper side of the filter, and cells are added to the upper chamber. Invasion is measured as the numbers of cells that reach the lower surface of the filter during a certain incubation period (**Figure 9**). For further analysis, migrated cells can be carefully removed from the lower surface, for example to isolate highly invasive subpopulations of tumour cells.

Other *in vitro* experimental designs use other types of gel preparations and observe cell migration within a gel using microscopic observation. However, the filter assay described yields better results, especially with Matrigel (or single ECM component preparations), and these artificial conditions are useful for comparisons with *in vivo* characteristics of tumour cell–ECM systems. A different approach to study tumour cell invasion is the use of

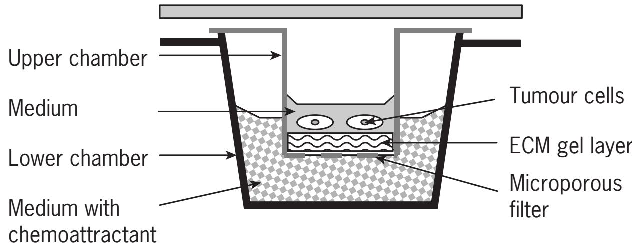


Figure 9 Principle of Boyden chamber assays for cell motility and invasion. Two chambers are separated by a microporous filter membrane. A chemotactic concentration gradient acts as stimulus for directed cell motility or invasion. For motility assays, the upper side of the microporous filter is coated with ECM substrates or EC, whereas invasion assays thin ECM gel layers are placed and completely cover the microporous filter.

xenograft models. In these *in vivo* experiments, mostly suspensions of established tumour cell lines are injected into the animals. Frequently a subcutaneous location is chosen, but intraperitoneal or orthotopic injections can be also used.

Release of Degradative Enzymes

During tissue invasion, malignant cells have to penetrate intertissue barriers, such as subepithelial and subendothelial basement membranes. As mentioned above, the molecular compositions of these basement membranes are organ specific. Moreover, the patterns of degradative enzyme expression by malignant cells may be important in determining organ-specific formation of metastases. Malignant cells can be induced to synthesize, release and activate various types of these enzymes by specific adhesive interactions with EC or components of the basement membranes. There are a variety of degradative enzymes, inhibitors and activators released by tumour cells.

Models for the investigation of degradative enzyme release and/or regulation have been developed based on adhesion assays, where the adhesive interactions between tumour cells and ECM substrates are well known. Two types of modifications can be used for specific questions regarding enzyme synthesis and release. For example, to examine the secretion of degradative enzymes by tumour cells after adhesive contacts the degradative activity in the adhesion medium can be measured during an adhesion assay. Enzyme activity can be detected by zymogram electrophoresis, and active proteolytic enzymes can be identified in the gels. This procedure allows characterization by molecular weight, enzymatic activity and substrate specificity. Alternatively, the enzyme activity can be directly measured using radiolabelled or fluorogenic substrates. It is also possible to determine enzyme amounts immunologically using microtitre assays, or to detect secreted protein amounts of enzymes using Western immunoblotting techniques.

Increase in enzymatic activity in adhesion medium can be the result of activation of proenzyme forms, reduced inhibition or increased secretion, and secretion because of post-translational signalling and/or activated transcription. There are, probably, overlapping signalling pathways involved in regulation of degradative enzyme secretion. Signalling pathways that are involved in the induction of degradative enzyme secretion and activation can be studied using various pharmacological interventions or genetically engineered tumour cells. A distinction between translational processing of downstream signals and cytosolic signalling after adhesive contacts can be examined in a second type of experiments. Using basic types of adhesion assays, the synthesis of mRNA coding for specific degradative enzymes can be measured by hybridization techniques, reverse transcriptase polymerase chain reaction (RT-PCR) or Northern blotting.

GENERAL APPLICATIONS

Models for tumour cell adhesion and invasion have been established for investigation of two essential cellular processes during the development of distant metastases. Historically, they were used to identify participating molecules, which are mainly located on the tumour cell surface. After characterization of various classes of adhesion molecules, research focused more on their functional properties. It became evident that adhesion molecules are not simply anchors for extracellular contact, but also important participants in cellular signalling. This led to further methods, and animal models were frequently replaced by molecular biology approaches to modify specific properties. Since valuable alternatives are available for many hypotheses, the use of animal models for cell adhesion/invasion is under restriction for many reasons.

The current spectrum of investigations with regard to cell adhesion, migration and organ invasion includes most fields of cancer research. However, understanding the basic function of adhesive cell systems is not restricted to malignant cells and has only scratched the surface. Differences between normal cell functions, such as cell–cell interactions during wound healing or leucocyte trafficking, and specific properties of invasive and metastatic tumour cells are still under investigation. Knowledge of these specificities of cancer cells may result in new therapeutic strategies, such as inhibition of cell signalling pathways for cancer treatment. For example, the advantages in understanding the role of metalloproteinases for the formation of secondary tumour locations have already reached clinical importance. Animal models and also phase I and phase II studies have shown that treatment with inhibitors of MMPs might be a valuable new alternative.

INTERPRETATION OF EXPERIMENTAL RESULTS

Many cellular functions change during cell adhesion, migration and invasion, and often these properties are dependent on each other. Although theoretical models describe the processes for a single cell in numerous steps, these are difficult to separate, and they occur in parallel if larger numbers of cells are used. For example, spreading and migration include successful adhesive interactions between cells and substrate. Therefore, assays for cell spreading, migration or invasion are strongly influenced by adhesive properties of the cells. In addition, time is a very critical point for investigation of these cell properties. Certain steps of adhesive interactions are found only in limited periods and models have to consider this time course. Optimal results can be obtained if initial events or plateau values, at least temporarily, are determined. Alternatively, the best approach might be to examine properties when cellular functions are in their linear phase. However, extension of assays beyond plateau values may interfere with secondary cell reactions, such as release of degradative enzymes in adhesion assays. Although assays for cell adhesion, spreading, migration or detachment focus on different points of cell behaviour, these techniques may overlap each other. There are variable advantages and problems for each method and results should be confirmed by different techniques.

Although *in vivo* experiments investigate cell behaviour in a more natural environment, specific problems have to be considered using these experimental designs. In addition to the described problems of invasion assays, the injected cells have to survive in the organ environment, and they have to maintain their invasiveness and tumorigenicity. After subcutaneous injection, these facts are extremely relevant in many tumour cell systems. A second argument against the use of *in vivo* studies is the existence of suitable, reproducible and reliable *in vitro* alternatives. Therefore, the exploitation of *in vivo* invasion assays should be restrictive.

As described above, activation and inhibition of degradative enzymes play essential roles during tumour cell invasion. For interpretation of results with regard to these enzymes, a clear distinction has to be made between detection of enzymatic activity and protein amount, because most degradative enzymes are synthesized and secreted as inactive molecules and are involved in complex regulations of their activity.

Over the last decade, improved computer technology has allowed sophisticated data analysis. However, this tends to describe minor differences without biological relevance. For example, image analysis for the evaluation of cell spreading may give quantitative results, but can also overestimate slight morphological changes. Optimal analysis, but not maximal amounts of data, should be obtained to investigate a certain hypothesis.

BIOLOGICAL AND TECHNICAL LIMITATIONS

Metastatic tumour cells usually show phenotypic instability, which leads to gradual shifts in their behaviour and diversification of cells to form heterogeneous subpopulations that differ in their metastatic properties (Nicolson, 1988b). However, most of the common techniques that are currently available for the characterization of carcinomas and the evaluation of their potential metastatic behaviour are based on statistical analyses and the use of average values. This difference between clonal growth and the possibility for single cells to establish distant metastasis on the one hand and the requirement of most procedures for the use of large numbers of cells for the quantitative examination of cell adhesion and invasion on the other renders it more difficult to answer the question of how a tumour cell becomes metastatic.

The use of cultured tumour cells inherits many advantages and disadvantages (see the chapter *Basic Tissue Culture in Cancer Research*). A significant problem is the stability of cell behaviour in culture systems. For example, evaluation of adhesive properties and their comparison between cell lines with different metastatic behaviour may be influenced by culture conditions, such as duration of passage. In **Figure 10** it is shown that, for example, the quantification of cell adhesion to ECM components depends on the age of culture. Other possible factors are

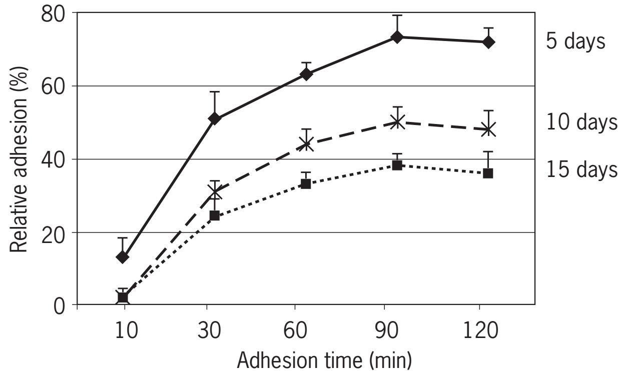


Figure 10 Influence of culture age on adhesive properties. HT-29P colon carcinoma cells were kept in culture for different times (5–15 days after passage) before use in a microtitre plate adhesion assay. With increasing age of culture (cells were leaving the logarithmic phase of growth), they lost their adhesive properties. The adhesion-time curve reached a plateau value where the number of adherent cells was at a maximum; after that time relative adhesion decreased because of release of degradative enzymes and subsequent destruction of the adhesive molecules or substrate (adhesive substrate: collagen I).

number of passages, selected growth and adhesion media, number of cells used for assays and type of cell preparation, among others. Therefore, all experimental data have to be considered with the background of this information. Additionally, once cell behaviour will be compared for cells with different metastatic properties, all conditions in the experimental design have to be analogous.

The choice of experimental designs must also respect technical aspects and possible bias of the method used. For example, in many assays washing steps are extremely critical and can cause false-negative results or a high nonspecific background. Adhesion assays require the removal of nonadherent cells for quantification of adhesive properties. This removal is often done by aspiration and/or washing with saline solutions where shear forces interfere with adhesive interactions. If fluid flow is too high, weak adhesion can be disrupted, but less washing strength increases nonspecific 'background adhesion.'

A further problem is that specific and nonspecific cellular events have to be distinguished. For example, to determine if chemotaxis towards a stimulating gradient is the motogenic factor for cell movement, it must be ruled out that enhanced migration is caused quantitatively by increased random migration. The effects on random migration can be estimated in a series of different concentrations of the supposed chemoattractant. Migration in a gradient of increasing concentrations that exceeds rates of random migration is evidence for chemotactic activity. Some experimental designs do not differentiate between enhanced migration and chemotaxis, and these results should be reported as 'stimulated (or enhanced) cell migration.'

All chemoattractants act by activation of receptors and signalling pathways, which might interfere with target functions, such as signalling mediated by adhesion molecules. Choosing an appropriate attractant is important in many migration or invasion assays. Conditioned medium obtained from cell cultures or various growth factors, such as EGF, bFGF and HGF, are widely used. Tissue extracts have also been found to act as chemoattractants with some cell types. Cells generally respond by migrating toward growth factors for which they have receptors.

Finally, some general remarks are necessary, but will not be discussed in detail. All adhesive or invasive interactions depend strongly on substrate concentrations, pH, cation contents and chosen medium. Antibodies or pharmacological substances may also require certain conditions for optimal action. When pharmacological stimulation or inhibition of cellular processes are investigated, these substances have to reach their target structures under living conditions (cell permeable). In addition, because of the assay's sensitivity, if possible stimulators or inhibitors are being assessed, several different concentrations of the test substance should be examined. Cell labelling has to avoid any interference with cell viability and adhesive/invasive properties. To exclude false-positive or false-negative results it is important to include controls in each assay. These controls have to determine nonspecific 'background' events (e.g. adhesion to nonspecific bovine serum albumin), and estimate effects of additives, such as reconstitution reagents (e.g. dimethyl sulfoxide). Negative results have to be confirmed by positive control experiments. Most assays should be performed in multiple parallel repeats for statistical analysis of test data.

PERSPECTIVES

The development of further models and methods for the examination of adhesive interactions will focus on consideration of *in vitro* conditions, which allow better simulation of *in vivo* characteristics. Real-time visualization of cellular processes, such as cytoskeletal activity or signal transmission, will gain functional understanding. Finally, analysis of complex interactions between different signalling cascades will also take part in investigation of metastatic behaviour.

REFERENCES

Akiyama, S. K., *et al*. (1994). Transmembrane signal transduction by integrin cytoplasmic domains expressed in single-subunit chimeras. *Journal of Biological Chemistry*, **269**, 15961–15964.

Atnip, K. D., *et al*. (1987). Chemotactic response of rat mammary adenocarcinoma cell clones to tumor-derived cytokines. *Biochemical and Biophysical Research Communications*, **146**, 996–1002.

Aznavoorian, S., *et al*. (1990). Signal transduction for chemotaxis and haptotaxis by matrix molecules in tumor cells. *Journal of Cell Biology*, **110**, 1427–1438.

Belloni, P. N., *et al*. (1992). Organ-derived microvessel endothelial cells exhibit differential responsiveness to thrombin and other growth factors. *Microvascular Research*, **43**, 20–45.

Birchmeier, W., *et al*. (1993) Molecular mechanisms leading to cell junction (cadherin) deficiency in invasive carcinomas. *Seminars in Cancer Biology*, **4**, 231–239.

Burtin, P., *et al*. (1983). Immunofluorescence study of the antigens of the basement membrane and the peritumoral stroma in human colonic adenocarcinomas. *Annals of the New York Academy of Sciences*, **420**, 229–236.

Cerra, R. F. and Nathanson, S. D. (1989). Organ-specific chemotactic factors present in lung extracellular matrix. *Journal of Surgical Research*, **46**, 422–426.

Conese, M. and Blasi, F. (1995). The urokinase/urokinase-receptor system and cancer invasion. *Baillieres Clinical Haematology*, **8**, 365–389.

Doerr, R., *et al*. (1989). Clonal growth of tumor-specific biomatrices and correlation with organ site specificity of metastasis. *Cancer Research*, **49**, 384–392.

Dorudi, S. and Hart, I. R. (1993). Mechanisms underlying invasion and metastasis. *Current Opinions in Oncology*, **5**, 130–135.

Fenyves, A. M., *et al*. (1993) Cultured microvascular endothelial cells (MVEC) differ in cytoskeleton, expression of cadherins and fibronectin matrix. *Journal of Cell Science*, **106**, 879–890.

Giavazzi, R., *et al*. (1993). Rolling and adhesion of human tumor cells on vascular endothelium under physiological flow conditions. *Journal of Clinical Investigation*, **92**, 3038–3044.

Haier, J., *et al*. (1999a). Influence of phosphotyrosine kinase inhibitors on adhesive properties of highly and poorly metastatic HT-29 colon carcinoma cells to collagen. *International Journal of Colorectal Diseases*, **14**, 119–127.

Haier, J., *et al*. (1999b). β1-Integrin mediated dynamic adhesion of colon carcinoma cells to extracellular matrix under laminar flow. *Clinical and Experimental Metastasis*, **17**, 377–388.

Haier, J., *et al*. (1999c). Different adhesion properties of highly and poorly metastatic HT-29 colon carcinoma cells with extracellular matrix components: role of integrin expression and cytoskeletal components. *British Journal of Cancer*, **80**, 1867–1874.

Inufusa, H., *et al*. (1995). Localization of oncofetal and normal fibronectin in colorectal cancer. *Cancer*, **75**, 2802–2808.

Kramer, R. H., *et al*. (1980). Metastatic tumor cells adhere preferentially to the extracellular matrix underlying vascular endothelial cells. *International Journal of Cancer*, **26**, 639–645.

Lawrence, M. B., *et al*. (1987). Effect of flow on polymorphnuclear leukocyte/endothelial cell adhesion. *Blood*, **70**, 1284–1290.

Lawrence, M. B., *et al*. (1990). Effect of venous shear stress on CD18-mediated neutrophil adhesion to cultured endothelium. *Blood*, **75**, 227–237.

Lengyel, E., *et al*. (1995). Induction of Mr 92,000 type IV collagenase expression in a squamous cell carcinoma cell line by fibroblasts. *Cancer Research*, **55**, 963–967.

Lichtner, R. B., *et al*. (1989). Differential adhesion of metastatic rat mammary carcinoma cells to organ-derived microvessel endothelial cells and subendothelial matrix. *Experimental Cell Research*, **57**, 146–152.

Liotta, L. A. and Stetler-Stevenson, W. G. (1991). Tumor invasion and metastasis: an imbalance of positive and negative regulation. *Cancer Research*, **51**, 5054s–5059s.

Liotta, L. A., *et al*. (1986). Tumor cell autokrine motility factor. *Proceedings of the National Academy of Sciences of the USA*, **83**, 3302–3306.

Marchetti, D., *et al*. (1993). Nerve growth factor effects on human and mouse melanoma cell invasion and heparanase production. *International Journal of Cancer*, **55**, 693–699.

Menger, M. D. and Vollmer, B. (1996). Adhesion molecules as determinants of disease: from molecular biology to surgical research. *British Journal of Surgery*, **83**, 588–601.

Menter, D. G., *et al*. (1992). Transglutaminase stabilizes melanoma adhesion under laminar flow. *Cell Biophysics*, **18**, 123–143.

Menter, D. G., *et al*. (1995). The role of trophic factors and autocrine/paracrine growth factors in brain metastasis. *Clinical and Experimental Metastasis*, **13**, 67–88.

Miyamoto, S., *et al*. (1995). Synergistic roles for receptor occupancy and aggregation in integrin transmembrane function. *Science*, **267**, 883–885.

Monsky, W. I. and Chen, W. T. (1993). Proteases of cell adhesion proteins in cancer. *Seminars in Cancer Biology*, **4**, 251–258.

Nicolson, G. L. (1988a). Cancer metastasis: tumor cell and host organ properties important in metastasis to specific secondary sites. *Biochimica Biophysica Acta*, **948**, 175–224.

Nicolson, G. L. (1988b). Organ specificity of tumor metastasis: Role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer Metastasis Reviews*, **7**, 143–188.

Nicolson, G. L. (1989) Metastatic tumor cell interactions with endothelium, basement membrane and tissue. *Current Opinions in Cell Biology*, **1**, 1009–1019.

Nicolson, G. L. (1991). Molecular mechanisms of cancer metastasis: tumor and host properties and the role of oncogenes and suppressor genes. *Current Opinions in Oncology*, **3**, 75–92.

Nicolson, G. L. (1995). Tumor interactions with the vascular endothelium and their role in cancer metastasis. In: Goldberg, I. D. and Rosen, E. M. (eds), *Epithelial–Mesenchymal Interactions in Cancer*. 123–156 (Birkhäuser, Basel).

Nicolson, G. L., *et al*. (1981). The role of fibronectin in adhesion of metastatic melanoma cells to endothelial cells and their basal lamina. *Experimental Cell Research*, **135**, 461–465.

Patton, J. T., *et al*. (1993). Computerized analysis of tumor cells flowing in a parallel plate chamber to determine their adhesion stabilization lag time. *Cell Motility and Cytology*, **26**, 88–98.

Pauli, B. U., *et al*. (1990). Organ-preference of metastasis: The role of endothelial cell adhesion molecules. *Cancer Metastasis Reviews*, **9**, 175–189.

Polette, M., *et al*. (1997). Induction of membrane-type matrix metalloproteinase 1 (MT-MMP) expression in human fibroblasts by breast adenocarcinoma cells. *Clinical and Experimental Metastasis*, **15**, 157–163.

Rice, G. E., *et al*. (1988). Tumor cell-endothelial cell interaction. *American Journal of Pathology*, **133**, 204–210.

Rodeck, U. and Herlyn, M. (1991). Growth factors in melanoma. *Cancer Metastasis Reviews*, **10**, 89–101.

Shapiro, S. D. (1998). Matrix metalloproteinase degradation of extracellular matrix: biological consequences. *Current Opinions in Cell Biology*, **10**, 602–608.

Shiozaki, H., *et al*. (1996). E-Cadherin mediated adhesion system in cancer cells. *Cancer*, **77**, 1605–1613.

Siletti, S., *et al*. (1991). Purifications of B16-F1 melanoma autokrine motility factor and its receptor. *Cancer Research*, **51**, 3507–3511.

Smith, T. W., *et al*. (1996). Computerized analysis of tumor cell interactions with extracellular matrix proteins, peptides, and endothelial cells under laminar flow. *Biotechnology and Bioengineering*, **50**, 598–607.

Tressler, R. J., *et al*. (1989). Correlation of inhibition of adhesion of large cell lymphoma and hepatic sinusoidal endothelial cells by RGD-containing peptide polymers with metastatic potential: role of integrin-dependent and-independent adhesion mechanisms. *Cancer Communications*, **1**, 55–63.

Virtanen, I., *et al*. (1990). Integrins in human cells and tumors. *Cell Differentiation and Development*, **32**, 215–228.

Watanabe, H., *et al*. (1991). Purification of human tumor cell autokrine motility factor and molecular cloning of its receptor. *Journal of Biological Chemistry*, **266**, 13442–13448.

Yun, Z., *et al*. (1997). Differential adhesion of metastatic RAW117 large-cell lymphoma under static conditions: role of the $\alpha v\beta 3$ integrin. *Clinical and Experimental Metastasis*, **15**, 3–11.

FURTHER READING

Belloni, P. N. and Tressler, R. J. (1990). Microvascular endothelial cell heterogeneity: interactions with leukocytes and tumor cells. *Cancer Metastasis Reviews*, **8**, 353–389.

Gumbiner, B. M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell*, **84**, 345–357.

Juliano, R. (1993). The role of $\beta 1$ integrins in tumors. *Seminars in Cancer Biology*, **4**, 277–283.

Nicolson, G. L. (1994). Tumor cell adhesion and growth properties in organ preference of tumor metastasis. *Advances in Molecular and Cell Biology*, **9**, 123–152.

Parson, J. T., *et al*. (1994). Focal adhesion kinase: structure and signaling. *Journal of Cell Science*, **18**, 109–113.

Schwarz, M. A., *et al*. (1995). Integrins: emerging paradigms of signal transduction. *Annual Reviews of Cellular Developmental Biology*, **11**, 549–599.

Weiss, L. (1992). Biomechanical interactions of cancer cells with the microvasculature during hematogenous metastasis. *Cancer Metastasis Reviews*, **11**, 227–235.

Yamada, K. M. and Geiger, B. (1997). Molecular interactions in cell adhesion complexes. *Current Opinions in Cell Biology*, **9**, 76–85.

Yamada, K. M. and Miyamoto, S. (1995). Integrin transmembrane signaling and cytoskeletal control. *Current Opinions in Cell Biology*, **7**, 681–689.

Zetter, B. (1993). Adhesion molecules in tumor metastasis. *Seminars in Cancer Biology*, **4**, 219–229.

Chapter 64

Tumour Metastasis Models

Richard Hill
Princess Margaret Hospital, Ontario Cancer Institute, Toronto, Ontario, Canada

CONTENTS

- Biological Background
- Principles of Model Establishment
- Interpretation of Experimental Results
- Biological Limitations
- Technical Considerations
- Perspectives

BIOLOGICAL BACKGROUND

Metastasis is a process in which tumour cells establish new growths at sites in the body distinct from the primary tumour. The spread of the tumour cells can occur either via lymphatic drainage of the primary site leading to local lymph node metastases or via the blood vasculature leading to metastases in distant organs. Further spread of tumour cells from these sites may also occur, particularly spread from the lymphatic system into the blood vasculature. These processes involve a number of specific stages, which include escape of the tumour cell from the primary site, survival in the circulation, arrest in the capillaries at a new location, migration from the capillary into the interstitial space and establishment of a new growth at this new location. These stages represent potential barriers which a tumour cell must be able to surmount if it is to be successful in forming a metastasis. The ability of tumour cells to complete the various stages of this process can be studied using *in vitro* assays, as discussed in earlier chapters (see *Models for Tumour Growth and Differentiation*; *Models for Tumour Cell Adhesion and Invasion*; *Invasion and Metastasis*), but *in vitro* assays do not allow an assessment of whether an individual cell is able to complete all the stages successfully. Such an assessment requires the use of an *in vivo* metastasis model to integrate all the various stages of the process. There is currently no substitute for such *in vivo* models.

The various stages of the metastatic process are illustrated in **Figure 1**. The initial stages involve detachment of cells from the primary tumour and their invasion into the lymphatic or systemic circulation. Since there are few (if any) functional lymphatics in tumours, cells near or at the invasive edge of the tumour are likely to be those which escape into the lymphatic system. Tumour blood vessels are often poorly formed owing to the rapid angiogenesis which may occur during tumour growth and this may provide for easier access of tumour cells into the systemic circulation. Reduced expression of homotypic cell adhesion molecules such as the cadherins may enhance the detachment of cells from the tumour mass. Concurrently, increased expression of proteolytic enzymes, such as matrix metalloproteinases, cathepsins or plasminogen activators, or decreased expression of their inhibitors, may assist invasion of the cells into the circulatory system. Once in the circulation the cells are most likely to be arrested in the first capillary bed that they encounter owing to size constraints and/or to an increased possibility of interacting with adhesion molecules expressed on the surface of the endothelial cells (or exposed basement membrane) at lower flow rates. Depending on the location of the tumour the first capillary bed will be in either the liver (for intestinal tumours growing in organs whose blood drains into the portal system) or the lung (for tumours in organs in other parts of the body). Some of the arrested cells may successfully transit through this initial ‘first-pass’ organ and move on to other sites, but studies with labelled cells have suggested that the proportion of the cells which do so is low and depends on the cell type. In both cases the cells must be able to survive both the physical stress of the bloodstream and also the various components of the immune system. Natural killer cells and classic T cell immunity have been implicated in reducing metastasis formation and this may be a particularly important factor if the tumour cells have any degree of immunogenicity. Following arrest, the cells may establish tight adherence to the vessel wall, extravasate into the interstitial space and possibly migrate within this space before starting to proliferate. The proteolytic enzymes discussed above in relation to intravasation are also likely to be involved in this extravasation stage. Some tumour cells can apparently bypass this stage by initiating growth whilst still in the blood vessel. Initiation of proliferation requires that the cells can either respond to locally available growth factors

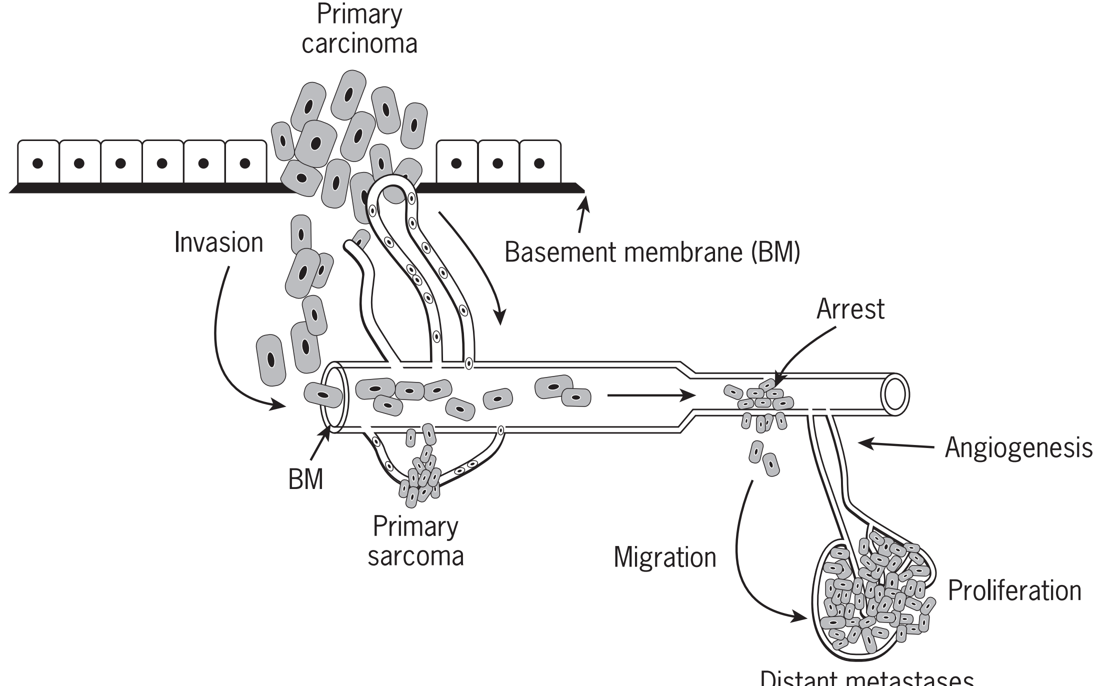


Figure 1 The major steps of the metastatic process. [From Tannock and Hill (eds.), 1998, *The Basic Science of Oncology*, 3rd edn (McGraw-Hill, New York).]

or can make sufficient growth factor themselves to achieve autocrine stimulation of growth. Meanwhile, they must also be able to avoid or ignore any negative growth signals. Finally, if progressive growth to a size greater than about 1 mm diameter is to occur the tumour cells must be able to initiate angiogenesis to stimulate the formation of new blood vessels. Proteolytic enzymes may also be involved in the growth stimulation stage, since some growth factors are sequestered in the interstitial matrix (e.g. VEGF) or are bound to binding proteins in the circulation (e.g. IGF). These factors can be released by proteolytic enzyme activity. (See also chapters on *Angiogenesis Models*; *Models for Tumour Cell Adhesion and Invasion*; *Models for Tumour Cell–Endothelial Cell Interactions*.)

Since these stages are sequential, only a few cells may be successful in completing all of them. Thus an important aspect of metastasis is that it is a very inefficient process. Following intravenous injection of cells even the most malignant cell populations rarely form metastases with an efficiency greater than about 1% and it is often several orders of magnitude lower. Similarly, if radiolabelled tumour cells are injected into the circulation and radioactivity in the host is monitored as a function of time, the remaining activity is usually below 1% within 2–3 days of injection of the cells. If the number of cells escaping into the circulation from a primary tumour is estimated by taking blood samples from efferent veins, it is found that spontaneous metastatic efficiency is also very low, although in this case the viability of the escaping tumour cells is difficult to assess. Whether the low efficiency of metastasis formation represents a situation in which there is a cumulation of low probabilities of surmounting a number of stages or whether it represents a very low probability that a tumour cell can surmount a specific stage of the metastatic process (a rate-limiting step) is currently unclear and certainly these are not mutually exclusive possibilities.

It has been argued that metastatic inefficiency is purely a result of the random chance that any individual tumour cell will surmount all the various stages of the process. This is consistent with the first possibility above and with the findings that cells derived from metastases do not necessarily have a metastatic efficiency much different from those derived from the primary tumour. However, as discussed below, there is increasing evidence that specific phenotypic characteristics can affect the metastatic potential of tumour cells and a number of 'metastasis-related' genes have been isolated in studies of clonal populations of cells, which have different metastatic potential but are derived from the same parental cell population. This evidence favours the 'rate-limiting step' model and the prevailing view is that arrest and extravasation are the primary rate-limiting steps. However, two recent studies have questioned this dogma. Chambers *et al*. (1999), who used intravital videomicroscopy (IVVM) (see **Figure 2**) to track the early stages of metastasis formation by B16 cells in mouse liver or chick embryo chorioallantoic membrane (CAM), found that a major barrier to successful metastasis formation in the system studied occurs at the final 'initiation of new growth' stage of the

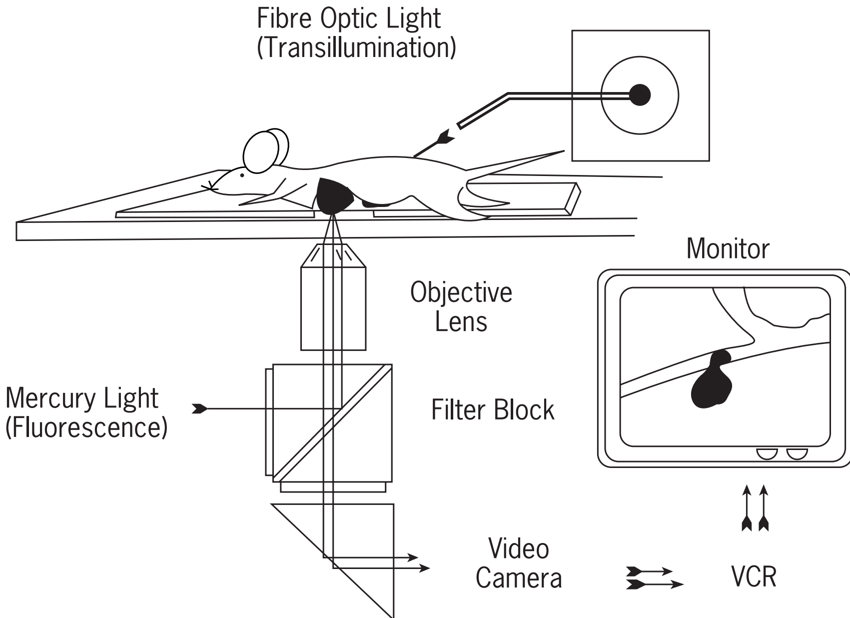


Figure 2 Schematic diagram of the intravital videomicroscopy system for examination of metastases *in vivo*. [From Tannock and Hill (eds.), 1998, *The Basic Science of Oncology*, 3rd edn (McGraw-Hill, New York).]

process. Similar findings, obtained using a careful cell counting technique, have recently been reported by the same group for tumour cells forming metastases in the lung (Cameron *et al*., 2000). In contrast, Al-Mehdi *et al*. (2000) using green fluorescent protein (GFP)-labelled HT-1080 tumour cells observed that arrest in the lung occurred in precapillary arterioles as a result of adhesion to the vessel wall and that initial growth of metastases occurred in the vessel without extravasation.

One of the most compelling pieces of evidence that specific phenotypes are associated with metastasis formation is the clinical observation that certain types of tumours have a propensity to metastasize to specific organs. Thus many tumour types will metastasize to lung or liver but prostate and breast carcinomas have a propensity to metastasize to bone whilst small cell lung carcinomas often metastasize to brain or bone marrow and melanomas to brain or bowel. These clinical observations have been supported by a number of experimental studies which have demonstrated that it is possible to select cells which will preferentially metastasize to specific organs. The now classical example of such studies is the selection of the B16–F10 population of B16 melanoma cells which have increased ability to form lung metastases (Fidler, 1973). These cells were selected by injecting B16 melanoma cells intravenously into syngeneic C57Bl mice, growing cells in culture that were isolated from the lung metastases which formed and then injecting the cells back into animals to form new lung metastases. This cycle was completed 10 times to produce the F10 population. Selection of populations of cells with increased propensity to form metastases in many different organs has been reported using similar techniques. It has been demonstrated that such cells will 'home' to the relevant organ even when it is ectopically transplanted in the animal.

It has also been shown that if multiple (clonal) populations of cells are grown from individual tumour cells there is substantial heterogeneity in the metastatic potential of the cells from the different clones. This finding was initially interpreted as indicating the presence of preexisting metastatic variants within a cell population but the finding that the metastatic properties of these clonal populations tends to be unstable has cast doubt on this interpretation. Tumour cell populations are genetically unstable and the cells rapidly become heterogeneous in their phenotypic properties during the growth of the 'clone.' Nevertheless, 'clonal heterogeneity' of metastasis formation is still consistent with the concept that specific phenotypes are associated with metastasis formation, even though these phenotypes may be only transiently expressed. Since most of the stages of the metastatic process are likely to occur over a short time period of a few hours to a few days, prolonged expression of many of the relevant phenotypes may not be necessary for successful metastasis formation.

These experimental results have provided support for a hypothesis about metastasis formation that was postulated more than 100 years ago (Paget, 1889). The so-called 'soil and seed' hypothesis was based on the idea that tumour cell–host cell interactions can occur and that these are more or less favourable for metastasis development. Thus the preference of specific tumours for metastasis to

specific organs can be explained by the concept that specific organs provide a more suitable growth environment for cells of a particular tumour type. Meanwhile, the propensity of many tumours to form lung or liver metastases can be explained by the fact that cells released into the circulation will tend to be arrested in the first-pass organ, thus producing a high cell burden in these organs.

PRINCIPLES OF MODEL ESTABLISHMENT

Although there are many variations in the applications of *in vivo* metastasis models, there are essentially two types of model (usually in rodents). In a spontaneous metastasis model, a primary tumour is initiated in some location in the body and the formation of metastases is monitored at one or more different locations (**Figure 3**). In an experimental metastasis model tumour cells, obtained from culture or by preparing a cell suspension from a tumour, are introduced directly into the circulation (e.g. by intravenous, intraportal or intra-arterial injection) and development of metastases is monitored at various times later (see chapter on *Human Tumours in Animal Hosts*). For most rodent models intravenous or intraportal injection results in metastases which are usually observed in the lung or liver, respectively, depending on the site of injection, unless specific selection has been undertaken for tumour cells which will form metastases in other organs. Following intra-arterial injection metastases in a wide range of organs may be observed. The experimental metastasis model is believed to provide an intermediate level of complexity, since it bypasses the initial events associated with escape of the tumour cells into the circulation. However, this concept of reduced complexity depends on the assumption that cells which escape from a primary tumour and go on to form metastasis have the same phenotypic characteristics as cells injected from a prepared cell suspension.

A very important factor in the choice of a metastasis model is that it should be appropriate for the question being asked. Issues which need to be considered include whether a syngeneic rodent tumour model or a xenografted human tumour model will be used. A rodent model is easier to handle and there are many highly metastatic rodent tumour models available. Such models are often rapidly growing and the primary tumour may become very large before spontaneous metastases are detectable macroscopically. One solution to this problem is to excise the primary tumour by surgery or ablate it with radiation to allow additional time for metastases to develop. Such a strategy can also be used to provide information about the kinetics of the seeding of spontaneous metastases from a primary tumour. It always needs to be established that such rodent models simulate the relevant clinical scenario. One situation where this may not be the case is in the use of chemically or virally induced tumours. Such tumours are usually immunogenic to a detectable degree. Such immunogenicity may not prevent the growth of a primary tumour from a large cell inoculum but may prevent the growth of small metastases. A rodent model to assess propensity for lymph node metastasis can be set up by transplanting tumour cells into the hind footpad, from where cells can metastasize to the popliteal lymph node.

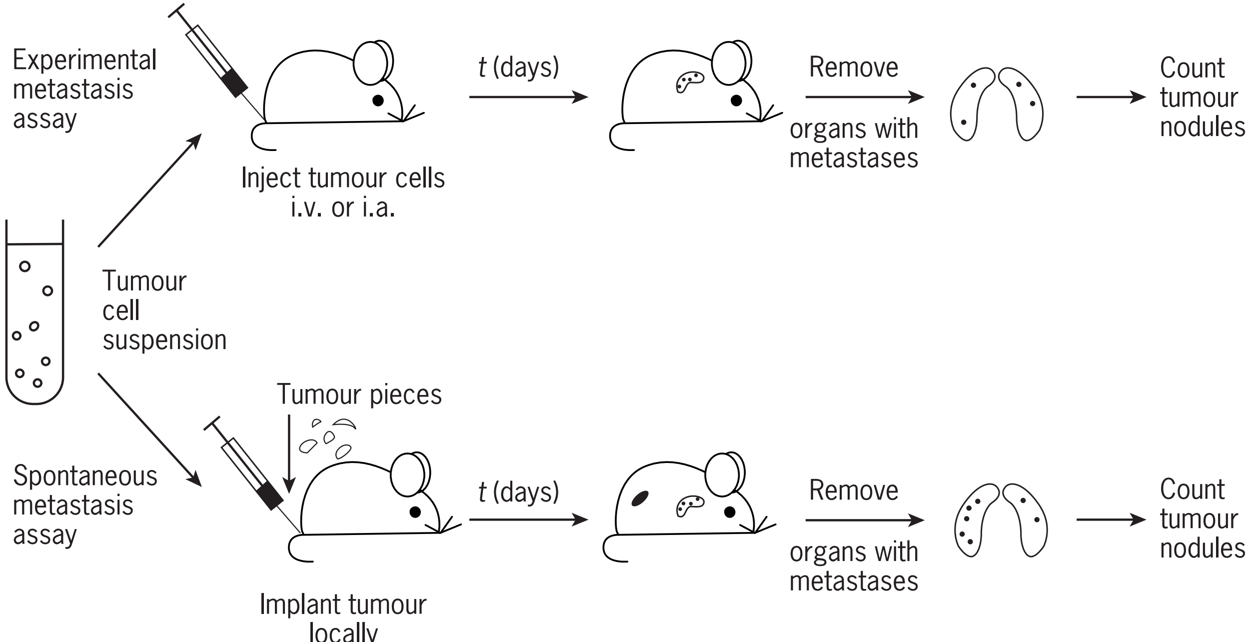


Figure 3 Representation of the experimental and spontaneous metastasis assays in mice. [From Tannock and Hill (eds.), 1998, *The Basic Science of Oncology*, 3rd edn (McGraw-Hill, New York).]

A refinement of rodent tumour model systems is the use of transgenic or knockout mice as hosts for the growth of the primary tumour or as recipient mice in experimental metastasis assays (see chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*). One example of this was the use of mice that were knocked out for the gene for Fas ligand (C3H/gld mice) in studies examining the importance of the expression of Fas on the surface of the tumour cells for metastasis formation. (See chapter on *Apoptosis*.) A difference in the number of experimental metastases which formed in normal C3H mice following injection of cells from two different sublines of a murine melanoma, which expressed different levels of Fas on their surface, was abolished when the cells were injected into C3H/gld mice (Owen-Schaub *et al*., 1998). A second example was the use of a transgenic model carrying the polyoma middle T antigen (PMTA) driven by the mouse mammary tumour virus promotor. These mice develop a high incidence of metastatic mammary tumours at an early age but this is reduced if the gene for Mgat5, a glucosaminyl transferase enzyme involved in a specific type of glycosylation of proteins, is knocked out in the mice (Granovsky *et al*., 2000). In another experiment with these PMTA transgenic mice, crossbreeding with a wide range of other inbred strains of mice demonstrated the importance of the genetic background of the mice on the latent period before tumour growth and on the extent of metastases that develop in the mice (Lifsted *et al*., 1998).

For xenografts various immune-deprived hosts are available, particularly nude or SCID mice (see chapter on *Human Tumours in Animal Hosts*). There have been reports that xenografted human tumours have a greater likelihood of metastasizing if grown in SCID mice (Xie *et al*., 1992). There is also strong evidence that the site of tumour implantation affects the formation of spontaneous metastases and that more metastases are likely if the tumour is transplanted into an orthotopic rather than ectopic site (i.e. if the tumour is growing in the organ of origin of the tumour) (Radinsky and Fidler, 1992). The sites of metastasis formation are also more likely to conform to the pattern observed in clinical practice for that tumour type. One refinement of this model is to use it in conjunction with tumour cells labelled with green fluorescent protein (GFP), which provides for the possibility of direct visualization of the growth of the tumour and its metastasis *in situ* using the GFP marker (Hoffman, 1999). A second refinement of this model is the transplantation of human tissue (e.g. fetal bone) into SCID mice (Shtivelman and Namikawa, 1995; Nemeth *et al*., 1999). These so-called SCID-hu models can then be used as recipients for growth of human xenografts and have been reported to show similar organ preference for metastasis formation to that seen in clinical practice. Such metastases appear to occur preferentially in the transplanted human tissue as opposed to the same murine tissue. These models are believed to recapitulate both the ‘soil’ and ‘seed’ aspects of Paget’s original hypothesis. A recent report has described the development of lymph node metastases following orthotopic transplantation of a human stomach Ca into SCID mice (Fujihara *et al*., 1998).

A nonrodent model that has been widely used for studies of the malignancy of tumour cells is the chick chorioallantoic membrane (CAM) model (**Figure 4**). The CAM functions to provide oxygen to the growing chick and can be exposed by carefully removing a portion of the egg shell. Since the developing chick has no functional immune system during its early growth, the CAM provides a site where human or other tumours can be transplanted and grown. Alternatively, tumour cells can be injected directly into the CAM veins. In both cases the internal organs of the chick can be examined at some time later for evidence of metastatic growth. This system is easy to handle and is amenable to direct microscopic observation but there is a time window of only about 7–10 days for tumour growth before the chick immune system starts to develop.

One advantage of an experimental metastasis model, over a spontaneous metastasis model, is the possibility of quantitating the efficiency of metastasis formation per cell. Such quantitation can be done at the end of the experiment by counting the number of metastatic nodules formed and relating it to the number of cells injected or it can also be done at earlier stages of the process using labelled cells. Tracking of radiolabelled cells immediately after intravenous injection has indicated trapping of essentially all the cells in the lung during their first pass through the capillary bed and has further suggested that there is rapid

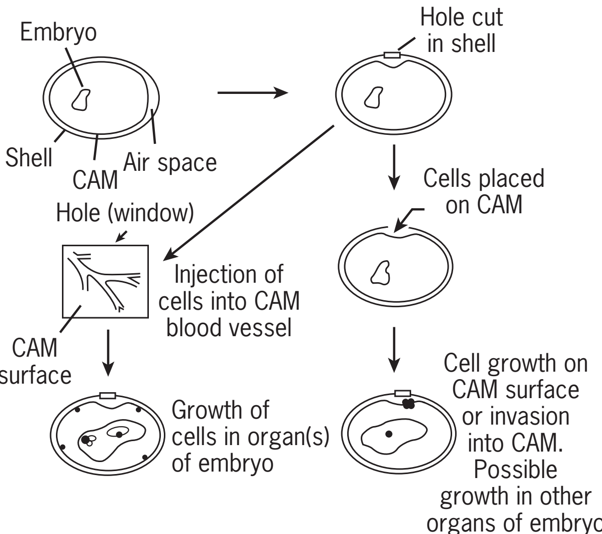


Figure 4 Assay for metastatic properties of tumour cells using the chick chorioallantoic membrane (CAM). [From Tannock and Hill (eds.), 1998, *The Basic Science of Oncology*, 3rd edn (McGraw-Hill, New York).]

death and loss of the majority of cells over the next 24–48 h. Experiments using the IVVM technique and/or with cells labelled with fluorescent markers have indicated a much slower loss of the cells, suggesting that the radiolabel may have played a role in the death of the injected cells. The availability of easily observable genetic markers, such as GFP, *lac Z* or luciferase, will allow further study of this issue and will potentially allow such observations to be extended to spontaneous metastasis models (Hoffman, 1999; Kruger *et al*., 1999; Rubio *et al*., 2001).

An alternative approach in such studies is to remove the lungs (or other organ) at various times after injection, to prepare a cell suspension and to reinject the cells to test for metastatic ability of the trapped cells. In one such study, Young and Hill (1986) found that tumour cells recovered within a few days of injection had increased metastatic ability relative to the initially injected cell population but that cells recovered at 3 weeks after injection, when there were macroscopic metastatic nodules, did not. They interpreted these findings as an indication of the presence of a transient metastatic phenotype which was expressed by cells remaining in the lung at short times after injection but which decayed over the course of the early microscopic growth of the metastases. The role of transient phenotypes in metastasis formation remains to be clearly delineated. It is possible that such phenotypes represent transient differences in the expression of relevant genes in individual cells and that the process of selection of cells for the formation of metastasis in specific organs represents the stabilization of such differences within the cell population. Since the initial establishment of a metastasis is likely to be a process which takes only a short time (minutes to hours) from cell escape from the primary tumour to lodgement and extravasation at a new location, it may not be necessary that phenotypes involved in the early stages of the metastatic process are stably expressed. Conversely, phenotypes associated with the growth of the metastasis at a new site may need to be more stable and, once the micrometastatic growth is established and angiogenesis initiated, such properties may be similar to those expressed by cells in the primary tumour.

INTERPRETATION OF EXPERIMENTAL RESULTS

As noted above, an attractive feature of an experimental metastasis model is that it is possible to quantitate the efficiency of metastasis formation per cell, since a known number of cells are injected (e.g. intravenously) and the number of metastatic nodules formed can be counted. However, there are a number of important issues which need to be considered if quantitative comparisons of different cell populations are to be made (Hill *et al*., 1986). One issue is whether the number of nodules formed is linearly related to the number of cells injected. For highly metastatic cell populations this is only true over a certain range of cell numbers injected ($\lesssim 2 \times 10^5$ tumour cells per mouse) and number of metastatic nodules formed (<100 per mouse lung). When high cell numbers are injected, nonspecific factors may increase nodule formation whilst large numbers of nodules per lung prevent accurate counting of all the nodules, leading to an underestimate of the true number. A second concern is that there is usually substantial heterogeneity in the number of metastases observed from one animal to another. This is particularly noticeable with spontaneous models but is also an issue for experimental metastasis models. Analysis of data from an experimental metastasis model demonstrated that the range in the numbers of metastases observed in different mice, injected with the same number of tumour cells from the same cell suspension, was much greater than that expected from statistical considerations. Furthermore, when two different cell types were injected into the same mice at the same time, the heterogeneity that was observed in the number of metastases arising from the two different cell populations was not correlated between the individual mice. The reason for this heterogeneity is not well understood but it implies that there are significant uncontrolled variables. Whether these are related to the animals or to the cells or both remains to be determined. The practical result of the heterogeneity is that it is necessary to do comparison studies with groups containing a minimum of 5–10 mice and to analyse the results using nonparametric statistical tests. It is also important to run controls at every stage of a series of experiments because there can be significant week to week variations in the efficiency of metastasis formation.

One of the most constant features of metastasis studies has been the heterogeneity of cellular properties which have been associated with metastasis formation (see chapter on *Advantages and Limitations of Models for Cancer and Malignant Cell Progression*). This was well illustrated some years ago by Maslow (1989), who reviewed the literature for studies of factors affecting metastasis of B16 melanoma cells, the tumour line used in the early studies of Fidler. He listed over 50 cellular properties that had been shown to affect metastasis formation by these cells using either spontaneous or experimental assays. These properties spanned the whole range of the metastatic process. He also noted that experimental and spontaneous metastasis assays do not always give the same results. An example is the B16 F10 cell line, which is 10-fold more metastatic than the B16 F1 population when tested in an experimental assay, but these cell lines show no difference in spontaneous metastasis formation. While such results emphasize the complexity of the metastatic process and the need for careful and controlled experimentation, one overall message from a wide range of observations is that different cell types manifest different properties associated with metastasis formation.

BIOLOGICAL LIMITATIONS

Biological limitations of *in vivo* metastasis models include those that limit the question that can be asked and those which affect the extent to which a model is similar to the clinical situation. *In vivo* models are limited by the fact that it is difficult to observe the progress of individual cells through the various steps of the metastatic process. In standard spontaneous or experimental metastasis studies what occurs between the time the cells are injected and metastases are observed can usually only be surmised, based on measured properties of the cell populations that are injected. Thus, if cells with decreased expression of adhesion molecules or increased expression of metalloproteinase activity form more metastases, then it is argued that arrest and/or invasion are important factors in controlling metastasis formation by such cell populations. Whether the cells which actually formed the metastases (a very small fraction of the total number of cells injected) were expressing these specific properties is not known. Furthermore, as shown by the IVVM studies, whether the association of expression of such properties with a specific stage of the metastatic process is necessarily correct is also not known (Chambers and Matrisian, 1997). This 'black box' aspect of the *in vivo* assays has been alleviated to some extent by the use of the IVVM technology and/or the use of labelled cells which allow observation of the tumour cells *in vivo* (Al-Mehdi *et al*., 2000), but the issue of whether the properties of the cell population as a whole reflects the properties of the small fraction of cells which form metastasis remains largely unknown. Addressing this question requires the examination of the properties of individual cells at various stages during the metastatic process. To date the IVVM approach has only been applied with the use of experimental metastasis models and this combination has provided important numerical information about cell seeding and early growth of metastasis. The application of this approach, or an alternative approach which could provide similar information, with a spontaneous metastasis model is clearly desirable.

The development of the orthotopic transplantation approach arose from investigation of the reason for the findings that human xenografts growing subcutaneously in SCID or nude mice rarely produced metastases. Orthotopic models now often provide good correspondence with the general patterns of metastatic spread for different types of human tumour and provide for opportunities to address specific questions associated with these specific patterns. The introduction of the SCID-hu model with transplanted human tissue adds further to this approach and provides for a link between *in vitro* studies of cell growth in such human tissue and *in vivo* formation of metastases in this tissue. Orthotopic transplantation in rodent tumour models has not been so extensively studied but there is evidence that there is increased metastasis following such transplantation. The ability to manipulate the genetic background of rodent models makes them attractive for studying specific aspects of the metastatic process.

TECHNICAL CONSIDERATIONS

Many different factors can affect the performance and reliability of metastasis assays. These factors have been extensively reviewed by Welch (1997) and a summary of his analysis is presented in **Table 1**. For both spontaneous and experimental metastasis assays, important considerations include the health and care of the animals being used and whether there is an immune incompatibility between the animals and the tumour cells. For spontaneous metastases the site of the tumour can affect the number and site of metastases, as discussed previously. For experimental metastases the handling and culture of the tumour cells prior to injection into the host animal can either increase or decrease metastasis formation, depending on the conditions. For example, both high confluency in the culture and the presence of clumps of cells in the suspension can increase metastasis formation whereas poor cell viability will reduce metastasis formation. A related issue is that it is desirable to maintain the tumour cell cultures with regular passages *in vivo* alternating with those *in vitro* or by frequent return to frozen stock. Tumour cells which have been maintained in long-term culture without *in vivo* exposure may demonstrate decreased metastasis efficiency, presumably due to the development of immune incompatibilities between the cells and the host animals.

Two aspects of the above may be considered more than technical issues. The first is the finding that treatment of the host animal with drugs (particularly cyclophosphamide) can modify the formation of metastases. This was particularly observed for experimental metastasis assays in which the metastases were formed in the lung and it was possible to modify metastatic levels by greater than 10-fold (Carmel and Brown, 1977). However, the effect observed with spontaneous metastasis assays was much reduced or absent. Similar observations were reported for irradiation of the lung prior to irradiation. The explanation for these results was suggested to be the damage to the endothelium caused by the treatments leading to easier access of the cells to the interstitial space. These results suggest that any form of treatment or perhaps infection that causes inflammation in the lungs may affect metastasis formation in that environment. A similar but opposite effect was observed with agents such as heparin that reduce blood clotting. Treatment of recipient animals with such agents was found to decrease experimental lung metastasis formation significantly. It was hypothesized

Table 1 Technical factors affecting metastasis assays

Technical factor	Issues	Comments
Recipient animal	Syngeneic (rodent tumours) Immune suppressed such as SCID or nude mice (human tumours)	Tumours grown in SCID mice have been reported to have higher metastatic propensity than when grown in nude mice
Health of recipient animals	Latent or viral infections Stress of recent transport	Good health very important for reproducible results
Immunogenicity of cells	Many chemically or virally induced tumours in rodents are immunogenic	Immunogenicity can reduce metastatic potential
Animal age and gender		Variations in metastatic potential with age and sex have been observed even for tumours that are not hormonally dependent
Site of injection of primary tumour	Orthotopic vs ectopic	Orthotopic transplantation usually leads to increased metastases
Ectopic organs (human)	Transplantation into SCID mice	Human organs may demonstrate metastatic deposits from human tumour cells when same mouse organ does not
Preparation of cells for experimental metastasis assay	Level of cell viability, no. of cells injected. Presence of clumps of cells. Method of harvesting cells	All factors can modify the no. of metastases observed
Culture conditions	Confluence of culture, acidity of culture, length of time in culture	Can modify the no. of metastases observed

that the explanation for these findings was that there was reduced retention of the cells due to reduced ability to form thrombi around the tumour cells. Again, it proved difficult to reproduce these findings using spontaneous metastasis models.

The second aspect of interest is the finding that deliberate treatment of tumour cells with a growth stress, such as exposure to hypoxia, low pH or low glucose or folate levels can also lead to a transient increase in metastatic efficiency. These results are best illustrated by exposures to hypoxia which have been shown to cause large increases in experimental metastases in both rodent and human tumour cell lines (Young *et al.*, 1988; Rofstad and Danielsen, 1999). These differences can also be observed in a spontaneous metastasis model and appear to link directly to observations in patients where it has been observed that those with more hypoxic tumours are more likely to have metastases. Again, the explanation for these findings is uncertain but it is likely that changes in gene expression associated with hypoxia (such as upregulation of VEGF) play an important role. Another factor may be that a low oxygen environment in tumours may produce oxic stress (reperfusion injury), which in turn may lead to increased DNA damage and mutations in the tumour cells.

PERSPECTIVES

A clear understanding of the critical steps in the process of metastasis and the differences which exist in the abilities of different tumour cells to form metastases remains elusive. The recent work using the IVVM technology and careful cell counting techniques to track cells during experimental metastasis formation has re-emphasized the later stages of the process. Whether the ability to grow at the new site is the limiting factor in metastasis formation in all organs, or with all tumours, remains to be established but experience over the last 30 years, which has identified a wide level of heterogeneity in metastasis-related factors, suggests that this is unlikely. These studies do, however, focus attention back on the seed and soil hypothesis of Paget (1889) and on issues associated with early growth of metastases. They are consistent with the increasing recognition of the importance of cell–cell and cell–matrix interactions in the control of cellular phenotype and with the fact that

many genes that have been isolated as metastasis-specific actually code for functions which relate to cell growth.

Metastases arise from single cells (or a small clump of a few cells), so it is the properties (and interactions) of individual cells rather than of populations of cells that are critical for metastasis formation. As noted above, the environment in which a cell exists can have a significant bearing on its behaviour, thus in the context of the soil and seed hypothesis it is appropriate to think of the properties of the seed in terms of the cell and its initial microenvironment. In parallel, the concept of soil should include the idea of the metastatic cell interacting with and modifying its new microenvironment. Many micrometastases apparently do not ever grow to form macrometastases. This raises the issue of whether phenotypes associated with metastasis formation necessarily represent stable changes and further whether they are truly metastasis-specific rather than reflecting more extreme versions of properties expressed by many of the cells in the primary tumour.

The newer models for metastases, such as the SCID-hu model and the transgenic and knockout mouse models, are opening up new ways to address some of these issues. The use of new technologies, such as DNA arrays, is providing more rapid ways to identify gene expression changes in cell populations. In combination with techniques such as laser capture microdissection, it will be possible to examine gene expression in small numbers, and perhaps individual, metastatic cells. The application of these new technologies will certainly increase our understanding of the complexities of the metastatic process and provide information which can be more directly related to human disease.

REFERENCES

Al-Mehdi, A. B., *et al*. (2000). Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis. *Nature Medicine*, **6**, 100–102.

Cameron, M. D., *et al*. (2000). Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency. *Cancer Research*, **60**, 2541–2546.

Carmel, R. J. and Brown, J. M. (1977). The effect of cyclophosphamide and other drugs on the incidence of pulmonary metastases in mice. *Cancer Research*, **37**, 145–151.

Chambers, A. F. and Matrisian, L. M. (1977). Changing views of the role of matrix metalloproteinases in metastasis. *Journal of the National Cancer Institute*, **89**, 1260–1270.

Chambers, A. F., *et al*. (1999). Preclinical assessment of anticancer therapeutic strategies using *in vivo* videomicroscopy. *Cancer and Metastasis Reviews*, **17**, 263–269.

Fidler, I. J. (1973). Selection of successive tumor lines for metastasis. *Nature New Biology*, **242**, 148–149.

Fujihara, T., *et al*. (1998). Establishment of lymph node metastatic model for human gastric cancer in nude mice and analysis of factors associated with metastasis. *Clinical and Experimental Metastasis*, **16**, 389–398.

Granovsky, M., *et al*. (2000). Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nature Medicine*, **6**, 306–312.

Hoffman, R. M. (1999). Orthotopic transplant mouse models with green fluorescent protein-expressing cancer cells to visualize metastasis and angiogenesis. *Cancer and Metastasis Reviews*, **17**, 271–277.

Kruger, A., *et al*. (1999). The bacterial lacZ gene: an important tool for metastasis research and evaluation of new cancer therapies. *Cancer and Metastasis Reviews*, **17**, 285–294.

Lifsted, T., *et al*. (1998). Identification of inbred mouse strains in harboring genetic modifiers of mammary tumor age of onset and metastatic progression. *International Journal of Cancer*, **77**, 640–644.

Maslow, D. E. (1989). Tabulation of results on the heterogeneity of cellular characteristics among cells from B16 mouse melanoma cell lines with different colonization potentials. A summary of sixty reports. *Invasion and Metastasis*, **9**, 182–191.

Nemeth, J. A., *et al*. (1999). Severe combined immunodeficient-hu model of human prostate cancer metastasis to human bone. *Cancer Research*, **59**, 1987–1993.

Owen-Schaub, L. B., *et al*. (1998). Fas and Fas ligand interactions suppress melanoma lung metastasis. *Journal of Experimental Medicine*, **188**, 1717–1723.

Paget, S. (1889). The distribution of secondary growths in cancer of the breast. *Lancet*, **1**, 571–573.

Radinsky, R. and Fidler, I. J. (1992). Regulation of tumour cell growth at organ-specific metastases. *In Vivo*, **6**, 325–331.

Rofstad, E. K. and Danielsen, T. (1999). Hypoxia-induced metastasis of human melanoma cells: involvement of vascular endothelial growth factor-mediated angiogenesis. *British Journal of Cancer*, **80**, 1697–1707.

Rubio, N., *et al*. (2001). Metastatic behavior of human breast carcinomas overexpressing the bcl-x(1) gene: a role in dormancy and organospecificity. *Laboratory Investigation*, **81**, 725–734.

Shtivelman, E. and Namikawa, R. (1995). Species-specific metastasis of human tumor cells in the severe combined immunodeficiency mouse engrafted with human tissue. *Proceedings of the National Academy of Sciences of the USA*, **92**, 4661–4665.

Welch, D. R. (1997). Technical considerations for studying cancer metastasis *in vivo*. *Clinical and Experimental Metastasis*, **15**, 272–306.

Xie, X., *et al*. (1992). Comparative studies between nude and scid mice on the growth and metastatic behavior of xenografted human tumors. *Clinical and Experimental Metastasis*, **10**, 201–210.

Young, S. D. and Hill, R. P. (1986). Dynamic heterogeneity: isolation of murine tumor cell populations enriched for metastatic variants and quantitation of the unstable expression of the phenotype. *Clinical and Experimental Metastasis*, **4**, 153–176.

Young, S. D., *et al*. (1988). Hypoxia induces DNA overreplication and enhances the metastatic potential of murine tumor cells. *Proceedings of the National Academy of Sciences of the USA*, **85**, 9533–9537.

FURTHER READING

Chambers, A. F. and Hill, R. P. (1998). Tumor progression and metastasis. In: Tannock, I. F. and Hill, R. P. (eds), *The Basic Science of Oncology*, 3rd edn (McGraw-Hill, New York).

Chambers, A. F. (1999). The metastatic process: basic research and clinical implications. *Oncology Research*, **11**, 161–168.

Hill, R. P., *et al*. (1986). Metastatic cell phenotypes: quantitative studies using the experimental metastasis assay. *Cancer Reviews*, **5**, 118–151.

Weiss, L. (1990). Metastatic inefficiency. *Advances in Cancer Research*, **54**, 159–211.

Welch, D. R. (1997). Technical considerations for studying cancer metastasis *in vivo*. *Clinical and Experimental Metastasis*, **15**, 272–306.

Chapter 65

Models for Tumour Cell–Endothelial Cell Interactions

Jörg Haier and Garth L. Nicolson
The Institute for Molecular Medicine, Huntington Beach, CA, USA

CONTENTS

- Biological Basics
- Principles of Model Establishment
- General Applications
- Interpretation of Experimental Results
- Biological and Technical Limitations
- Perspectives

BIOLOGICAL BASICS

The endothelial lining is distributed throughout the body, represents an organ that serves multifunctional purposes and is of approximately 1.5 kg tissue weight in an adult. The primary function of this organ is the preservation of a morphological barrier between the blood circulation and other organs and tissues. The structure is made up of a continuous but semipermeable cell layer throughout the vascular tree as well as an underlying subendothelial basement membrane. Furthermore, endothelial cells (EC) play an essential role in numerous cellular interactions, such as those involved in haemostasis, inflammation, wound healing and tissue regeneration. EC interact with platelets and leucocytes in the bloodstream and also cells from surrounding tissues, such as fibroblasts, macrophages and other immunocompetent cells. Moreover, EC take part in various haemostatic and metabolic functions, including tissue nutritional support and gas exchange.

In addition to the role of EC in angiogenesis and vascular support of primary tumours, the most obvious role of EC is in the formation of distant metastases where it is the primary cellular contact between circulating blood-borne tumour cells and host organs (Auerbach, 1992). The vascular endothelium is not a passive anatomical structure during the blood-borne spread of malignant tumours. Although the locations of many regional metastases can be explained on the basis of anatomical and/or mechanical factors, clinical and experimental observations indicate that many malignant tumours preferentially metastasize to particular distant organ sites that would not be predicted solely from mechano-anatomical considerations and blood flow patterns (Hart, 1982; Nicolson, 1982). Some organs and tissues, such as bones, adrenals and brain, are served by a small fraction of the circulating blood volume, but they are often involved in metastases from certain primary malignancies, whereas others with a high fraction of blood supply, such as heart, muscles, kidneys and others, are only sporadically colonized by cancer metastases (Sugarbaker, 1981; Nicolson, 1988).

For some time now, two major hypotheses on the organ preference of tumour metastasis have existed. Ewing (1928) proposed that the mechanical lodgment of circulating tumour cells in the first capillary system encountered was the determining factor, and that the haemodynamics of the vascular system based on specific anatomical structures predicted the locations of secondary tumours. In contrast, Paget (1889) postulated in his ‘seed and soil’ hypothesis that successful interactions of tumour cells (‘seeds’) with the microenvironment of a particular target organ (‘soil’) lead to formation of distant metastases in specific organs. These hypotheses are not mutually exclusive, and the development of distant metastases at specific organ sites is probably determined by anatomical structures with their resulting haemodynamics as well as specific interactions between unique circulating tumour cells and host organ structures.

The primary barrier between the circulation and extravascular tissues consists of the vascular endothelium and its subendothelial basement membrane (Nicolson, 1989). For the maintenance of normal function, specific interactions are required between circulating cells and these structures, and comparable interactions between tumour cells and organ EC may be necessary for organ-specific formation of secondary tumours (Weiss *et al.*, 1988). For example, specific parts of the microcirculation known as high endothelial venules (HEVs) are specialized in their interactions with other cells, such as lymphocytes, and this may allow transmigration of lymphocytes between the

bloodstream and lymphatics across the vascular barrier (Ford *et al*., 1976). The specialized HEVs, which are morphologically and antigenetically distinct from other EC and also express different antigenic patterns between HEVs from different organs, can be specifically recognized by circulating lymphocytes. Specific molecules are used by lymphocytes to initiate adhesion and migration into particular organs (Gallatin *et al*., 1983). Similarly, malignant tumour cells that preferentially colonize certain host organs must adhere to and invade in a similar way through endothelium and subendothelial basement membranes during haematogenic metastasis formation. Alternatively, for some blood-borne tumour cells it is known that they can adhere to microvascular EC and proliferate within the blood vessels (Belloni and Tressler, 1990) (**Figure 1**).

The endothelium is a single EC layer of considerable biological significance that lines the interior of blood/lymph vessels and the heart. Many of its complex functions appear to involve specific structural and biochemical characteristics of the EC surface. These specific molecular structures are found as permanent or transient, depending

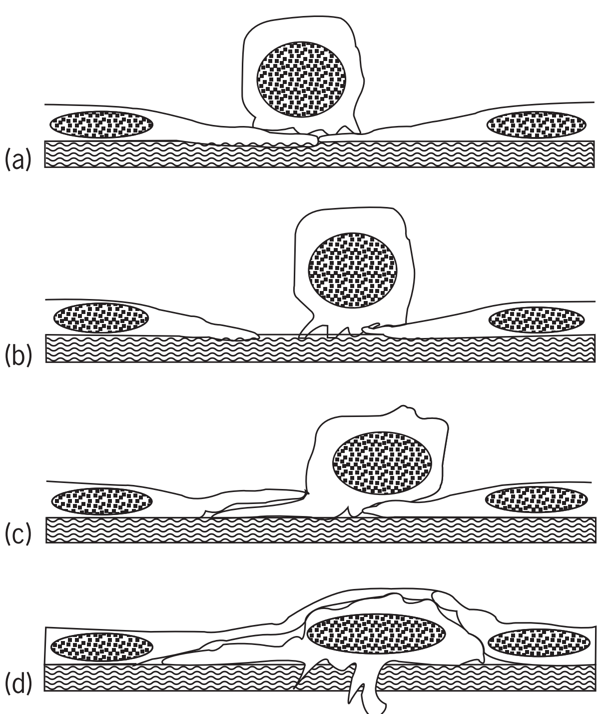


Figure 1 Sequence of events during metastatic tumour cell interactions with EC monolayer. (a) Tumour cell attachment to EC; (b) disruption of EC-EC contacts and retraction; (c) active migration of the tumour cell under the EC in contact with the underlying basement membrane; (d) invasion through basement membrane into host organ and reformation of intercellular EC junctions. (For further steps of invasion see the previous chapter, *Tumour Metastasis Models*.)

on EC functional demands and organ type (Pauli *et al*., 1990). The microcirculation is characterized by its own specific vascular structures, including differences in vessel length, diameter, wall thickness, receptors and ultrastructure that vary remarkably between different tissues and organs (Manjo, 1965; Rhoden, 1980). For example, the continuous endothelium can be either thick or thin depending on tissue and the type of EC. For example, continuous thick capillaries (thickness $>2\,\mu m$) can be found in skeletal tissue, cardiac smooth muscle and genital organs. In contrast, continuous thin capillaries (thickness $<1\,\mu m$) are typically located in tissues of the central nervous system, lung, dermis, vasa recti of the kidney, spleen, thymus, bone marrow and bone. Morphological studies *in vivo* and in cell cultures have shown that microvascular EC differ structurally according to their microvascular origin. EC from thin capillaries usually grow in culture as flat cells of hexagonal shape with large amounts of membrane-related cytoplasmic and pinocytotic vesicles on their apical and basolateral surfaces. Various EC phenotypes can be characterized according to their shape, ranging from cuboidal to elongated and stellate, and contents of vacuoles and shapes of nuclei.

The endothelium maintains its integrity and continuity through communication via gap junctions and desmosomes. The exception is in the central nervous system where tight junctions are the predominant structures. In addition to homotypic contacts, EC also have multiple points of attachment to the subendothelial basement membrane that contains various matrix proteins, such as laminin, proteoglycans (predominantly heparan sulfate proteoglycans), fibronectins, vitronectin and collagens (primarily types III, IV and V). Since variations can be found in the distribution of extracellular matrix (ECM) components within the basement membranes of different organs, such variations likely play a role in vascular heterogeneity and in the regulation of organ haemostasis and growth (Carley *et al*., 1988).

Although most EC are morphologically similar, there are distinctive regional molecular and structural differences in the vascular endothelium (Fenyves *et al*., 1993) For example, the vesiculo-vacuolar organelles, recently described organelles found only in the cytoplasm of EC that line tumour microvessels and normal venules, span the entire thickness of the vascular endothelium. These organelles provide potential trans-endothelial connections between the vascular lumen and the extravascular space. Weibel–Palade bodies, which are specific intracellular storage organelles for von Willebrand factor (vWF) in large vessel endothelia, are rarely found or absent in EC derived from the central nervous system. Moreover, differences in size and thickness of EC have been found in different organs or tissues. Generally, EC from large vessels appear to be thicker than those from microvascular endothelium. Additionally, variable surface structures that include adhesion molecules for homotypic cell–cell

interactions and for anchorage to the subendothelial basement membrane as well as other molecules, such as lectins, can be differentially expressed in large and small vessels (Simionescu *et al*., 1982; Wang *et al*., 1985).

EC in various tissues are functionally and metabolically distinct as well as structurally dissimilar (Schnitzer et al., 1994). They can participate to various degrees in metabolic and haemostatic functions, such as the following: (1) maintenance of a non-thrombogenic luminal surface, in part, via production of prostacyclins; (2) regulation of the coagulation pathway, clot formation and fibrinolysis; (3) presentation of antigens to lymphoid cells; (4) secretion of chemotactic monokines; (5) active participation in leucocyte adhesion and extravasation; (6) secretion of angiogenic factors and other growth factors that coordinately stimulate neoangiogenesis; (7) clearance and metabolism of vasoactive substances, such as bradykinin, endothelin, ADP, angiotensin I and thrombin; (8) control of vasomotion; (9) regulation of perfusion and permeability at the microvascular level and responses to and secretion of autocoids such as nitric oxide, prostaglandins, epinephrine and vasopressin; and (10) regulation of macromolecular transport between plasma and interstitial fluids via large pores, vesicular transport and specific cell surface receptors. Although the highly regulated endothelial functions associated with haemostasis and metabolism of vasoactive substances undoubtedly contribute to the metastatic process of certain tumours, we refer the reader to other reviews for more detailed information about these topics (Gerritsen, 1987; Weiss *et al*., 1988; Belloni and Tressler, 1990).

Transcapillary movements of solutes, such as hormones, drugs, sugars, amino acids, lipids and electrolytes, are regulated by the endothelium. The mechanisms responsible for the regulation of transport are (1) EC microdomains that attract or exclude molecules based on their size and charge, (2) EC surface receptor-specific membrane transport, (3) EC surface ectoenzymes that chemically alter solutes transferred into, out of or between EC and (4) EC synthesis and secretion of solutes that establish concentration gradients between blood plasma and the interstitial tissues. Thus transcapillary transport regulates vascular tone and organ blood supply as well as lymphatic flow and expression of surface receptors for the activation of leucocytes, such as during inflammation. In different organs it has additional organ-specific functions, such as regulation of the cerebral endothelial lining/blood–brain barrier and endothelium-mediated changes in renal, splenic and hepatic function and in skeletal muscle perfusion (Bassenge, 1996).

The microvascular EC membrane has been described as a static pore in some older vascular EC permeability models, with regulation of solute and fluid flow by changes of hydrostatic and hydrodynamic pressure and surface area. Although these models may be suitable for basic conditions, they cannot explain the enormous changes in secretion that occur after trauma or during inflammation. During the last two decades, the regulation of microvascular porosity has been investigated using various inflammatory mediators, vasoactive substances and intracellular signal inhibitors (Bevilacqua *et al*., 1985, 1986; Guretzki *et al*., 1994). These studies led to the development of a new concept of variable mediator-regulated pathways that control endothelial permeability (Lum and Malik, 1994). Investigation of the regulation of the permeability properties of the endothelium has yielded evidence that supports the concept of a dual regulation of EC gap formation and barrier function.

Tumours can dramatically alter the permeability and transport properties of blood vessels (Zhang and Olsson, 1997). For example, in brain tumours a common problem is vasogenic oedema that can cause subsequent neurological symptoms. The peritumoural brain parenchyma shows structural and functional changes of the intracerebral microvessels and oedema (Groothius and Vick, 1982). Additionally, the EC of peritumoral microvessels in brain tumours express glucose transporter protein in the same way as the normal tissue. However, a reduction in immunostaining for the glucose transporter occurred in the microvessels located within metastases but not in surrounding tissues. This indicates abnormalities of the blood–brain barrier in tumour vessels but normal barrier function in peritumoural regions. These effects might be related to the tumour-induced production of vasoactive substances. Reactive astrocytes and activated microglial cells produce numerous biologically active compounds, including endothelin-1, which can influence the structure and function of the peritumoural brain tissue. Structural alterations are also present in tumour-induced blood vessels and peritumoural brain microvessels, such as increased numbers of pinocytotic vesicles, large vacuoles and open junctions, which may also cause changes in permeability. Other possible mechanisms might be related to the destruction of microvessel structure and subsequent changes in hydrostatic pressure (Roy and Chitra, 1989).

One of the interesting tumour properties associated with organ-specific metastasis is that tumour cells may respond to local concentrations of growth factors produced by or present in the target organs. A likely source of organ-associated tumour cell motility factors is organ microvascular EC (Sawada *et al*., 1996) The EC-produced motility factors might interact with other soluble tissue factors to promote organ-specific tumour cell invasion (Orr *et al*., 1980; Hujanen and Terranova, 1985). Examples are nerve growth factors and neurotrophins that play an important role in migration through the blood–brain barrier (Menter *et al*., 1995). EC are also known to produce specific motility-enhancing components of the basement membrane, such as fibronectin, laminin, collagen IV, thrombospondin and fragments derived from these molecules. As described above, EC from various organs possess different amounts and proportions of subendothelial matrix

components. Therefore, tumour cell chemotaxis may be specifically influenced by the composition of the subendothelial basement membrane in the target organ (Cerra and Nathanson, 1991).

The response to various systemic and paracrine growth factors and their inhibitors determines the growth of malignant cells at particular organ sites (**Figure 2**). Differentially expressed stimulating and inhibitory factors influence tumour growth at particular organs and their ability to synthesize and secrete autocrine growth factors. These paracrine growth factors and inhibitors are produced and secreted specifically in each organ, and they appear to play an important role in organ preference of tumour cell proliferation and development of metastases. Some of the paracrine growth factors from certain tissues have been identified. For example, a lung-derived growth factor that stimulated mammary adenocarcinoma cells was purified from porcine lung-conditioned medium and later sequenced and identified as a transferrin (Cavanaugh *et al.*, 1991). Transferrin and transferrin-like factors are produced by microvessel EC, and differences have been demonstrated in the secreted amounts with the highest amounts being produced by brain EC. Liver and lung EC produced lesser amounts of transferrin-like factors (Inoue *et al.*, 1993).

Since some tumour cell lines showed growth inhibition during experiments using EC-conditioned medium, the release of inhibitory factors was assumed. Growth inhibition correlated with the metastatic behaviour of these cell lines (Nicolson and Dulski, 1986). Using organ-conditioned medium, various types of growth inhibition have been found to be mediated by the secretion of growth-inhibitory paracrine factors. One of these inhibition factors, which was first found in kidney-conditioned medium, was purified and identified as transforming growth factor-β (TGF-β) (Holley et al., 1980). TGF-β and other cytokines that are released from EC, such as interleukin-6 (IL-6), might have dual roles in the growth regulation of malignant cells (**Figure 3**). IL-6 has potent stimulatory effects in the activation of EC and enhancement of tumour cell adhesion, but it was also found that this cytokine decreased the cell growth of some tumour cell lines. Other sources of growth-inhibitory potential are immune-competent cells that are often found in the tumour

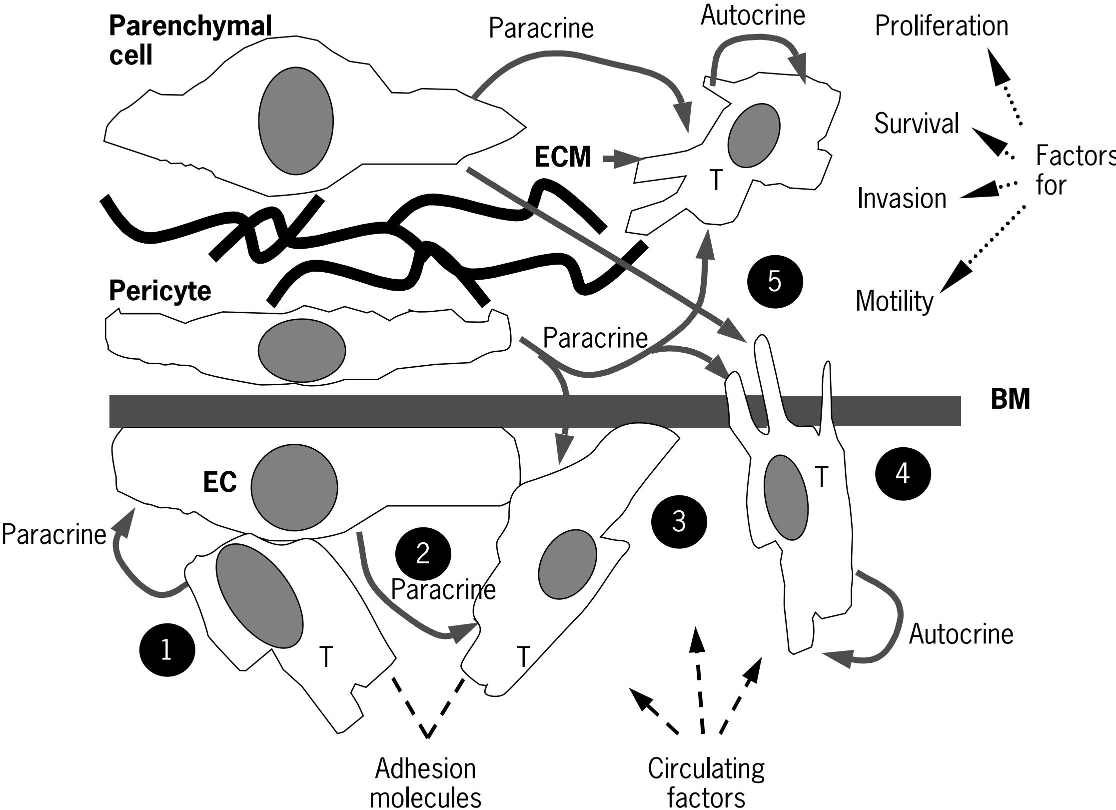


Figure 2 Paracrine and autocrine factors in tumour cell–host organ interactions during various steps of invasion and metastasis. (1) Tumour cell adhesion to EC; (2) adhesion-mediated retraction of EC with exposure of subendothelial basement membrane; (3) adhesion of tumour cells to basement membrane; (4) basement membrane degradation and migration of tumour cells into host organ; (5) survival and growth of tumour cells in host organ. EC, endothelial cell; T, tumour cells; ECM, extracellular matrix; BM, basement membrane; full lines, autocrine and paracrine factor interactions; dashed lines, cofactors that can interact with autocrine and paracrine signalling and modify cellular reactions.

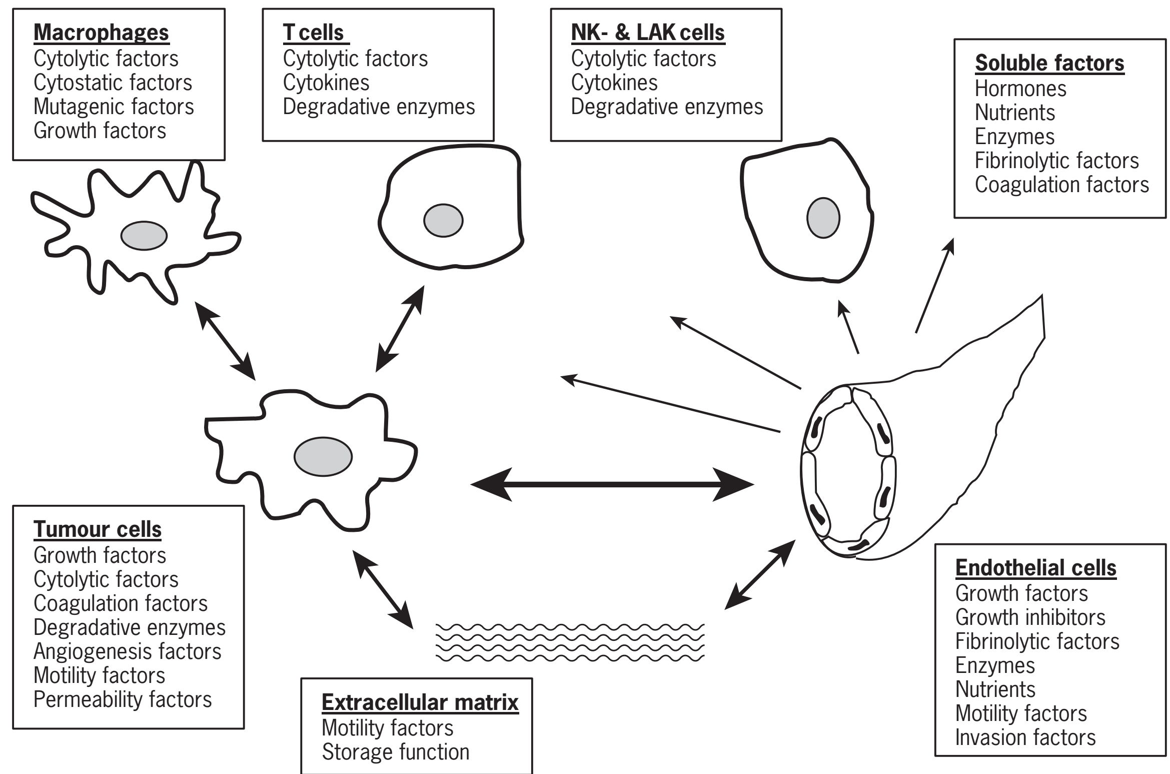


Figure 3 Bidirectional or reciprocal tumour cell–EC interactions. The cellular microenvironment is influenced by factors released from tumour cells, EC, parenchymal cells or invading leucocytes. These factors can affect tumour cell survival, proliferation, motility, invasiveness and metastatic properties.

microenvironment, such as tumour-infiltrating lymphocytes, natural killer cells and macrophages. For example, these cells as well as EC produce activated oxygen derivatives, such as oxygen free radicals (O_2^-). These radicals are rapidly converted into highly reactive radicals (H_2O_2, $HO^·$) that are cytotoxic or cytostatic.

Another mechanism has been proposed for cytostatic or cytotoxic effects of EC on circulating tumour cells (LaBiche and Nicolson, 1993). To examine whether vascular EC can be treated with various cytokines to become cytotoxic against tumorigenic target cells, mouse EC monolayers were treated with interferon-γ (IFN-γ) or tumour necrosis factor-α (TNF-α). The activated EC but not untreated control cells produced significant cell lysis of mouse reticulum sarcoma cells and two different mouse melanoma cell lines. This effect was dependent on the presence of the endothelial cells because treatment of the malignant cells with IFN-γ and/or TNF-α did not produce direct target cell lysis. Continuous presence of the cytokines was required for the activated EC-mediated tumour cell lysis (Li *et al*., 1991). However, this result was subsequently found not to be due to direct tumour cell cytolysis by activated EC. Although EC were able to inhibit the growth of tumour cells, they were not able to lyse tumour cells (Nicolson, 1999).

PRINCIPLES OF MODEL ESTABLISHMENT

Isolation and Cultivation

The data discussed above indicate that microvascular EC and ECM from different organs possess unique characteristics that are probably related to their different functions. These characteristics are dynamic and dependent on the functional requirements and microenvironmental conditions of the endothelium. First EC cultures were obtained from relatively accessible sources. For example, human umbilical vein is a useful source of EC (HUVEC) because it is easy to obtain, cells can be isolated in quantity and the cells are generally pure and grow well. However, it has the disadvantage that it is not usually a target site for metastatic involvement. Other blood vessels and tissue sources have also been used to obtain EC. For example, EC have been prepared from nasal mucosa, bovine capillary, pancreatic islet, mouse haemangioendothelioma, adrenal cortex, human dermis, bovine aorta or human neonatal foreskin. In contrast, the culture of tumour-associated EC continues to be difficult, because microvascular EC change during prolonged cultivation and growth conditions, and cellular properties can vary with their particular origin.

Cultivation of HUVEC is well established, and such cultures are widely used. Many studies on the function and interactions of EC have been performed using HUVEC cultures. However, in the evaluation of these results consideration should be given to the neonatal and macrovascular large vessel origins of these cells. As discussed above, this type of EC differs from microvascular EC. For example, the reactions of HUVEC to growth factors (ECGF, EGF, TGF-α, thrombin) and other cytokines (interleukins, TGF-β) were different from those found with microvascular EC. Additionally, differences were found between the release of cytokines by HUVEC and adult macrovascular EC. Also, organ-specific patterns of adhesion molecules expressed on microvascular EC were demonstrated *in vivo* and *in vitro*. The morphology and antigen expression of EC can also be influenced by microenvironmental conditions, such as infiltrating tumour cells. In conclusion, the use of microvascular EC cultures derived from specific target tissues for models of metastasis appears to reflect *in vivo* conditions better than EC from other sources.

The isolation of EC is generally performed using collagenase digestion of ECM and mass gradient selective centrifugation followed by several washing steps. Once cells are isolated, their identity as EC has to be confirmed by their characteristic morphology and expression of EC-specific markers, such as factor VIII, CD31 or vWF. Functional criteria, such as uptake of acetylated LDL, can also be used to confirm cell identity in EC cultures.

Activation of EC

The functions expressed by EC are strongly influenced by the extracellular environment. The presence of interleukins, interferons, tumour necrosis factor, lipopolysaccharides (LPS) and other mediators of inflammation can lead to changes in the expression of adhesion molecules on EC surfaces as well as on other cells (macrophages, monocytes, lymphocytes, etc.). This activation of EC also generates various functional changes. For example, soluble substances, such as TNF-α, IL-1, IL-4, IL-13, IFN-γ or TGF-β, initiate expression and enhanced affinity of homing receptors, such as integrins or selectins for neutrophils which results in increased adhesive interactions between these cells and EC. TGF-β, however, plays a dual role and inhibited E-selectin expression on HUVEC at the mRNA and protein levels. The basal adhesion of various tumour cells to unstimulated EC was very low, and cytokine stimulation augmented these interactions by upregulation of E-selectin expression on the EC surface. Other factors, such as plasminogen activator, released from normal or neoplastic cells cause retraction of EC with subsequent exposure of the underlying ECM. Thus, chemotactic reactions and transendothelial migration of neutrophils are possible. Furthermore, haematopoietic growth factors take part in these complex activation and inhibition processes. For example, granulocyte macrophage colony-stimulating factor (GM-CSF) and granulocyte colony-stimulating factor (G-CSF) can modulate integrin-mediated adhesion of neutrophils to activated EC.

The microvasculature is susceptible to perturbation from environmental agents, host cells and cancer cells. There is clinical and experimental evidence that physical trauma, irradiation, chemotherapeutic agents or environmental toxins can induce or enhance the metastatic process. Using established animal models of pulmonary microvasculature injury it has been shown that endothelial damage promotes the localization of circulating cancer cells to the lung with increased development of metastatic lesions, and this effect is lost after endothelial repair. These models generally use the pharmacological or chemical induction of EC damage, such as application or release of free radicals. In this example, oxidative stress is an effector of vascular damage in several of the experimental models. While endothelial cells appear to be directly susceptible to free radical attack, basement membranes are much less sensitive. However, oxidative injury of EC causes release of degradative enzymes with subsequent degradation of ECM components. This event is associated with the generation of tumour cell chemoattractants and enhanced cancer cell invasion of vascular basement membranes.

Adhesion Molecules

For investigations on the adhesive properties of various malignant cell lines, EC monolayers from human umbilical vein (HUVEC) or bovine aorta (BAEC) have been widely used. Although useful for general adhesion studies, differences were usually not observed for metastatic and nonmetastatic tumour cell adhesion to large-vessel EC. The organ preference of metastasis did correlate, however, with the adhesion of tumour cells to microvascular EC in many but not all experimental systems. For example, high lung-colonizing rat rhabdomyosarcoma, melanoma and lymphoma cell lines showed high rates of adhesion to lung-derived EC monolayers, whereas cell clones of poor metastatic potential had low adhesion rates. Metastatic tumour cells adhered to target organ EC and stimulated their retraction at cell junctions adjacent to the adherent tumour cells. Subsequently, the tumour cells penetrated the EC monolayer, adhered to components of the underlying basement membrane and eventually penetrated the subendothelial matrix. This behaviour was normally not found in most of the noninvasive, nonmalignant cells examined. Although different metastatic systems may vary in their adhesive properties to EC and their underlying matrix, their adhesive behaviours are often related to their metastatic properties (Belloni and Nicolson, 1988).

In general, the subendothelial matrix was found to be a better substrate for cell adhesion of most tumour cells than the EC surface (see the previous chapter, *Tumour*

Metastasis Models). Also, the adhesive behaviours of tumour cells could be modified by the organ-specific composition of the subendothelial ECM. Using bovine aortic EC grown on ECM preparations from various organs tumour cells showed adhesion to EC that was dependent on the organ matrix. Thus, EC behaviour can be modified by ECM composition, possibly due to changes in expression of adhesion and other molecules.

Some of the organ-associated EC surface determinants involved in the adhesion of metastatic tumour cells have been identified using affinity procedures (**Figure 4**). For example, ^{125}I-labelled lysates of EC were adsorbed on lightly glutaraldehyde-fixed tumour cells. After binding, the components were solubilized in detergent and analysed electrophoretically. In these experiments, the majority of the adsorbed radiolabelled proteins migrated in the molecular mass region known to contain integrins (~120–180 kDa). However, there were also low molecular mass proteins bound to the fixed tumour cells corresponding to the molecular masses of known glycoproteins, such as annexins, lectins and other receptors. Moreover, it was found that the individual protein profiles were unique for each EC type (Belloni and Tressler, 1990).

A similar technique for identification of adhesion molecules on EC surfaces was developed (tumour cell–Western adhesion) where detergent-solubilized EC preparations were electrophoretically separated, and EC-binding components were detected by adding viable biotinylated or radiolabelled tumour cells. For example, using this method five EC-derived glycoproteins were identified from liver or lung microvessel EC that interacted with murine large-cell lymphoma cells. These lymphoma cells bound preferentially to hepatic EC, and they adhered quantitatively based on their organ-specific metastatic behaviour. These results suggest that tumour cell adhesion is directed, in part, to specific tissue-associated surface molecules expressed by the microvascular EC in different organs. Although these adhesion molecules appear to be involved in tumour cell interactions with microvascular EC, it is unlikely that these molecules alone can confer adhesive specificity to metastatic tumour cells (Nicolson, 1988; Belloni and Tressler, 1990).

Another approach for studies of EC surface molecules that are involved in interactions with tumour cells was the development of serological procedures, including monoclonal antibodies against organ-associated antigens on EC (**Figure 5**). Hybridomas were selected from animals immunized with EC from various organs and different species. Screening and further selection using binding studies of the hybridomas and cells from different organs or species resulted in monoclonal antibodies directed against but not limited to organ-specific antigens. Using this method, antibodies reactive against EC and other cells, such as fibroblasts from one organ but not reactive against cells derived from other organs, were obtained (Doer *et al*., 1989).

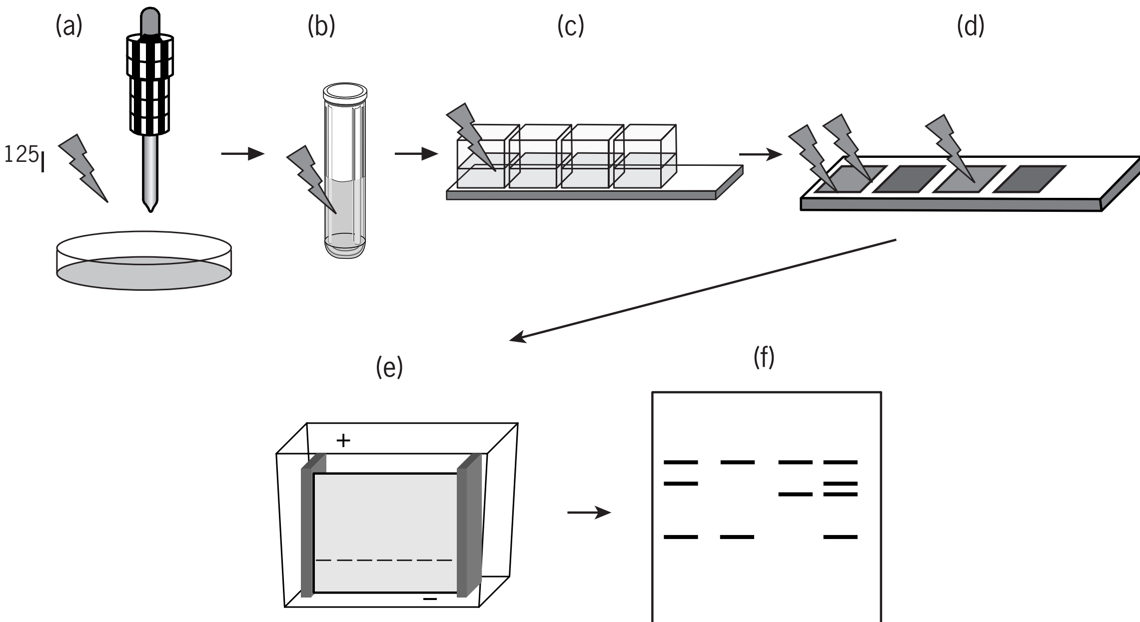


Figure 4 Principle of identification of tumour cell-binding proteins on microvascular EC surfaces. (a) Labelling (^{125}I) of microvascular EC surface molecules; (b) solubilization using sodium dodecyl sulfate or CHAPS detergent; (c) adsorption of labelled lysate to glutaraldehyde-fixed tumour cells; (d) bound EC adhesion molecules occur as radioactive sources on tumour cells; (e) electrophoretic separation of bound adhesion molecules; and (f) visualization of EC adhesion molecules by autoradiography. Microvascular EC cell lysates from different organs showed different patterns of adhesion molecule expression.

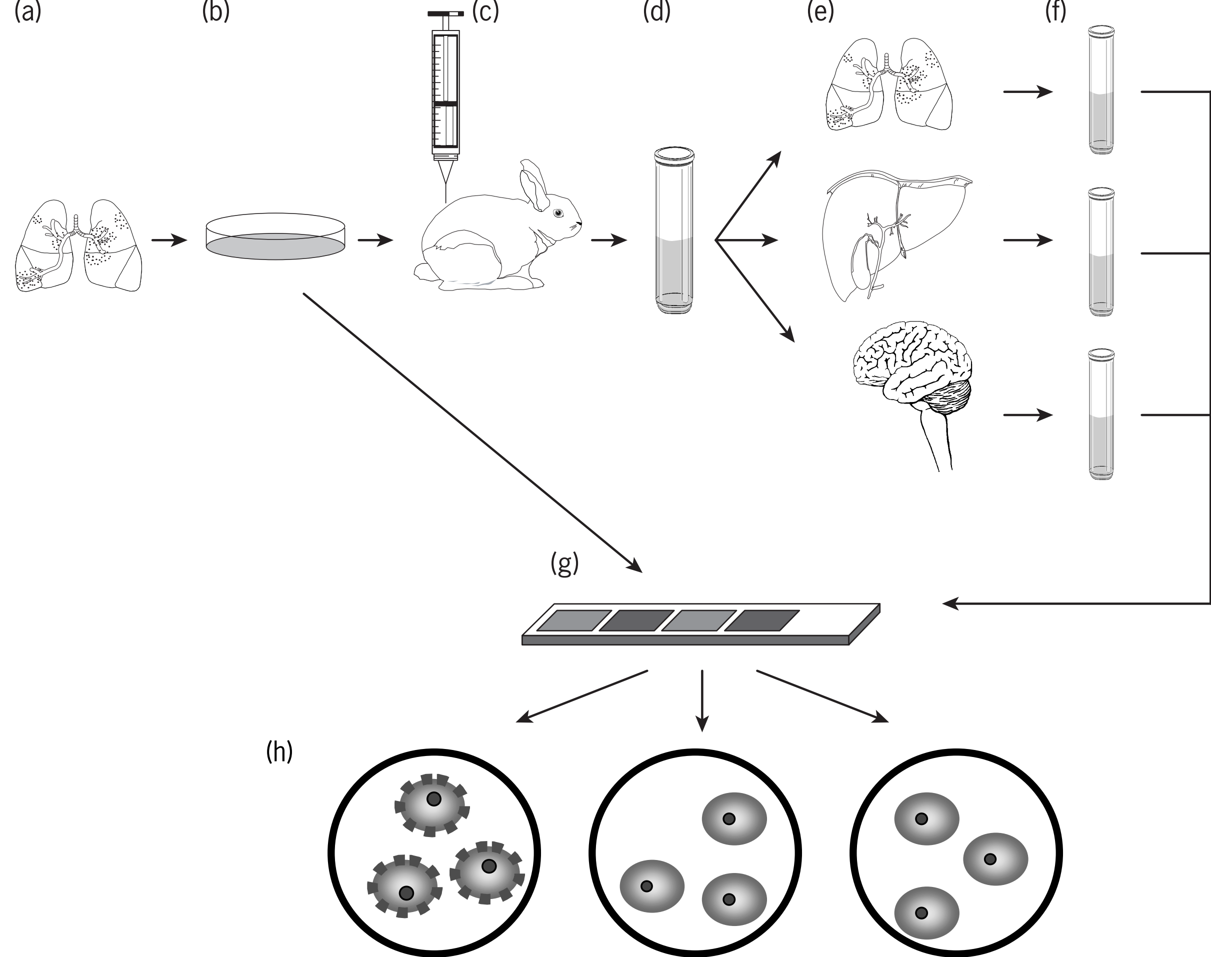


Figure 5 Selection of organ-specific antibodies against EC surface molecules. (a) Isolation of organ-specific EC (e.g. from lung); (b) culture of organ-specific EC; (c) immunization of an animal with organ-specific EC; (d) isolation of antibodies against EC; these polyclonal sera contain antibodies against organ-specific, species-specific and nonspecific EC surface antigens; (e) selection of organ-specific antibodies against EC from various organs; (f) remaining preparations contain organ-specific but not nonspecific antibodies; (g) EC monolayers from target organ can be used for immunocytochemical analysis; (h) organ-specific antibodies react with EC derived from target organ, but with EC from other organs.

Immunohistological studies using polyclonal antisera against various EC-binding glycoproteins have shown that specific organ distribution of EC adhesion molecules is quantitatively not qualitatively different in most tissues. At least one of these glycoproteins was identified as a galectin (∼33 kDa) (Lotan *et al*., 1994). Other components were identified as annexin II (∼35 kDa). Antibodies against these adhesion molecules inhibited adhesion of tumour cells to microvascular EC monolayers by about 40% (Tressler *et al*., 1993), suggesting that multiple adhesion systems (integrins, selectins, annexins) are involved in tumour cell–EC adhesive interactions.

Functional studies on adhesive interactions of EC with circulating cells can be performed using static or dynamic adhesion systems. The principles of these techniques are described in the chapter *Models for Tumour Cell Adhesion and Invasion*, and these procedures can be adapted with the modification that EC monolayers are the adhesive substrates. Since EC–tumour cell interactions appeared to be mediated primarily by glycoproteins where the binding takes place between the lectin domain of one molecule with a corresponding carbohydrate, different types of carbohydrates were assumed to be important in determining organ-specific adhesive interactions. Therefore, modification of terminal sugar moieties or blocking using antibodies against specific carbohydrates were used for identification of the involved lectins. Thus an important group of adhesion molecules identified by these methods were a type of lectins called selectins. Selectins are adhesion molecules found on tumour cells and normal cells that use carbohydrates as receptor ligands, such as leucocytes/lymphocytes (L-selectin), platelets (P-selectin) and

EC (E-selectin). The action or expression of EC selectins depends on cell activation by ILs, TNF or toxins. For example, selectins mediate the recognition of carbohydrates via Sialyl-Lewis[a], Sialyl-Lewis[x] or the MECA-70 antigen and they play a central role in targeting and initial contacts between circulating tumour cells and EC of the host organ (Irimura *et al*., 1986). (See also the chapter *Models for Tumour Cell Adhesion and Invasion*.)

The presence of platelets in adhesive interactions can also result in enhancement in EC adhesion, and their involvement in tumour cell targeting to microvascular vessel walls has been demonstrated. Platelets take part in different steps of the adhesion cascade, but they appear to be especially important in stabilization of tumour cell–EC and tumour cell–ECM adhesion. Specific interactions between tumour cells and platelets are mediated by release of activating substances and cross-linking of platelets during adhesion of tumour cells to EC and ECM.

Paracrine Motility Factors

Almost two decades ago studies showed that tumour cells can preferentially grow in medium conditioned by the target organs for metastatic colonization. For example, the ability of liver- and brain-metastasizing tumour cells to migrate towards a concentration gradient of soluble components extracted from brain or liver was demonstrated. Liver-metastasizing murine colon carcinoma cells preferentially migrated towards a concentration gradient of liver extract rather than to gradients made from extracts of lung or brain (Sawada *et al*., 1996).

Since differences in the motility and growth responsiveness to soluble factors derived from mouse liver, brain, and lung microvessel EC were found, conditioned media were used to examine the effects of secreted factors on various tumour cell systems. For example, mouse liver hepatocyte sinusoidal EC (HSE) and microvascular lung EC (MLE) were used to demonstrate organ-specific motility factors that lead to specific enhancement of motility of colon carcinoma and large-cell lymphoma cells with different metastatic potentials (Hamada *et al*., 1992; Sawada *et al*., 1996) (**Figure 6**). Highly liver-metastatic cells showed a higher stimulation of motility after exposure to HSE-conditioned medium than poorly metastatic cells. Additionally, MLE-conditioned medium failed to stimulate the motility of these cells. In contrast, the growth properties of colon carcinoma cells were not altered after stimulation of poorly and highly metastatic cell lines with conditioned medium but highly metastatic lymphoma cells were preferentially growth stimulated by conditioned medium from liver EC (Nicolson, 1988). EC derived from lung and liver secrete different tumour cell paracrine motility factors, such as C3b complement factor or monocyte chemotactic factor-1 (MCP-1) that appears to

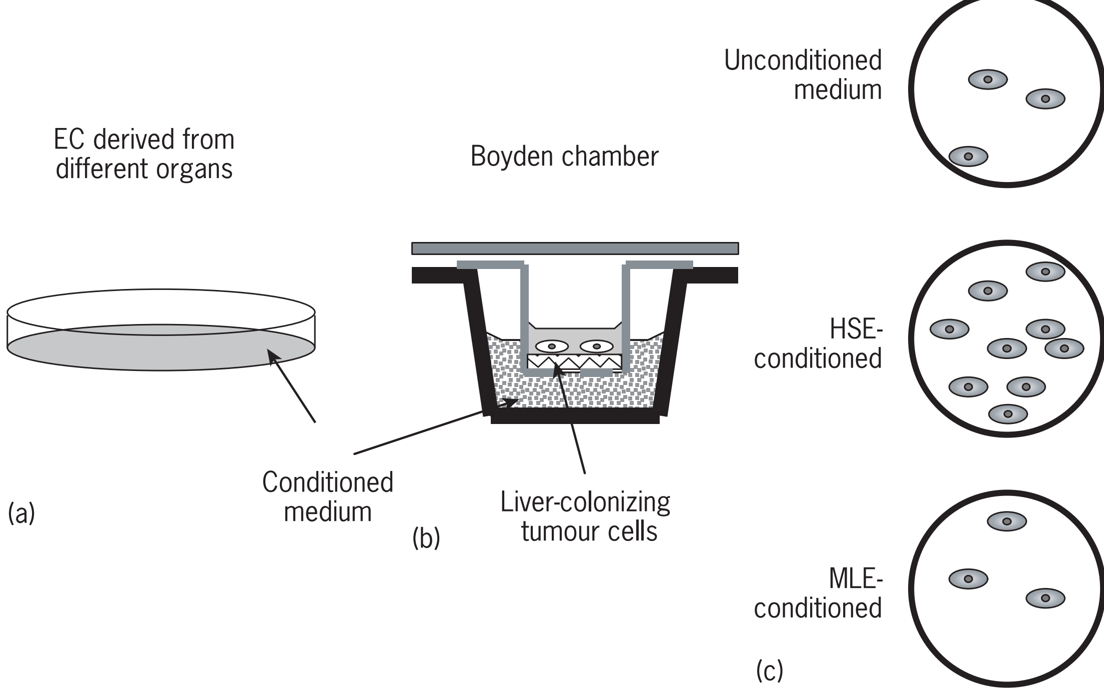


Figure 6 Paracrine motility factors secreted by microvascular EC enhance tumour cell migration. (a) Conditioned media are obtained from organ-specific microvascular EC cultures; (b) conditioned media are used as chemoattractants in the lower chamber of a Boyden chamber; (c) HSE medium from liver EC contains specific motility factor(s) and increases the number of migrated liver-metastasizing cells, whereas MLE medium from lung EC has no effect (see text).

contribute to organ preference of tumour cell invasion of these tissues by lymphoma cells (Hamada *et al*., 1992).

Various factors can influence specific steps of tumour cell adhesion, migration and invasion. Some of these have been examined using a two-chamber system separated by a porous barrier with chemoattractants in one chamber. The experimental conditions used in the Boyden chamber assay, such as various EC-conditioned media as chemoattractants, ECM components or gel preparations (as invasion barriers), determine the experimental outcome. For example, the effects of nerve growth factor (NGF) on the invasive potential of brain-metastatic melanoma cells through Matrigel (a mixture of basement membrane components that forms a gel layer in the Boyden upper chamber) were investigated during chemotactic stimulation using brain microvessel EC-conditioned medium. It was shown that the presence of brain EC motility factors was essential for invasion of the melanoma cells stimulated by NGF or other neurotrophins (Menter *et al*., 1995).

In a modification of the motility/invasion assays, possible pathways of action for these factors can be examined. For example, a possible pathway for tumour cell–EC and –ECM interactions is the stimulation of degradative enzyme secretion, such as collagenase and heparanase, and it was shown that melanoma cells exposed to NGF in the experiments described above are stimulated to release these degradative enzymes (Marchetti *et al*., 1993). Similar experimental designs can be used for the identification and evaluation of factors that determine the growth properties of tumour cells. It is known that EC can secrete growth stimulatory and inhibitory factors, and this can be obtained from the conditioned medium of organ-specific EC cultures. Tumour cell cultures can then be incubated with these conditioned media, and the number of cells or cell colonies can be measured using various detection methods.

Monolayer Wounding

A common application for the investigation of EC motility and functional consequences of disturbances of the integrity of EC monolayers is the use of wounded EC monolayers. The principle of this model is simple, and it is highly reliable. After EC are grown to confluence, the integrity of monolayers can be disrupted using sharp tools which are dragged across the monolayer. Subsequently the reconstruction of EC monolayer integrity can be observed under a microscope. Time-lapse microscopy can be applied to visualize time-dependent functions during the re-establishment of the EC monolayer integrity. This principle also allows the evaluation of cell functions which are required for EC monolayer re-establishment and the analysis of environmental conditions that are related to the presence of infiltrating tumour cells, such as paracrine growth and motility factors. Functional consequences can also be analysed by harvesting the cells for further biochemical investigations, such as examination of phosphorylation/dephosphorylation of cellular proteins.

Release of Degradative Enzymes

EC are a very important source of degradative enzymes required for the penetration of tumour and normal cells through the subendothelial basement membranes. Although tumour cells themselves are able to secrete various types of matrix degradative enzymes, they also can induce secretion and/or activation of various enzymes by surrounding EC. This stimulation occurs either directly through adhesive tumour cell–EC interactions or indirectly by release of soluble factors that modify the balance between degradative enzymes and their inhibitors to enable subendothelial basement membrane penetration by tumour cells. Models for investigation of degradative enzyme release in relation to adhesive interactions are described in the previous chapter, *Tumour Metastasis Models*. The analysis of soluble factors released by tumour cells that affect the balance between degradative activity and inhibition can be performed at the transcriptional level or by evaluation of conditioned medium. The first approach allows the investigation of surface receptors and subsequent signalling cascades that are involved in the mediation of enzyme release. EC monolayers can be incubated with different types of soluble factors, such as epidermal growth factor (EGF), platelet derived growth factor (PDGF) or various cytokines. Subsequent changes in mRNA levels of degradative enzymes and their inhibitors can be analysed, and specific drug-induced modifications of these changes allow the identification of candidate signalling mechanisms. (See also the chapters *Models for Tumour Cell Adhesion and Invasion* and *Modelling Tumour Tissue Interactions*.)

In contrast, conditioned media have been used to investigate post-translational events that are related to EC stimulation. EC monolayers can be stimulated with biochemicals (soluble factors), radiation or UV exposure or by physical interference (wounding), and the subsequent release of degradative enzymes into the culture medium can be examined at the protein level by immunological procedures, such as Western blotting. Since degradative enzymes are normally released as inactive pro-forms of the enzymes, detection of enzyme activity is required for further analysis. A common and very simple way to achieve semiquantitative results of released enzyme activity is provided by zymogram electrophoresis. This method uses different types of substrates that are added to polyacrylamide electrophoresis gels to visualize degradative activity by the disappearance of stainable substrates in the gel. Using zymogram electrophoresis the secretion of degradative enzyme can be characterized with regard to their molecular masses and substrate specificities. However, several experimental procedures, such as sodium dodecyl sulphate polyacrylamide gel electrophoresis, can

activate certain enzymes and activities observed do not necessarily represent EC function. Other experimental designs, however, allow investigators to overcome these limitations. For example, using the release of radioactively labelled heparan sulfate from ECM preparations after γ-irradiation of EC the secretion of heparanase was examined (Nicolson *et al.*, 1991). Most recently, fluorogenic substrates were introduced that allow measurements of enzyme activity, substrate specificity and kinetics. This method is based on the degradation of substrates where the resulting products are fluorogenic and can be detected using fluorescence spectrometers.

Barrier Function

The vascular endothelium is essential in maintaining proper haemostasis, selective barrier and permeability properties of blood vessels, and in mediating responses to various physiological and pathological stimuli. However, malignant tumour cells can influence EC barrier functions, including fluid exchange and selective passage of larger molecules. As an example of these changes, it is frequently seen that tumour masses are surrounded by oedema and the fluid exchange between blood circulation and tissue is disturbed (Criscuolo *et al.*, 1988; Zhang and Olsson, 1997).

Interactions with Coagulation System

Fibrin deposits in tumour beds are an unintriguing phenomenon. These deposits have been suggested to promote tumour growth by providing a provisional matrix in which tumour cells can grow, invade and induce the development of a vascular network that is crucial for tumour growth. It has been shown that procoagulant factors, such as thrombin, are not only involved in coagulation, but they can also stimulate tumour growth and neovascularization. Furthermore, the fibrin matrix can protect the newly formed micrometastases from host defence systems. Several mechanisms seem to contribute to peritumoural fibrin formation (Rickles *et al.*, 1992). For example, mechanical damage of EC integrity and subsequent exposure of the underlying subendothelial basement membrane can trigger fibrin formation by activation of the extrinsic coagulation cascade. Additionally, immunological and inflammatory stimuli in the tumour microenvironment can initiate responses of the intrinsic coagulation system. It has also been found that certain tumour cells are able to induce procoagulatory responses, and they can release proteases that activate various coagulation factors. These enzymes are serine or cysteine proteases that, for example, can cleave factor X.

The procoagulant response of EC alone or cocultured with tumour cells can be determined using conditioned media, as described above. Release or activation of different types of coagulation factors can be monitored using a variety of techniques that will not be explained in further detail in this chapter. Alternatively, fibrinolysis, or the cleavage of thrombin or plasminogen, can be investigated. The modification of plasminogen pathways by degradative activities is closely related to the degradative properties of tumour cells and/or EC (Ordinas *et al.*, 1990).

Tumour cells in addition to EC can be influenced by various circulating cells of the blood in different types of complex interactions. For example, platelets are frequently found in tumour or tumour-surrounding vessels. Their interactions with EC or tumour cells include reciprocal activation, stimulation of soluble factor release, such as growth factors or procoagulatory factors, and secretion of degradative enzymes (Honn *et al.*, 1992). For investigations of interactions between platelets, tumour cells and EC, coculture techniques have been widely used. Their principle is based on techniques where one cell type is grown to confluence (mostly EC), and other cells are added to the culture (usually tumour cells and/or platelets). Further analysis can be performed using microscopy for morphological analyses or immunological and/or biochemical methods for functional investigations. The evaluation of functional responses involves examination of signalling cascades within the different cell types and/or secretion of soluble factors into the medium. However, experiments involving several types of cells are often difficult to interpret, and slight changes in the experimental design or different cell preparations often produce different results.

GENERAL APPLICATIONS

One of the more interesting phenomena in cancer research is that tumours are collections of heterogeneous cells. Interestingly, EC are also heterogeneous in their properties to certain degrees. EC heterogeneity was apparent when immunohistochemical studies demonstrated that EC could be negative for expression of classical EC markers, such as factor VIII-related antigen von Willebrand factor. It was found that these negative EC were located in specific parts of the vascular tree, and they were functionally distinctive in sinusoidal organs, such as liver, lymph nodes or renal glomeruli. This EC heterogeneity was later also found in organ metastases. For example, EC in different tumours and sometimes within the same tumour demonstrate differential expression of surface molecules. Moreover, it has been shown that tumour cells demonstrate different adhesive interactions to EC derived from different organs, and these properties appear to be related to their metastatic behaviour.

The development of *in vitro* or *in vivo* models to study tumour cell–EC interactions should consider EC heterogeneity as well as their organ-specific microenvironments. Unfortunately, many if not all EC cell characteristics can be time dependent and transient. This includes the expression and function of various surface antigens and

secreted components and their complex interactions during the metastatic process. Furthermore, the endothelial origin of and response to various factors that stimulate or inhibit cell growth, motility and/or invasive potential must be considered. Moreover, autocrine, paracrine and hormonal mechanisms can occur transiently and change over time.

Currently, our knowledge about endothelium-dependent vascular responses, microcirculatory regulation and tissue regeneration is limited, and further investigations in these areas and their relevance to understanding pathological phenomena such as cancer invasion and metastasis are necessary. Functional responses and cellular signal transduction stimulation after tumour cell adhesion to EC are only partially understood, in addition to their downstream responses and ultimately changes in gene regulation. Exactly how EC promote lysis of their basement membrane and the subsequent penetration of tumour cells into the host organ is still under investigation, but these mechanisms appear to be very important in determining organ-specific formation of metastases.

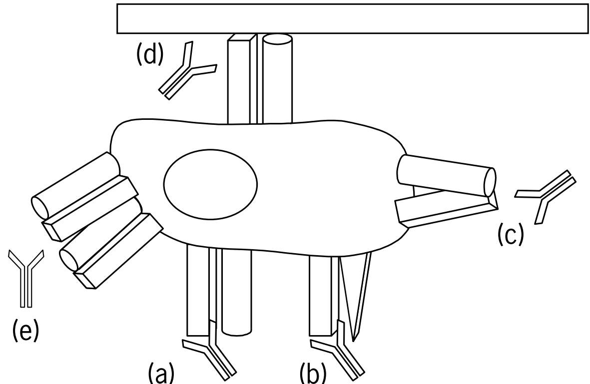


Figure 7 Detection of surface molecules by monoclonal antibodies. The scheme shows possible interference of various factors with the ability of a monoclonal antibody against one subunit of a heterodimeric molecule (for example integrins). (a) Antibody recognizes the target molecule; (b) antibody binds to a different heterodimer with the same subunit; (c) altered conformation of the molecule (e.g. after cell activation) inhibits antibody recognition; (d) steric inhibition after cell binding protects antibody-binding site; (e) 'clustering' of surface molecules results in steric inhibition.

INTERPRETATION OF EXPERIMENTAL RESULTS

Since at least some of the surface molecules on EC are able to change their activation status in relation to their functional properties, alteration of EC molecular structures under relatively native conditions is an important issue in experimental designs. For example, integrins and other adhesion molecules can be modified in their binding affinities, which can be accompanied by alterations in their quaternary structures. On the other hand, binding of antibodies is often strongly dependent on steric considerations and sometimes on the recognition of quaternary structures. Therefore, binding of antibodies to specific surface antigens is closely related to the receptor activation status. If the binding sequence is located within a critical domain, antibody attachment may not occur. In contrast, binding of antibodies to various domains of adhesion molecules can result in activation or inhibition of their function, and adhesive properties can be modified. A second problem, which is related to antibody recognition and functional consequences of antibody binding, is that various surface molecules consist of heterodimeric complexes. Therefore, antibodies against single subunits can differ in their effects compared with antibodies against more than one type of subunit in heterodimers, especially if multiple combinations of various subunits are possible, such as in the integrins. The choice of antibodies and interpretation of experimental results must consider these functional consequences of antibody binding, such as in adhesion inhibition studies used for identification of adhesion molecules involved in tumour cell attachment to EC or ECM.

Similar problems can arise if surface molecules are investigated using immunocytochemical procedures (**Figure 7**). The detection of receptor expression can depend on their functional status or the steric availability of the surface antigen. For example, surface molecules cannot be visualized by immunocytochemistry if these molecules are bound to their receptors in such a way that results in steric hindrance or occupation of the recognition domain for the antibody used. However, the ability of antibodies to recognize certain functional forms of cellular molecules can also be used in specific experimental designs for the investigation of changes in these functional properties.

BIOLOGICAL AND TECHNICAL LIMITATIONS

The organ-specific heterogeneity of EC and their potential for functional and structural alterations are among the most important problems using *in vitro* EC models. The characteristics of the EC phenotype are dependent on culture conditions, adhesive substrates and cell–cell interactions in virtually every experimental system. For example, expression of CD31 on EC surfaces is inducible by various ECM components. Additionally, the expression of various antigens (e.g. von Willebrand factor or glucosaminoglycans) and other molecules (e.g. plasminogen activator, prostaglandins or angiotensin-converting enzyme) are influenced by the age of the cultures, number of passages

and activation status of EC. Moreover, microvascular EC can deposit an insoluble ECM when cultured on plastic or collagen surfaces. Under certain culture conditions such matrices exhibit a filamentous and multilayered ultrastructure with typical characteristics of the basement membranes formed *in vivo*. Some types of EC also undergo substantial morphological and functional changes during *in vitro* culturing. For example, HUVEC cannot be used beyond passage 5–6 owing to significant changes in their phenotype. This requires frequent new preparations from different human sources. Studies on the influence of source variability on the success of EC cultures demonstrated a strong negative influence of various factors, such as smoking behaviour. EC characteristics must be frequently tested for functional criteria to obtain comparable EC. Finally, it was demonstrated that some EC lines, such as EC derived from bovine corpus luteum, express different EC types within the same vascular segments. Differentiation of these cell types by intensive cultivation resulted in EC cultures with differences in cytoskeleton structures, expression of ECM and display of adhesion molecules. Growth factors, such as EC-derived growth factor (ECGF) and cytokines (e.g. GM-CSF or interferons), can influence the growth characteristic of EC cultures, depending on the cell type.

The complex involvement of EC in microenvironmental regulation and their interactions with other cells cannot be easily mimicked by *in vitro* models. This implies that there are advantages and disadvantages in various model systems, especially if they are used for functional studies. *In vitro* systems are able to focus on individual steps in complex processes or functions but this can also result in over simplification. For example, cross-links between EC signalling cascades and cellular functions can occur and their separation may result in disparate cellular behaviour *in vitro*, which may not reflect *in vivo* behaviour. On the other hand, *in vivo* models, such as intravital video microscopy, can examine tumour cell–EC interactions under physiological conditions. However, some influencing factors might not be controllable in such *in vivo* settings, resulting in problems in reproducibility of the observed effects. Furthermore, injection of human tumour cells into animal models does not consider species-related interactions between the tumour cells and the host. For example, intravenously injected tumour cells can be attacked by the host immune system mounting a xenogenic response because many cell surface molecules demonstrate genetic and structural variability between different species. Therefore, molecular interactions between molecules of different origin may be influenced by differences in protein structure. Approaches to eliminate these problems include the injection of human tumour cells into immunodeficient animals or the use of tumour cells and animals of the same species. Also, tumour cells must maintain their tumorigenic and metastatic potentials for reliable results in these experimental designs.

In vitro culturing of EC requires specific and expensive medium to propagate the cells. Although the EC growth supplements are costly, they ensure optimal results for cell survival and growth. Other media, such as those derived from various tumour cell cultures, may vary dramatically and their effects on reproducibility can be troublesome.

Many bacterial toxins are potential inducers of EC responses, and they are common contaminants of commercial sera. They can cause release of various EC cytokines, procoagulative molecules and other factors as well as induction of EC functional changes, such as EC activation. Therefore, it is of special importance in experiments using EC to maintain an endotoxin-free environment during cell culture (media, sera) and in the experimental assays to be used.

PERSPECTIVES

Although our understanding of EC interactions with tumour cells and their involvement in organ-specific formation of metastases is increasing, much remains to be learned about the pathobiology of these interactions. Many endothelial properties are very complex, and the changes caused by tumour cell interactions, such as induction of signal transduction, and cross-talk between various signalling cascades remain to be investigated. Microcirculatory regulation and involvement of multiple cell systems, such as EC, leucocytes and platelets, in this process make relevant experimental studies a challenge. Eventually, differences between normal vascular function and pathological events may provide new therapeutic strategies for cancer treatment.

REFERENCES

Auerbach, R. (1992). Endothelial cell heterogeneity: its role as a determinant of selective metastasis. In: Simionescu, N. and Simionescu, M. (eds), *Endothelial Cell Dysfunctions.* 427–437 (Plenum Press, New York).

Bassenge, E. (1996). Endothelial function in different organs. *Progress in Cardiovascular Diseases*, **39**, 209–228.

Belloni, P. N. and Nicolson, G. L. (1988). Differential expression of cell surface glycoproteins on various organ-derived microvascular endothelia and endothelial cell cultures. *Journal of Cell Physiology*, **136**, 398–410.

Belloni, P. N. and Tressler, R. J. (1990). Microvascular endothelial cell heterogeneity: interactions with leukocytes and tumor cells. *Cancer Metastasis Reviews*, **8**, 353–389.

Bevilacqua, M. P., *et al*. (1985). Interleukin-1 activation of vascular endothelium. Effects on procoagulant activity. *American Journal of Pathology*, **121**, 393–403.

Bevilacqua, M. P., *et al*. (1986). Grimbone. Recombinant TNF induce procoagulant activity in cultured human

endothelia: characterization and comparison with the action of IL-1. *Proceedings of the National Academy of Sciences of the USA*, **83**, 4533–4541.

Carley, W. W., *et al*. (1988). Extracellular matrix specificity for the differentiation of capillary endothelial cells. *Experimental Cell Research*, **178**, 426–434.

Cavanaugh, P. G. and Nicolson, G. L. (1991). Lung-derived growth factor for lung-metastasizing tumor cells: identification as transferrin. *Journal of Cellular Biochemistry*, **47**, 261–267.

Cerra, R. F. and Nathanson, S. D. (1991). Chemotactic activity present in liver extracellular matrix. *Clinical and Experimental Metastasis*, **9**, 39–49.

Criscuolo, G. R., *et al*. (1988). Further characterization of malignant glioma-derived vascular permeability factor. *Journal of Neurosurgery*, **69**, 254—262.

Doer, R., *et al*. (1989). Clonal growth of tumors on tissue-specific biomatrices and correlation with organ site specificity of metastasis. *Cancer Research*, **49**, 384–392.

Ewing, J. (1928). *A Treatise on Tumors, Neoplastic Diseases*, 3rd edn. (Saunders, Philadelphia).

Fenyves, A. M., *et al*. (1993). Cultured microvascular endothelial cells (MVEC) differ in cytoskeleton, expression of cadherins and fibronectin matrix. *Journal of Cell Science*, **106**, 879–890.

Ford, W. D., *et al*. (1976). The migration of lymphocytes across specialized endothelium. *Cell Tissue Kinetics*, **9**, 351–361.

Gallatin, W. M., *et al*. (1983). A cell surface molecule involved in organ-specific homing of lymphocytes. *Nature*, **304**, 30–34.

Gerritsen, M. E. (1987). Functional heterogeneity of vascular endothelial cells. *Biochemical Pharmacology*, **36**, 2710–2721.

Groothius, D. R. and Vick, N. A. (1982). Brain tumors and the blood–brain barrier. *Trends in Neuroscience*, **5**, 232–235.

Guretzki, H. J., *et al*. (1994) Heparin induces endothelial extracellular matrix alterations and barrier dysfunction. *American Journal of Physiology*, **267**, 946–954.

Hamada, J., *et al*. (1992) Separable growth and migration factors for large-cell lymphoma cells secreted by microvascular endothelial cells derived from target organs for metastasis. *British Journal of Cancer*, **66**, 349–354.

Hart, R. (1982). 'Seed and soil' revisited: mechanisms of site specific metastasis. *Cancer Metastasis Reviews*, **1**, 5–16.

Holley, R. W., *et al*. (1980) Purification of kidney epithelial cell growth inhibitors. *Proceedings of the National Academy of Sciences of the USA*, **77**, 5989–5992.

Honn, K. V., *et al*. (1992) Platelets and cancer metastasis: a causal relationship? *Cancer Metastasis Reviews*, **11**, 325—351.

Hujanen, E. S. and Terranova, V. P. (1985). Migration of tumor cells to organ-derived chemoattractants. *Cancer Research*, **45**, 3517–3521.

Inoue, T., *et al*. (1993) Differences in transferrin response and numbers of transferrin receptors on rat and human mammary carcinoma lines of different metastatic potentials. *Journal of Cell Physiology*, **156**, 212–217.

Irimura, T., *et al*. (1986) Sialoglycoproteins of murine RAW117 large cell lymphoma/lymphosarcoma sublines of various metastatic colonization properties. *Experimental Cell Research*, **165**, 403–416.

LaBiche, R. A. and Nicolson, G. L. (1993). Modulating the metastatic potential of murine RAW117 large cell lymphoma by selection for resistance to interferon-γ. *International Journal of Cancer*, **54**, 1003–1010.

Li, L., *et al*. (1991). Direct *in vitro* lysis of metastatic tumor cells by cytokine-activated murine vascular endothelial cells. *Cancer Research*, **51**, 245–254.

Lotan, R., *et al*. (1994). Expression of galectins in endothelial cells and their involvement in tumor cell adhesion. *Glycoconjugate Journal*, **11**, 462–468.

Lum, H. and Malik, A. B. (1994). Regulation of vascular endothelial barrier function. *American Journal of Physiology*, **267**, 223–241.

Manjo, G. (1965). Ultrastructure of the vascular membrane. In: Hamilton, W. F. and Dow, P. (eds), *Handbook of Physiology, Circulation Vol. III*. 2293–2375 (Waverly Press, Baltimore).

Marchetti, D., *et al*. (1993). Nerve growth factor effects on human and mouse melanoma cell invasion and heparanase production. *International Journal of Cancer*, **55**, 693–699.

Menter, D. G., *et al*. (1995). The role of trophic factors and autokrine/parakrine growth factors in brain metastasis. *Clinical and Experimental Metastasis*, **13**, 67–88.

Nicolson, G. L. (1982). Cancer metastasis: organ colonization and the cell-surface properties of malignant cells. *Biochimica Biophysica Acta*, **695**, 113–176.

Nicolson, G. L. (1988). Organ specificity of tumor metastasis: role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer Metastasis Reviews*, **7**, 143–188.

Nicolson, G. L. (1989). Metastatic tumor cell interactions with endothelium, basement membrane and tissue. *Current Opinions in Cell Biology*, **1**, 1009–1019.

Nicolson, G. L. (1999). Cell contact-dependent cytostasis but not cytolysis of tumor cells by target syngenic microvascular endothelial cells in vitro correlates with organ-specific metastasis. *Proceedings of the American Association Cancer Research*, **40**, 231.

Nicolson, G. L. and Dulski, K. M. (1986). Organ specificity of metastatic tumor colonization is related to organ-selective growth properties of malignant cells. *International Journal of Cancer*, **38**, 289–294.

Nicolson, G. L., *et al*. (1991). Effects of gamma irradiation on cultured rat and mouse microvessel endothelial cells: metastatic tumor cell adhesion, subendothelial matrix degradation, and secretion of tumor cell growth factors. *Clinical and Experimental Metastasis*, **9**, 457—468.

Ordinas, A., *et al*. (1990) The role of platelets in cancer metastasis. *Blood Coagulation and Fibrinolysis*, **1**, 707–711.

Orr, F. W., *et al*. (1980). Partial characterization of bone-derived chemotactic factors for tumor cells. *American Journal of Pathology*, **99**, 43–52.

Paget, S. (1889). The distribution of secondary growth in cancer of the breast. *Lancet*, **1**, 571–573.

Pauli, B. U., *et al*. (1990) Organ-preference of metastasis: the role of endothelial cell adhesion molecules. *Cancer Metastasis Reviews*, **9**, 175–189.

Rhoden, L. (1980). Structure and metabolism of connective tissue proteoglycans. In: Lennarz, W. J. (ed.), *The Biochemistry of Proteoglycans*. 180–221 (Plenum Press, New York).

Rickles, F. R., *et al*. (1992). Hemostatic alterations in cancer patients. *Cancer Metastasis Reviews*, **11**, 237–248.

Roy, S. and Chitra, S. (1989). Ultrastructural study of microblood vessels in human brain tumors and peritumoral tissue. *Journal of Neurooncology*, **7**, 283–294.

Sawada, H., *et al*. (1996). Differential motility stimulation but not growth stimulation or adhesion of metastatic human colorectal carcinoma cells by target organ-derived liver sinusoidal endothelial cells. *Clinical and Experimental Metastasis*, **14**, 308–314.

Schnitzer, J. E., *et al*. (1994). Segmental differentiation of permeability, protein glycosylation, and morphology of cultured bovine lung vascular endothelium. *Biochemical and Biophysical Research Communications*, **199**, 11–19.

Simionescu, M., *et al*. (1982). Preferential distribution of anionic sites on the basement membranes and abluminal aspect of the endothelium in fenestrated capillaries. *Journal of Cell Biology*, **95**, 425–434.

Sugarbaker, E. V. (1981). Patterns of metastasis. *Cancer Biology Reviews*, **2**, 235–278.

Tressler, R. J., *et al*. (1993). Extracellular annexin II is associated with divalent cation-dependent tumor cell–endothelial cell adhesion of metastatic RAW117 large-cell lymphoma cells. *Journal of Cellular Biochemistry*, **53**, 265–276.

Wang, Z. W., *et al*. (1985). Characterization of the ECM-associated GAG produced by untransformed and transformed bovine corneal endothelial cells in culture. *European Journal of Biochemistry*, **153**, 125–130.

Weiss, L., *et al*. (1988). Interaction of cancer cells with the microvasculature during metastasis. *FASEB Journal*, **2**, 12–21.

Zhang, M. and Olsson, Y. (1997). Hematogenous metastases of the human brain – characteristics of peritumoral brain changes: a review. *Journal of Neurooncology*, **35**, 81–89.

FURTHER READING

Auerbach, R. (1991). Vascular endothelial cell differentiation: organ-specificity and selective affinities as the basis for developing anti-cancer strategies. *International Journal of Radiology*, **60**, 1–10.

Belloni, P. N. and Nicolson, G. L. (1992). The role of the vascular endothelium in cancer metastasis. In: Simionescu, N. and Simionescu, M. (eds), *Endothelial Cell Dysfunctions*. 395–425 (Plenum Press, New York).

Bicknell, R. (1993). Heterogeneity of the endothelial cell. I. Physiological aspects. *Behring Institut Mitteilungen*, **92**, 1–7.

Haier, J. and Nicolson, G. L. (1999). Role of the vascular endothelium in cancer metastasis. In: Skouteris, G. and Nicolson, G. L. (eds), *Intermolecular Cross-talk in Tumor Metastasis* (ISO Press, Geneva).

Nicolson, G. L., *et al*. (1992). Differential stimulation of the growth of lung-metastasizing tumor cells by lung (parakrine) growth factors: identification of transferrin-like mitogens in lung-tissue-conditioned medium. *National Cancer Institute Monographs*, **13**, 153–161.

Nicolson, G. L. (1993). Paracrine and autocrine growth mechanisms in tumor metastasis to specific sites with particular emphasis on brain and lung metastasis. *Cancer Metastasis Reviews*, **12**, 325–343.

Nicolson, G. L. (1993). Cancer progression and growth: relationship of paracrine and autocrine mechansims to organ preference of metastasis. *Experimental Cell Research*, **204**, 171–180.

Nicolson, G. L. (1995). Tumor cell interaction with the vascular endothelium and their role in cancer metastasis. In: Goldberg, I. D. and Rosen, E. M. (eds), *Epithelial–Mesenchymal Interactions in Cancer*. 123–156 (Birkhäuser Verlag, Basel).

Shepro, D. and Dunham, B. (1986). Endothelial cell metabolism of biogenic amines. *Annual Review of Physiology*, **48**, 335–368.

Stemmerman, M. B. (1980). General properties of blood vessels: vascular endothelia. In: Abramson, D. and Dobrin, P. (eds), *Blood Vessels and Lymphatics in Organ Systems.* 25–31 (Academic Press, New York).

Chapter 66

Modelling Tumour Tissue Interactions

Fred R. Miller
Barbara Ann Karmanos Cancer Institute and Wayne State University School of Medicine, Detroit, MI, USA

CONTENTS

- Introduction
- The Extracellular Matrix and Cell Function
- Mesenchymal–Epithelial Interactions in Embryonic Organogenesis
- Interactions in the Adult Mammary Gland
- Growth of Primary Tumours
- Metastasis
- Conclusion

INTRODUCTION

Tumours are complex tissues consisting of many variant tumour subpopulations which, if isolated, have widely differing growth rates, metastatic potential, response to therapy and genetic stability. The phenotype of an intact tumour results from interactions amongst its subpopulations and from interactions with tissue elements including, in the case of carcinoma, remaining normal and premalignant epithelial cells, stromal cells, and extracellular matrix components (**Figure 1**). Stromal cells include a number of cell types including fibroblasts, myofibroblasts (positive for alpha-smooth muscle actin and negative for cytokeratin), endothelial cells, adipocytes and infiltrating lymphocytes and macrophages. Myofibroblasts are common in cancer stroma and are probably derived from fibroblasts responding to cytokines produced by tumour cells. The extracellular matrix composition varies and contains a variety of growth factors as well as components for cellular attachment factor interactions. Premalignant progression, growth of primary cancers, invasion and metastasis are all altered by stromal interactions. The growing appreciation of the importance of orthotopic implantation in experimental oncology reflects the recognition of these interactions as critical determinants of cancer behaviour. However, far too often models to study stromal interactions simply mix tumour cells with fibroblasts. This chapter is intended to provide an overview of the interactions occurring in cancer tissues and instill an appreciation of the complexity of these tissue interactions, in those wishing to model them. Interactions occurring in the breast will often be discussed in detail but the principles apply to other organs. (See also chapter *Extracellular Matrix: the Networking Solution*.)

THE EXTRACELLULAR MATRIX AND CELL FUNCTION

Gene expression in epithelial cells is dependent upon cell shape and extracellular matrix components. Corneal epithelial cells respond to fibroblast growth factor if grown on plastic but on a collagen matrix these cells respond to epidermal growth factor, not fibroblast growth factor (Gospodarowicz *et al*., 1978). An assay for diffusible growth factors in which both producer and responder cells grow in a three-dimensional array in a collagen gel matrix (**Figure 2**) reflects the effect of normal mammary stromal and epithelial cells on neoplastic and preneoplastic cell growth *in vivo*; that is, both normal mammary epithelium and normal mammary stroma stimulate growth of mammary tumour cells but not growth of normal mammary epithelium or preneoplastic hyperplastic alveolar nodule cells (Miller *et al*., 1989). However, in parallel experiments using monolayer cultures, neoplastic cells are inhibited by normal mammary cells *in vitro*. Such results argue for the use of an extracellular matrix and allowing for three-dimensional structure in assays of cellular interactions via diffusible factors. However, the composition of the extracellular matrix is also important. For example, the prolactin and insulin response of mammary epithelium is greater on a basement membrane type extracellular matrix than on collagen type I, whereas response to epidermal growth factor is greater on collagen I than on basement membrane (Lee and Streuli, 1999).

The extracellular matrix is connected to the nucleus by a network of proteins consisting of transmembrane adhesion molecules, the cytoskeleton and the nuclear matrix. Cytoskeletal components interact with signal transduction pathways and are linked to both the extracellular matrix

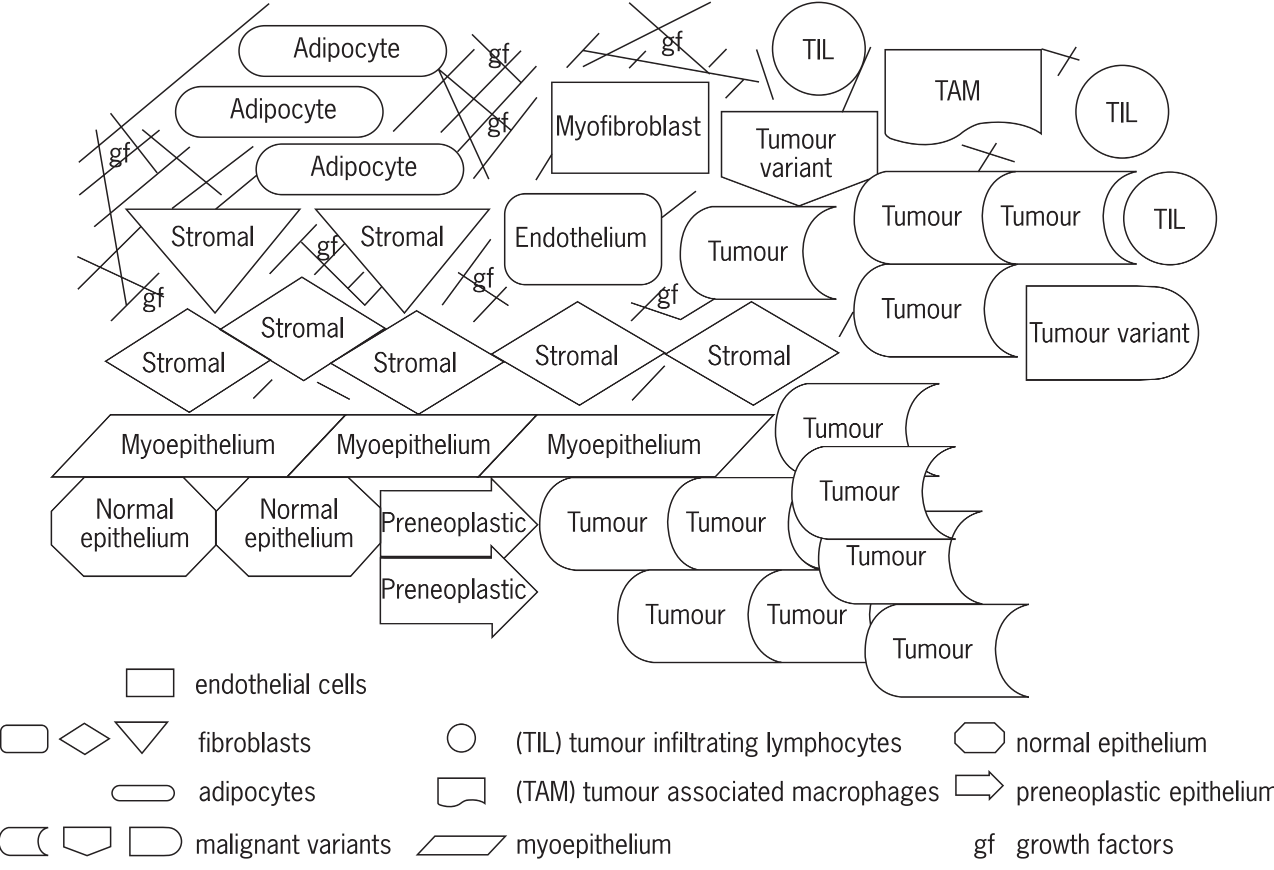


Figure 1 Interactions in a malignant tumour involve multiple cell types.

and the nuclear matrix. The matrix networks are different between normal and transformed cells and modification of the extracellular matrix environment affects the protein composition of the nuclear matrix (Getzenberg *et al.*, 1991). The nuclear matrix is associated with DNA replication and regulation of gene expression.

The stromal extracellular matrix is altered in cancer. The tumour stromal extracellular matrix contains more collagens, glycosaminoglycans and proteoglycans which may be structurally altered as well as increased in quantity. Many components which are restricted to the ductular–stromal interface in normal breast are diffusely expressed throughout the cancer stroma. Tenascin-C is highly expressed during embryogenesis, during wound healing and in tumours. Tenascin-C is weakly expressed in normal breast and varies with the menstrual cycle, being lowest during weeks three and four of the cycle. As well as increased quantities, localization is altered in breast cancer. In the normal breast tenascin-C expression is primarily in the basement membrane and just outside the ductule but in cancer tenascin is absent from the basement membrane and present diffusely in intralobular and interlobular stroma (Ferguson *et al.*, 1990).

Cell adhesion molecules involved in cell–extracellular matrix interactions include integrins which consist of an α and a β chain. Various combinations of α and β chains are receptors for different matrix components. The $\beta1$ group bind collagens, fribronectin and laminin, whereas the $\beta3$ integrins bind thrombospondin and vitronectin. The integrins $\alpha2$, $\alpha6$, $\beta1$ and $\beta4$ are decreased in most breast cancers. For example, 80% of breast cancers had complete loss of detectable $\alpha6\beta4$. Upon binding to specific matrix components, signal transduction is triggered through focal adhesion proteins to regulate proliferation, apoptosis and gene expression. Thus, alterations in extracellular matrix components and adhesion molecules in cancer result in altered gene expression in the cancer cells which can lead to altered motility and invasive properties as well as loss of attachment itself.

MESENCHYMAL—EPITHELIAL INTERACTIONS IN EMBRYONIC ORGANOGENESIS

Mesenchymal–epithelial interactions play a fundamental role in tissue development in the embryo. Developmental response of epithelium to differentiation factors, including steroid hormones, is mediated through stromal/mesenchymal cells. If the mesenchyme is unable to respond, for example owing to a genetic disorder, organogenesis is aberrant. Formation of the mammary gland epithelial bud,

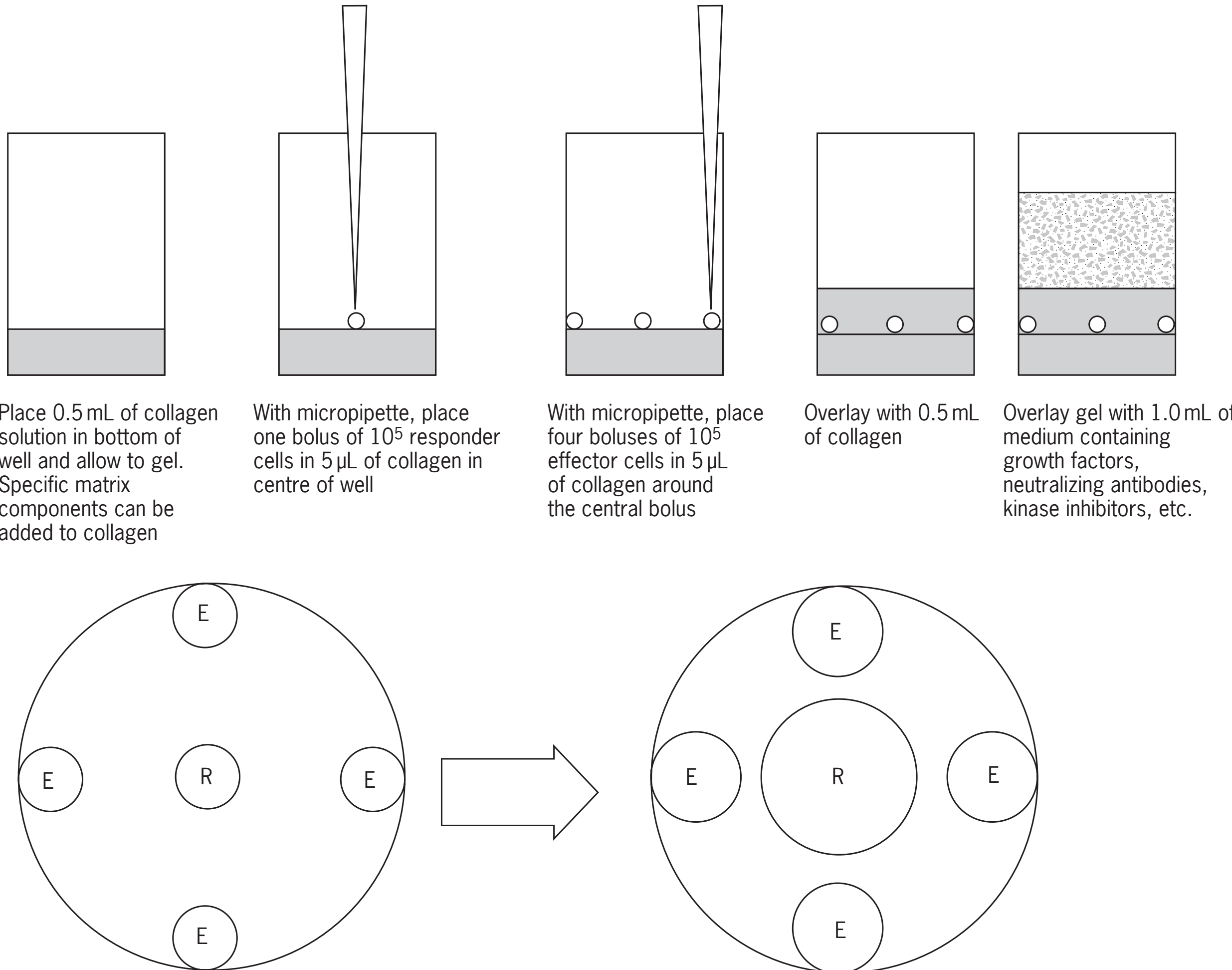


Figure 2 Assay for interactions via soluble factors in three-dimensional arrays in matrix.

appearance of hormone receptors and formation of the primary mammary sprout and ductal elongation are regulated by mesenchymal–epithelial interactions. Mesenchymal fibroblasts and mesenchymal adipocytes have separate, distinctive roles in the regulation of embryonic development of the mammary gland. Mesenchymal fibroblasts respond to testosterone to induce regression of the foetal mammary bud in developing males while the mesenchymal adipocytes regulate the initial mammary ductal elongation in females.

Most of our understanding of these interactions comes from tissue-recombination studies in which mutant or knockout mouse mesenchyme is mixed with wild-type epithelium or vice versa. Mesenchyme from *tfm* (testicular feminization) mice do not induce regression of wild-type mammary buds in response to testosterone but wild-type mesenchyme induces regression of *tfm* epithelial mammary buds. The *tfm* mutation is probably a nonfunctional androgen receptor and results in a lack of all secondary male sex characteristics in male mice. Thus, functional androgen receptors are unnecessary in the mammary epithelium but must be present in the mesenchyme for mammary bud involution in the male embryos. Tissue recombination experiments with knockout mouse tissues illustrate that epidermal growth factor receptor must be expressed by the stromal element for mammary gland ductal morphogenesis (Wiesen *et al*., 1998) and progesterone receptor is needed in the stroma for ductal development but in the epithelium for lobuloalveolar development (Humphreys *et al*., 1997).

Both mesenchyme and epithelium retain organogenic inductive response competence in the adult. The latter phenomenon has been particularly well documented with respect to the mammary gland. Morphogenetically normal outgrowths result from transplanting fetal mammary epithelium into adult mammary stroma and adult mammary epithelium responds to fetal mesenchyme. The epithelial portion of the mouse mammary gland may be surgically removed prior to the age of 21 days to obtain cleared fatpads consisting of normal stroma but lacking epithelial components. If fetal mammary epithelium is transplanted into a cleared adult mammary gland, morphogenetically normal

outgrowths result, but growth of fetal epithelia from other organs is not supported by adult mammary stroma. Thus, adult stroma retains specific inductive capacity. Adult mammary epithelium from virgin mice responds to fetal mesenchyme, and the morphogenetic branching pattern induced is dependent upon the organ source of the mesenchyme (Sakakura *et al*., 1979). The retention of organogenic inductive responses in the adult may reflect the fact that normal breast must be able to remodel glandular architecture repeatedly in response to hormonal stimuli during oestrus (mouse) or menstrual (human) cycles and successive pregnancies. However, orthotopic effects have been observed for many other types of cancer including colorectal, melanoma, prostate, lung, renal, bladder and pancreatic carcinoma, indicating that stromal–epithelial interactions persist in many adult tissues. *In vitro* studies have elucidated both contact-dependent and soluble factor-mediated mechanisms of interactions.

INTERACTIONS IN THE ADULT MAMMARY GLAND

In the mouse, both normal and preneoplastic mammary epithelia are stroma dependent, that is, they only grow in mammary fatpads and not at ectopic sites. The growth of both normal and preneoplastic cells is inhibited by normal mammary epithelium. Thus, to obtain growth, normal and preneoplastic cells must be transplanted into cleared mammary glands. Mammary tumours are obviously not mammary stroma dependent because tumours grow after subcutaneous implantation and at metastatic sites. However, tumour formation in the fatpad requires ∼10-fold fewer cells than at subcutaneous sites (Miller *et al*., 1981) and thus even tumour cells are ‘stromal-responsive.’ Furthermore, mouse mammary tumour cells are stimulated rather than inhibited by normal mammary epithelium, as evidenced by their enhanced growth in intact versus growth in cleared fatpads (Miller *et al*., 1981). Perhaps the most important orthotopic tissue effect is the enhanced metastasis displayed by mouse mammary tumour cells (Miller, 1981) and xenografted human breast cancer (Price *et al*., 1990). Fibroblasts can stimulate growth of epithelial cells as well as increase the invasiveness of malignant epithelium. Epithelium can stimulate growth of stromal cells and induce stroma to produce proteases important in invasion and release of growth factors from the extracellular matrix. Reciprocal interactions between stroma and epithelium are common.

GROWTH OF PRIMARY TUMOURS

Angiogenesis

Perhaps the most important role of stroma in the growth, invasion and metastasis of cancer is to provide additional blood vessels. Nutrients required to support a growing mass of neoplastic cells can diffuse through tissue up to a distance of only 1–2 mm. Thus, tumours will grow only to 2–3 mm in diameter unless vascularization of the tumour occurs. The ability to induce an angiogenic response is one critical difference between premalignant and malignant mammary lesions. Pieces of tissue from human ductal carcinoma *in situ* and invasive breast cancers both induce angiogenic activity when implanted in the rabbit cornea, whereas normal breast and fibroadenoma tissues do not (Brem *et al*., 1978). Thus, angiogenic activity is a stromal–epithelial interaction which is altered at an early stage in progression (i.e. carcinoma *in situ*). (See also chapter *Angiogenesis*.)

In some cases, the malignant epithelial cells produce angiogenic factors that directly induce the host to produce the vascular network. Vascular endothelial growth factor (VEGF) is one such molecule which may be released by cancer cells, possibly as a response to hypoxia. In some cases cancer cells may induce VEGF production by the stroma (Fukumura *et al*., 1998). In the latter study transgenic mice expressing green fluorescent protein under the control of the VEGF promoter show strong green cellular fluorescence in the stromal fibroblasts, but not tumour cells, of spontaneous mammary tumours. Other factors with possible roles in inducing angiogenesis include acidic and basic fibroblast growth factors, platelet-derived growth factor and angiogenin.

The angiogenic activity of specific factors can be assesssed by implanting pellets containing the factor in the cornea of mice (Kenyon *et al*., 1996). Neovascularization occurrs between days 3 and 8 after bFGF or VEGF are implanted. Systemic angiogenesis inhibitors can also be detected with this model. One advantage of the mouse cornea model is the availibilty of immunodeficient mice for xenograft studies. Some, but not all, human tumours xenografted into immunodeficient mice inhibit the angiogenic response to bFGF pellets in the mouse cornea (Chen *et al*., 1995). Transgenic and knockout mice also allow the study of the role of specific gene products in angiogenesis.

A disadvantage of using nude mice for the cornea assay is that the host defence mechanisms, which also may contribute to neovascularization of a tumour mass, are deficient. For example, a graft versus host reaction induces an angiogenic response after injecting lymphocytes intradermally into irradiated allogeneic mice (Sidky and Auerbach, 1975). Activated macrophages are also angiogenic, as indicated by the induction of vascular proliferation when injected into guinea pig corneas. Cytokines produced by monocytes that have been shown to have angiogenic activity include tumour necrosis factor alpha (TNFα) transforming growth factor beta (TFGβ) and interleukin-8. Although TNFα inhibits endothelial cells in culture, this factor induces angiogenesis *in vivo* when placed in the rabbit cornea, suggesting that the angiogenic response is due to infiltrating leucocytes induced by TNFα rather than due to a direct action of TNFα on the

endothelium. Whether a specific T cell-mediated response or a general inflammatory response is mounted, host defence efforts may be instrumental in allowing the further growth of a neoplastic lesion by inducing the neovascularization process.

Stromal Heterogeneity

Studies to understand the mechanisms of interactions have relied upon *in vitro* experiments in which fibroblasts and epithelial tumour cells are grown in mixtures or in which conditioned medium from one cell type is supplied to the other. It is important to remember that such systems are highly simplified and may not be very representative of conditions *in situ*. There are two basic problems in modeling stromal interactions *in vitro*. One is the effect of matrix and cell shape on cellular responses and interactions. Before utilizing an *in vitro* assay to study mechanisms of interactions, it should be verified that that the product of the interaction successfully models *in vivo* results. For example, mouse mammary tumour cells are stimulated by both stroma and epithelium *in situ* (Miller, *et al*., 1981). However, in monolayer cocultures, primary mammary gland cultures inhibit the growth of tumour cells, indicating that monolayer culture fails to reproduce the environment which *in vivo* results in stimulation of tumour cell proliferation by normal cells. The three-dimensional assay in collagen matrix does detect paracrine stimulation of neoplastic (but not normal or preneoplastic) mouse mammary cells by normal mammary epithelium and stromal fibroblasts (Miller *et al*., 1989). While the three-dimensional collagen gel assay is more difficult to perform than monolayer cultures and conditioned medium, it has the single advantage that it more accurately reflects tissue interactions in the mammary gland. An additional disadvantage is that production of differentiated structures and cell motility and invasion may increase the size of the cell mass being measured without an accompanying increase in cell number. Hence it is sometimes difficult to discriminate between effects of interactions on growth from those which alter differentiation state.

The other problem in modelling stromal interactions *in vitro* is the heterogeneity of the stroma and also tumour cell heterogeneity. Stroma is often represented *in vitro* by immortalized mouse fibroblast cell lines or primary skin fibroblast cultures. Mouse mammary stroma is grossly less complex than human breast stroma (**Figure 3**). In the mouse, a thin rim of fibroblasts surrounds the epithelial ducts and adipocytes constitute the majority of the stroma between the ducts. In the human breast there are many more fibroblasts with both an interlobular and an intralobular type and pockets of adipocytes. Ectoenzymes differ for inter- and intralobular fibroblasts and neutral endopeptidase is produced only by stroma from breast cancer (Atherton *et al*., 1994). In fact, based on differential gene expression, there are at least seven types of breast fibroblasts derived from normal, fibrotic and tumoural breast including fibroblasts and myofibroblasts from breast skin, normal stroma, tumour-adjacent stroma, postradiation fibrosis lesions and benign or malignant tumours (Spanakis and Brouty-Boye, 1997).

These stromal fibroblast types do not alter epithelial behaviour uniformly. Conditioned medium from malignant stroma was found to be more stimulatory for human breast cancer cell line MCF-7 than conditioned medium from normal stroma or from stroma adjacent to malignant breast tissue (van Roozendaal *et al*., 1996). Early full-term pregnancy significantly lowers the risk of breast cancer. It has been suggested that the pregnancy causes differentiation of the epithelial cells, thus reducing the number of stem cells at risk of carcinogenesis. However, the reduced risk may be due to altered stroma rather than, or in addition to, altered epithelial cells. In the rat, uniparous mammary stroma is more inhibitory to development of mammary tumours than is normal mammary stroma after carcinogen-treated rat mammary epithelium is transplanted into untreated uniparous or virgin fatpads (Abrams *et al*., 1998).

One cannot ignore stromal elements other than fibroblasts. A common perception is that host defence mechanisms act as a natural surveillance system to recognize and eradicate neoplastic cells and it is often assumed or hoped that infiltrating lymphocytes and macrophages are inhibitory or, at worst, ineffective. However, lymphocytes are more stimulatory to breast cancer cells *in vitro* than are fibroblasts (Ogmundsdottir *et al*., 1993) and tumour-associated macrophages have aromatase activity and thus may act as a local source of oestrogen (Mor *et al*., 1998). Hence, these cell types must also be regarded as potential sources of tumour growth stimulation.

The Role of the Stroma in Hormonal Responsive Tumours

The lifetime exposure to oestrogen is an important factor in breast cancer risk and it is generally thought that mitogenic activity is important in this process. Mouse mammary stroma may play an important role in the oestrogen response of mixed mammary fibroblasts and epithelium (Haslam and Counterman, 1991). A likely role of stromal cells is the local production of oestradiol within the breast so that a paracrine mechanism independent of systemic oestradiol levels can drive epithelial expansion. Breast fibroblasts plus testosterone accelerate growth of breast cancer xenografts in ovariectomized nude mice perhaps due to increased aromatase (Koh *et al*., 1998). Human breast stroma produces a 50-kDa protein which stimulates MCF-7 production of reductive 17β-oestradiol dehydrogenase activity, thus allowing the conversion of oestrone to oestradiol (Adams *et al*., 1988). The human oestrogen-dependent breast cancer line MCF-7 releases

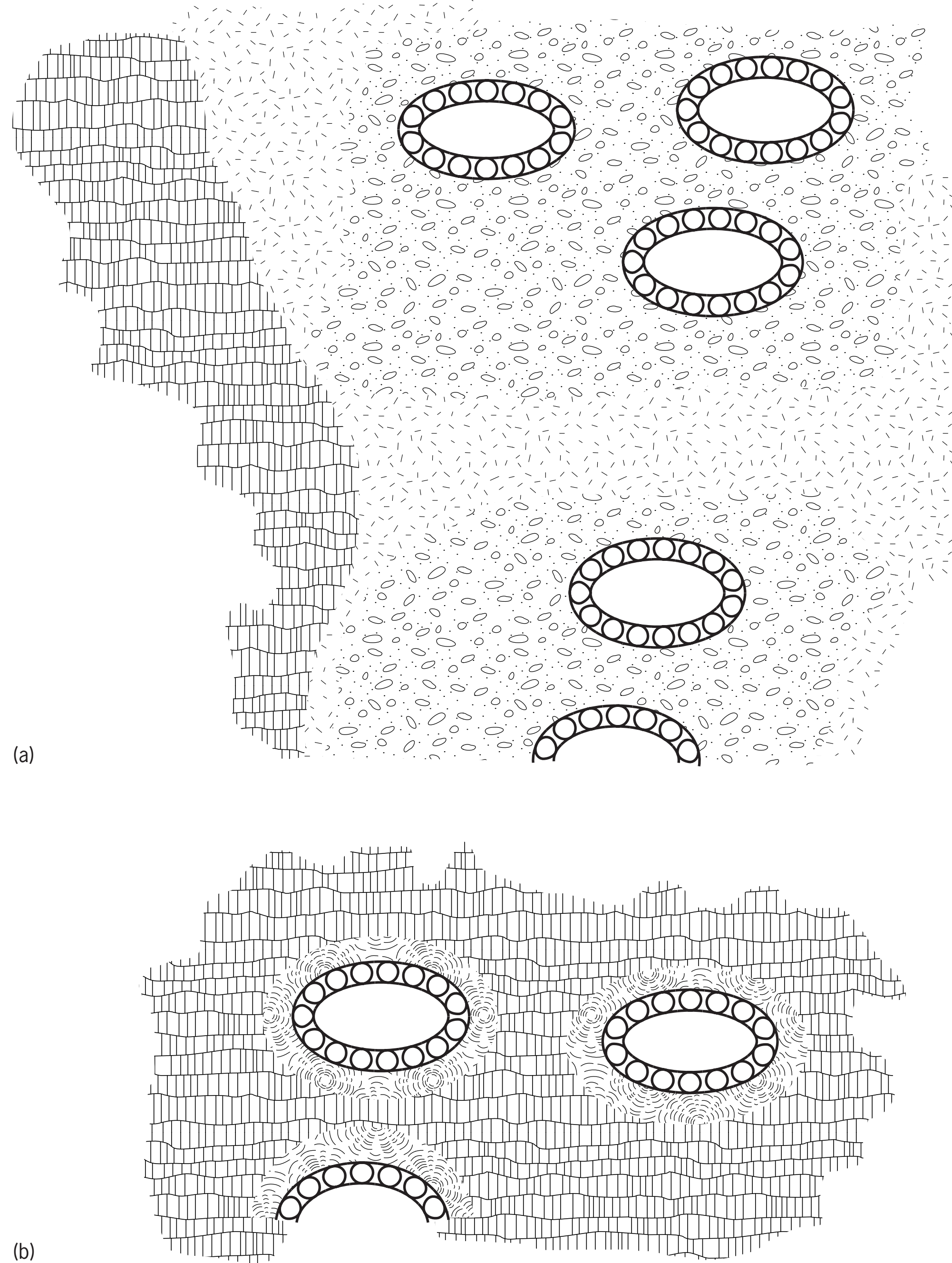


Figure 3 (a) Human breast terminal epithelial ductules (ovals with circles), intralobular stroma (), interlobular stroma (), and adipocytes (). (b) Mouse mammary gland ductules (ovals with circles), stromal fibroblasts () and adipocytes ().

both interleukin-6 and its soluble receptor to up-regulate local aromatase activity (Singh *et al*., 1995).

Stroma Stimulates Growth of Epithelium

Mammary stroma from either cancer or normal breast stimulates growth of MCF-7 *in vitro* (Gache *et al*., 1998). Normal human skin fibroblasts or Matrigel (a basement membrane extracellular matrix available commercially) mixed with human breast cancer cell lines increases xenograft incidence and decreases latency; fibroblasts plus Matrigel are even more stimulatory. Matrigel depleted of growth factors does not work. Stromelysin may be the enzyme which releases active growth factors from extracellular matrix because fibroblasts from stromelysin 3 knockout mice do not increase growth of MCF-7 xenografts (Masson *et al*., 1998). Alternatively, fibroblasts may release matrix metalloproteinases which release growth

factors from matrix (because fibroblasts added to Matrigel have no effect on xenografts of MCF-7 cells which overexpress an inhibitor of metalloproteinase (Noel *et al.*, 1998). One disadvantage of Matrigel is that it is prepared from a mouse sarcoma. It varies in growth factor composition among batches but growth factors can be depleted allowing the addition of specific factors of interest. However, when studying tumours of other species one must remember that the extracellular components are mouse and, possibly more important, are from a malignant sarcoma. A possible alternative for human cell studies is Humatrix, which is made by human myoepithelial cells (Kedeshian *et al.*, 1998). Humatrix also contains a number of growth factors, such as EGF and insulin-like growth factor I, but is reported to contain protease inhibitors and angiogenesis inhibitors rather than proteases and angiogenic factors.

Progression of various carcinomas is often accompanied by increased expression of growth factor receptors such as EGFR or erbB2. In prostate, keratinocyte growth factor, which is produced by stromal cells, is found in 65% of prostate cancer tissues but not in benign hyperplasias. The receptor is expressed by epithelium in both benign and malignant epithelium so the observed change with progression is stromal rather than epithelial (Leung *et al.*, 1997).

Reciprocal Stromal–Epithelial Interaction

Epithelium also controls stromal behaviour in tumours resulting in reciprocal interactions in which each compartment is both sending and receiving signals from the other. Stromal proliferation is a common occurrence in breast cancer and mitotic activity is increased in stromal cells near the epithelium. Cancer cells often produce growth factors the receptor for which is expressed in the stroma. Breast cancer epithelium produces basic fibroblast growth factor whereas breast stromal cells express FGFR (Hasebe *et al.*, 1997). Breast cancer cells make platelet-derived growth factor but have no receptor. Cancer makes platelet-derived growth factor which induces stroma to make insulin-like growth factors which are mitogenic for cancer cells (Yee *et al.*, 1991). Mouse mammary tumour cells produce prostaglandin E2, which induces human fibroblasts to produce hepatocyte growth factor which is mitogenic for the tumour cells (Matsumoto-Taniura *et al.*, 1999).

Normal breast fibroblasts produce insulin-like growth factor I (IGFI), not IGFII, whereas tumour-derived fibroblasts produce IGFII, not IGFI (Cullen *et al.*, 1991). MCF-7 induces normal breast fibroblasts to produce IGFII *in vitro* (Singer *et al.*, 1995) and *in situ* expression of IGF2 is localized to stroma specifically surrounding malignant breast epithelium (Rasmussen and Cullen, 1998). Since both IGFI and IGFII are potent mitogens for breast cancer cells, the advantage to the tumour cell of local production of IGFII rather than IGFI is not clear but might relate to progression to oestrogen independence (Daly *et al.*, 1991).

Contact-dependent Interactions

Cell to cell contact has long been recognized as an important element in homeostasis and loss of contact inhibition *in vitro* is a hallmark of transformation. However, growth of transformed cells may be inhibited by contact with normal cells *in vitro*. Contact-dependent regulation may involve gap junction-mediated intercellular communication (GJIC), cellular adhesion molecules, and juxtacrine growth factor/receptor binding.

Gap Junctions

Gap junctions allow the passage of small regulatory molecules such as cAMP and inositol triphosphate to diffuse among cells. A gap junction consists of localized plaques of multiple intercellular channels each of which consists of a six-protein oligomer (connexon) in the membrane of each of the communicating cells. More than a dozen of these proteins, called connexins, have been identified. Different tissues express a limited number of these connexins but an individual cell type generally expresses more than one. A connexon may be homotypic, consisting of six identical proteins, or heterotypic. Some connexins only form homotypic connexons and some combinations of connexins form nonfunctional heterotypic connexons. Passage through connexon pores is limited to molecules under 1 kDa and additional selectivity may be provided by the connexin species constituting the connexon. Loewenstein first suggested that regulatory molecules may be transmitted between cells via gap junctions. Since then, the role of GJIC in transformation, promotion and progression to a metastatic phenotype has been the subject of numerous investigations. Carcinogens and tumour promoters, such as 12-*O*-tetradecanoylphorbol-13-acetate, inhibit gap junction mediated communication. The chemopreventive agent all-*trans*-retinoic acid increases connexin 43 in renal cell cancer and prevents disruption of GJIC by 12-*O*-tetradecanoylphorbol-13-acetate and carcinogen *in vitro* (Watanabe *et al.*, 1999).

Measurement of GJIC between cells has depended upon passage of low molecular mass dyes or passage of small molecules or ions in metabolic cooperation assays. The simplest assay based on the transfer of the dye lucifer yellow from stained to unstained cells coupled by gap junctions involves scraping the monolayer with any sterile object in the presence of fluorescent dyes. Monolayers can be 'scrape-loaded' with lucifer yellow and rhodamine isothiocyanate dextran, incubated at 37°C for 20 min, and cell suspensions prepared and analysed by fluorescence cell cytometry (Kavanagh *et al.*, 1987). Lucifer yellow readily passes through gap junctions but rhodamine isothiocyanate dextran, owing to its large size, does not. Scraped cells should fluoresce at both 545 nm (lucifer yellow stained) and 590 nm (rhodamine isothiocyanate dextran stained), cells which have been coupled to scrape-loaded cells should be stained with lucifer yellow only, and

cells which did not communicate via gap junctions should contain neither dye. Metabolic cooperation assays utilize either 'kiss of death' or 'kiss of life' strategies (**Figure 4**). The kiss of death assay most commonly utilizes a mixture of hypoxanthine–guanine phosphoribosyltransferase deficient (HPRT−) and wild-type cells. The HPRT− cells are resistant to 6-thioguanine, which forms toxic thioguanine nucleotides in wild-type cells. If GJIC between the two cell types occurs, in the presence of 6-thioguanine, thioguanine nucleotides pass from the wild-type cells to the HPRT− cells, resulting in cell death. If GJIC does not occur, the wild type-cells will be killed but the HPRT− cells will survive and replicate. The number of wild-type cells needed to reduce the number of HPRT− cells in co-culture can be used as a quantitative estimate of relative efficiency of GJIC of different wild-type cells with a given HPRT− indicator cell. Kiss of life metabolic cooperation assays may utilize mixtures of HPRT− and wild-type cells in HAT medium (hypoxanthine, aminopterin, thymidine) or mixtures of ouabain-resistant and wild-type cells. HAT medium kills HPRT− cells because of the inability to salvage purines required by aminopterin block of *de novo* synthesis. If GJIC occurs with an HPRT + cell, the latter can produce sufficient purines from the active salvage pathway to rescue the HPRT− cell. Ouabain inhibits the active transport of sodium ions out of the cell via the Na^+:K^+ pump. If GJIC is established, ouabain-resistant cells can remove excess Na^+ from wild-type cells, allowing their growth in the presence of lethal levels of ouabain. In mixtures of wild-type cells which are HPRT + and are sensitive to ouabain and cells that are ouabain resistant and HPRT deficient, a reciprocal kiss of life is necessary for any growth in medium containing HAT and ouabain.

Aberrant communication competence has been described for some cancer cells. Some have been found to have a diminished ability to communicate and others have lost selectivity in communication (Fentiman *et al*., 1979) which might allow a response to improper signals. This may be the result of changed connexin species being expressed, loss of connexin expression or inappropriate processing of connexins. Loss of intercellular communication competence has been implicated during essentially all steps of initiation, promotion and progression of neoplasia. Cells transformed by oncogenes often have reduced GJIC. Transformation with H-*ras* reduced communication between transformed epithelial cells but not between transformed 3T3 fibroblasts (Vanhamme *et al*., 1989). Growth inhibition of *ras*- or *neu*-transformed cells by non-transformed cells required GJIC (Esinduy *et al*., 1995). Sequential loss of coupling was reported during progression from (1) normal rat ovarian granulosa cells to (2) immortalized cells to (3) contact independent cells to (4) cells tumorigenic in nude mice (Stein *et al*., 1991).

In addition to loss of GJIC or loss of selectivity, a third class of communication defect which might result in or allow cancerous growth is the loss of responsiveness to modulators of coupling. Activity of individual connexons may be regulated by tyrosine phosphorylation and net GJIC can be regulated by altering turnover rates (synthesis and/or degradation) as gap junctions have a half-life of 5 h (Laird, 1996). Both EGF and TGFβ may inhibit GJIC but TGFβ has also been reported to increase GJIC in keratinocytes (Madhukar *et al*., 1989) as well as in fibroblasts (Gibson *et al*., 1994). In the latter study, TGFβ treatment of quiescent 10T1/2 cells elevated proliferation and increased expression of cx43 with elevated GJIC, illustrating that increased cell proliferation can occur in the presence of GJIC. However, TGFβ more typically is involved in the GJIC-mediated inhibition of transformed cells by normal, untransformed cells (Hofler *et al*., 1993).

Hormones may function as promoters for initiated, hormone-responsive cells by altering GJIC. Because testosterone, but not oestradiol, uncouples two human transitional cell carcinoma lines, Kihara *et al*. (1990) suggested that the androgen might be an important factor in the higher incidence of urothelial cancer in males. Because communication occurs very efficiently between some tumour subpopulations and poorly between others, growth signals transmitted via gap junctions may be nonrandomly distributed throughout a tumour. Studies of mixtures of tumour cells in collagen gel culture suggest that, if cells within the mammary gland are extensively coupled, first messages, such as hormones and growth factors, might need to act directly on very few cells ($< 10\%$) to induce an effect throughout the gland (Miller, B. E., *et al*., 1986). Control of tissue in which only a few cells respond directly to a diffusible factor might be accomplished by gap junction coupling of receptor-positive and receptor-negative cells. An *in vitro* model which demonstrates this possibility uses a mixture of rat ovarian granulosa cells and mouse myocardial cells (Lawrence *et al*., 1978). Rat ovarian granulosa cells produce plasminogen activator in response to follicle-stimulating hormone but make no response to noradrenaline, whereas myocardial cells change beat frequency and action potential in response to noradrenaline but do not respond to follicle-stimulating hormone. In mixed cultures, follicle-stimulating hormone changes the beat frequency and action potential of myocardial cells and noradrenaline induces granulosa cells to produce plasminogen activator.

It is possible that GJIC in a heterogeneous tissue allows exchange of small molecules between cell types unable to communicate by GJIC directly. Although bovine mammary epithelial and fibroblasts both express cx43 and are able to self-communicate, GJIC between epithelial and fibroblastic cells does not occur. An additional cell type with properties of both epithelium (E-cadherin expression) and fibroblasts (no cytokeratin but express vimentin) also expresses cx43 and is able to communicate with both epithelium and fibroblasts. When all three cell types co-exist, GJIC occurs throughout the monolayer so that small molecules can be passed via the

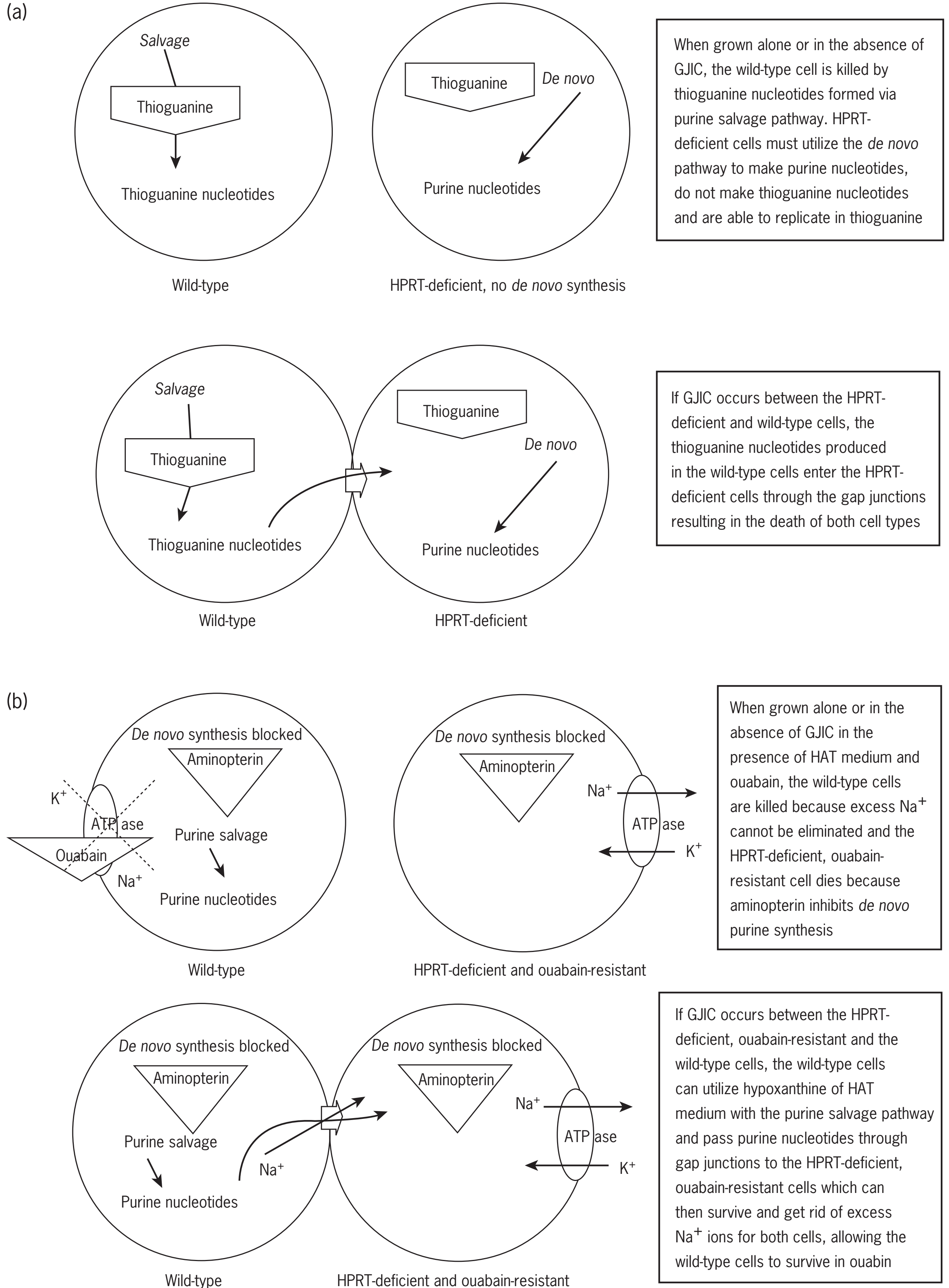


Figure 4 Metabolic cooperation assays for GJIC. (a) The kiss of death and (b) the reciprocal kiss of life.

intermediate cell between fibroblasts and epithelium (Woodward *et al*., 1998). If fibroblasts are transfected to express E-cadherin, direct GJIC with epithelium still does not occur.

Juxtacrine Signalling

Juxtacrine signalling requires contact between a receptor-positive cell and one which expresses the appropriate membrane-bound ligand. For example, the growth factors of the EGF family initially exist with a transmembrane domain and are proteolytically cleaved to form the soluble forms. In the transmembrane form, they act as both an adhesion molecule and growth factor in a juxtacrine manner. Through cell–cell contact the transmembrane form of the growth factor can activate the appropriate receptor on a neighbouring cell (**Figure 5**). Juxtacrine signalling was first demonstrated by pairing a bone marrow stromal cell line transfected with TGFα with a haematopoietic cell line transfected with EGFR to achieve both adhesion of haematopoietic cells to the stromal cells and a mitogenic response in the haematopoietic cell line (Anklesaria *et al*., 1990). Addition of EGF inhibited the adhesion. Tumour necrosis factor also exists as both a membrane integrated and soluble form, allowing both juxtacrine and paracrine signalling (Haas *et al*., 1999). Juxtacrine signalling allows very rigorous control so that the signal is delivered only to specific cells. One can imagine that shared adhesion molecules as well as a match between growth factor on one cell and receptor on the other provide a means to deliver precisely homeostatic/differentiation signals to the appropriate cell. Paracrine systems are somewhat promiscuous because no requirement for response past the expression of the growth factor receptor is necessary.

Cell adhesion molecules may also act as signalling receptors. A diverse system of transmembrane glycoproteins have been identified that mediate the cell–cell and the cell–extracellular matrix adhesion. The main families of adhesion molecules are the cadherins, integrins and selectins. The observation that the fibroblast growth factor receptor (FGFR) contains an evolutionarily conserved sequence with homology to adhesion molecules such as N-cadherin suggests they might interact with, and signal via, FGFR tyrosine kinases (Doherty *et al*., 1996). The E-cadherin–catenin complex transduces signals from the microenvironment to other molecular complexes. Signal transduction pathways starting from pp60src include E-cadherin-associated β-catenin (Noe *et al*., 1999). In addition to interactions with receptor tyrosine kinases, signal transduction via receptor protein tyrosine phosphatases (RPTPs) may be involved. The extracellular domain of RPTPs can mediate either homophilic or heterophilic interactions and a role for cadherin-mediated cell–cell adhesion has been suggested (Zondag and Moolenaar, 1997). RPTPs transduce extracellular signals by dephosphorylating tyrosine-phosphorylated intracellular substrates. For example, RPTP alpha activates pp60c-src due to direct dephosphorylation of the regulatory Tyr residue at position 527 in pp60c-src (den Hertog *et al*., 1993). Regulation of the src-related fyn protein activation by RPTP alpha appears to be important in neuronal–glial cell interactions (Zeng *et al*., 1999) and knockout mice deficient for the LAR RPTP express a deficiency in mammary gland differentiation such that lactation does not occur (Schaapveld *et al*., 1997).

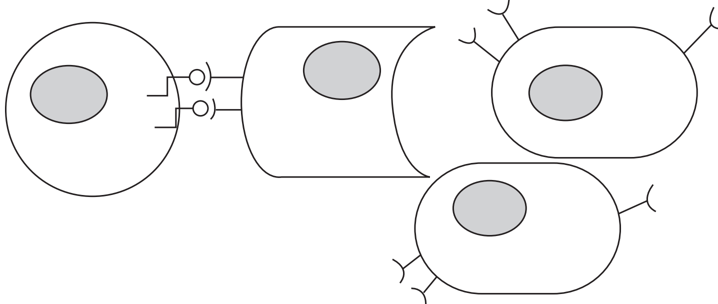
The juxtacrine mechanism may be far more important than we realize. We know of the factors which are active when cleaved from the membrane because an activity can be measured. In fact, detection and identification of the active growth factor preceded the realization that membrane-bound forms are also active in a juxtacrine configuration if cell contact is allowed. It is far more difficult to detect cleaved forms which are inactive. Perhaps the so-called orphan receptors do not have soluble ligands but act only in a juxtacrine fashion. If so, cell co-cultures or cell membranes should be screened for active ligands rather than conditioned media.

METASTASIS

The metastatic process is a complex sequence of events which involve stromal–epithelial interactions at every stage. Invasion by tumour cells in the primary site is aided by proteases made by stromal cells, sometimes in response to signals from the cancer cells. Intravasation is aided by the production of new blood vessels. Transport is a traumatic event for cancer cells moving through the circulation and survival may be aided by the formation of heterotypic emboli of cancer cells and host cells. Arrest and extravasation involves interactions between tumour cells and endothelium. Growth of extravasated cells to form metastatic deposits requires an appropriate environment, a phenomenon referred to as seed and soil. (See also chapter on *Tumour Metastasis Models*.)

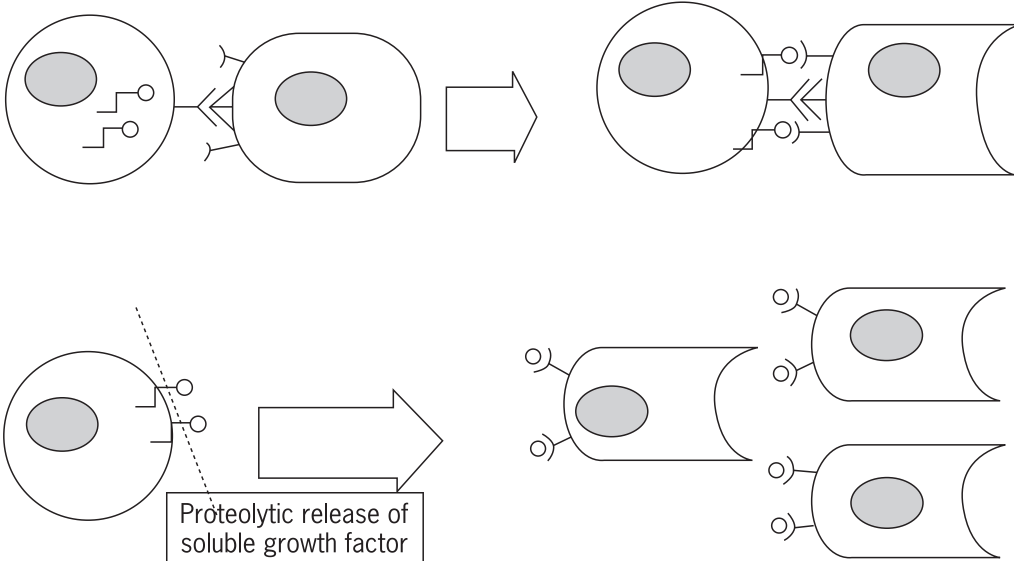
Invasion and Intravasation

Tumour cells make a variety of proteases necessary for extracellular matrix degradation and invasion but stromal cells, primarily myofibroblasts which make up the bulk of the reactive tumour stroma, also contribute several important proteases, including collagenase 1, matrix metalloproteinase 2 (MMP2) and metalloproteinase 9 (MMP9), stromelysin 1, stromelysin 3, urokinase-type plasminogen activator and cathepsin D. Tumour cells induce production by the stromal cells via release of soluble factors such as EGF, basic fibroblast growth factor and TNFα. Stromal cells other than myofibroblasts may aid in extracellular matrix degradation. In colorectal cancer, MMP2 mRNA is expressed in both peritumoural stromal and epithelial cells

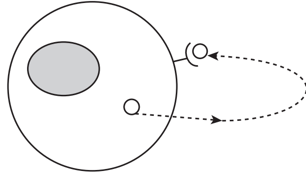


Juxtacrine interaction. Only the cell with specific receptor and in contact with cell producing the membrane-bound growth factor is stimulated.

Reciprocal juxtacrine interaction. The illustration below depicts a case in which an additional recognition molecule is necessary for production and/or transport of the growth factor to the membrane. This mechanism would allow very restrictive regulatory control to insure that only the appropriate cell expressing the appropriate adhesion/recognition molecule as well as the growth factor receptor is stimulated.



Paracrine interactions require only the presence of the growth factor receptor for response to growth factors which diffuse through the tissue. This mechanism is contact independent and allows the factor-producing cell to stimulate many more cells than possible in juxtacrine interactions but is less well controlled.



Autocrine stimulation can occur when a cell which produces a growth factor also expresses the receptor for that factor. This can occur without actual externalization of the growth factor. (See chapters *Signalling by Cytokines*; *Signalling by Tyrosine Kinases*; *Signalling by TGF beta*.)

Figure 5 Juxtacrine, paracrine and autocrine signalling.

but MMP9 is localized only in peritumoural stroma and is particularly expressed by tumour-associated macrophages at the tumour–stroma interface (Zeng and Guillem, 1996). Urokinase-type plasminogen activator in breast cancer is primarily a product of stromal myoepithelial cells, not cancer cells (Nielsen *et al*., 1996).

In some cases stroma may induce protease production by tumour cells. Mixed rat embryo fibroblasts and human cancer cell lines, neither of which produce MMP9 alone, resulted in the production of human MMP9. Fibroblast-conditioned medium contained a 30–100-kDa protein responsible for induction of MMP9 (Himelstein and Muschel, 1996).

Cathepsin D was not expressed in benign lesions (fibrocystic disease, fibroadenomas), but was positive in both the stroma and epithelium in about one half of breast cancers. About 15% of carcinoma *in situ* lesions were positive but production was found to be restricted to the epithelium (Zheng *et al*., 1999). Thus, expression in the stroma may be an important step in progression to malignancy. However, another study found that cathepsin D was increased in invasive cancer but was restricted to the epithelial component. Stromelysin 3 was mostly expressed in peritumoural stroma, and urokinase-type plasminogen activator was detected in both cancer cells and stromal cells. In normal and benign breast lesions, all three proteases were primarily found in epithelial cells, not stromal cells (Escot *et al*., 1996).

Tumour-infiltrating macrophage may result in increased release of tumour cells by releasing lysosomal hydrolases that cause necrosis, and the products of necrosis may cause an increase in the motility of neoplastic cells. Macrophages, neutrophils and lymphocytes have all been reported to induce collagenase activity *in vitro* of both cancer stromal fibroblasts and cancer cells. One should recognize that the neovascularization process in response to angiogenic factors is itself an invasive process that could augment malignant cell invasiveness. Because tumour vasculature often includes sinuses with incomplete basement membranes and because this vascularized tissue coexists with degradative enzyme-rich zones of necrosis, intravasation is greatly encouraged.

Transport

Aggregates of cells are more metastatic than single tumour cells. This may be due to more efficient arrest of the larger emboli or to protection of the inner cells to shear forces during transport. Emboli often consist of several cancer cells, but heterotypic emboli of tumour cells and lymphocytes, leucocytes or platelets might also enhance metastasis.

Arrest and Extravasation

In some cases the arrest of tumour cells is organ specific, suggesting that cell-surface receptors may exist on tumour cells and target organ endothelium. *In vivo* phage display has revealed that a number of peptide motifs selectively homed (3–35-fold increased localization over background) to the endothelium of specific organs (Rajotte *et al*., 1998). It has been reported that tumour cells and endothelial cells form gap junctions prior to extravasation and that the expression level of the receptor/ligand pair that mediates adhesion between tumour cells and the endothelium is rate limiting (el-Sabban and Pauli, 1995).

Seed and Soil

Depending upon the organ site of the primary cancer, certain organs are most likely to be involved with metastatic disease. Vascular connections are clearly very important in determining which sites have metastases, as increasing numbers of tumour cells which reach an organ increase the likelihood of the establishment of proliferative metastatic nodules. Preferential arrest due to organ-specific endothelial cell adhesion molecules may determine metastatic sites and there are differential abilities of organs to provide an appropriate environment (soil) for growth of the tumour cell (seed) and different seeds have specific soil requirements found only in certain organs. Stroma may promote metastasis by increasing cancer cell migration and invasion in addition to stimulating growth. For example, osteonectin produced by bone marrow stroma is chemotactic for breast and prostate cancer cells, both of which commonly metastasize to bone, but not for cancer cell types which do not metastasize to bone (Jacob *et al*., 1999). Of course, the metastatic tumour cells release a number of factors which alter the host organ. In particular, bone structure may be altered significantly. Breast cancer forms osteolytic metastases whereas prostate cancer forms osteoblastic metastases.

Dormancy and Recurrence

Therapy sometimes appears to cure cancer but recurrences of dormant tumours, both at the surgical site and at distant metastatic sites, may occur years later. It is not known whether dormancy represents a balance of tumour cell replication and death or whether the tumour cells simply go into an extended period of cell cycle arrest. It has been suggested that the immune system is responsible for holding these malignancies in a dormant state and that subsequent weakening of immunity, possibly via an additional disease, emotional stress or simply ageing, allows the deposits of tumour cells to regrow. Other evidence suggests that extracellular matrix components may be involved in allowing tumour cells to avoid apoptosis. Vitronectin inhibits topoisomerase-induced apoptosis in glioma cells (Uhm *et al*., 1999) and $\beta 1$ integrin-binding suppressed chemotherapy-induced apoptosis in small cell lung carcinoma (Sethi *et al*., 1999). Tumour dormancy in

human cancer xenografts was induced by low-level expression of urokinase receptor and involved reduced integrin $\alpha5\beta1$ binding to fibronectin and subsequent low basal levels of extracellular regulated kinase (ERK) (Hiragun *et al*., 1985).

CONCLUSION

The complexity of cancer is far greater than described in this chapter. New factors and interactions are being discovered rapidly. Cytotoxic agents are still the first line of defence following surgical ablation of primary cancers, but the understanding of cellular and matrix interactions provides new therapeutic targets. Anti-angiogenic factors and inhibitors of proteases are obvious candidates for chemoprevention and therapeutic targets. Chemopreventive agents such as retinoic acid altered the composition of the extracellular matrix in rat mammary glands with increased tenascin and fibronectin but decreased laminin (Schedin *et al*., 1995). Effects on GJIC are also a likely target. The chemopreventive agent all-*trans*-retinoic acid increases connexin 43 in renal cell cancer and prevents disruption of GJIC by promoters and carcinogens *in vitro* (Watanabe *et al*., 1999). Perhaps the chemotherapeutic index could be improved if $\beta1$ integrin-mediated protection from apoptosis could be blocked. Tumour cells might be held in the dormant stage by reagents which interfere with plasminogen activator receptor-$\beta1$ activation of ERK. The complex nature of cancer tissue provides ample opportunities to intercede in the disease. However, the same complexity allows for multiple paths of cancer evolution. If therapy blocks one route, a variant cell which is not dependent upon that mechanism takes a different route in progression. Nevertheless, it may be possible to identify relatively few interactions which are bottlenecks in cancer progression, allowing efficient therapy.

REFERENCES

Abrams, T. J., *et al*. (1998). Changes in the parous rat mammary gland environment are involved in parity-associated protection against mammary carcinogenesis. *AntiCancer Research*, **18**, 4115–4121.

Adams, E. F., *et al*. (1988). Paracrine influence of human breast stromal fibroblasts on breast epithelial cells: secretion of a polypeptide which stimulates reductive 17β-oestradiol dehydrogenase activity. *International Journal of Cancer*, **42**, 119–122.

Anklesaria, P., *et al*. (1990). Cell–cell adhesion mediated by binding of membrane-anchored transforming growth factor α to epidermal growth factor receptors promotes cell proliferation. *Proceedings of the National Academy of Science of the USA*, **87**, 3289–3293.

Atherton, A. J., *et al*. (1994). Ectoenzyme regulation by phenotypically distinct fibroblast sub-populations isolated from the human mammary gland. *Journal of Cell Science*, **107**, 2931–2939.

Brem, S. S., *et al*. (1978). Angiogenesis as a marker of preneoplastic lesions of the human breast. *Cancer*, **41**, 239–244.

Chen, C., *et al*. (1995). A strategy to discover circulating angiogenesis inhibitors generated by human tumours. *Cancer Research*, **55**, 4230–4233.

Cullen, K. J., *et al*. (1991). Growth factor messenger RNA expression by human breast fibroblasts from benign and malignant lesions. *Cancer Research*, **51**, 4978–4985.

Daly, R. J., *et al*. (1991). Autocrine production of insulin-like growth factor II using an inducible expression system results in reduced estrogen sensitivity of MCF-7 human breast cancer cells. *Cell Growth Differentiation*, **2**, 457–464.

den Hertog, J., *et al*. (1993). Receptor protein tyrosine phosphatase alpha activates pp60c-src and is involved in neuronal differentiation. *EMBO Journal*, **12**, 3789–3798.

Doherty, P., *et al*. (1996). Shared cell adhesion molecule (CAM) homology domains point to CAMs signalling via FGF receptors. *Perspectives in Developmental Neurobiology*, **4**, 157–168.

Escot, C., *et al*. (1996). Cellular localisation by *in situ* hybridisation of cathepsin D, stromelysin 3, and urokinase plasminogen activator RNAs in breast cancer. *Breast Cancer Research and Treatment*, **38**, 217–226.

Esinduy, C. B., *et al*. (1995). *In vitro* growth inhibition of neoplastically transformed cells by non-transformed cells: requirement for gap junctional intercellular communication. *Carcinogenesis*, **16**, 915–921.

Fentiman, I. J., *et al*. (1979). Junctional intercellular communication pattern of cultured human breast cancer cells. *Cancer Research*, **39**, 4739–4743.

Ferguson, J. E., *et al*. (1990). Tenascin distribution in the normal human breast is altered during the menstrual cycle and in carcinoma. *Differentiation*, **42**, 199–207.

Fukumura, D., *et al*. (1998). Tumour induction of VEGF promoter activity in stromal cells. *Cell*, **94**, 715–725.

Getzenberg, R. H., *et al*. (1991). Modifications of the intermediate filament and nuclear matrix networks by the extracellular matrix. *Biochemical and Biophysical Research Communications*, **179**, 340–344.

Gibson, D. F., *et al*. (1994). The mitogenic effects of transforming growth factors beta 1 and beta 2 in C3H/10T1/2 cells occur in the presence of enhanced gap junctional communication. *Cell Growth and Differentiation*, **5**, 687–696.

Gospodarowicz, D., *et al*. (1978). Determination of cellular shape by the extracellular matrix and its correlation with the control of cellular growth. *Cancer Research*, **38**, 4155–4171.

Haas, E., *et al*. (1999). Continuous autotropic signaling by membrane-expressed tumour necrosis factor. *Journal of Biological Chemistry*, **274**, 18107–18112.

Hasebe, T., *et al*. (1997). Significance of basic fibroblast growth factor and fibroblast growth factor receptor protein expression in the formation of fibrotic focus in invasive ductal

carcinoma of the breast. *Japanese Journal of Cancer Research*, **88**, 877–885.

Haslam, S. Z. and Counterman, L. J. (1991). Mammary stroma modulates hormonal responsiveness of mammary epithelium *in vivo* in the mouse. *Endocrinology*, **129**, 2017–2023.

Himelstein, B. P. and Muschel, R. J. (1996). Induction of matrix metalloproteinase 9 expression in breast carcinoma cells by a soluble factor from fibroblasts. *Clinical and Experimental Metastasis*, **14**, 197–208.

Hiragun, A., *et al*. (1985). Isolation of two syngeneic cell lines from a rat mammary carcinoma: growth factor production by neoplastic epithelial cells. *Journal of the National Cancer Institute*, **75**, 471–482.

Hofler, P., *et al*. (1993). TGF-beta induces an inhibitory effect of normal cells directed against transformed cells. *International Journal of Cancer*, **54**, 125–130.

Humphreys, R. C., *et al*. (1997). Mammary gland development is mediated by both stromal and epithelial progesterone receptors. *Molecular Endocrinology*, **11**, 801–811.

Jacob, K., *et al*. (1999). Osteonectin promotes prostate cancer cell migration and invasion: a possible mechanism for metastasis to bone. *Cancer Research*, **59**, 4453–4457.

Kavanagh, T. J., *et al*. (1987). Flow cytometry and scrape-loading/dye transfer as a rapid quantitative measure of intercellular communication *in vitro*. *Cancer Research*, **47**, 6046–6051.

Kedeshian, P., *et al*. (1998). Humatrix, a novel myoepithelial matrical gel with unique biochemical and biological properties. *Cancer Letters*, **123**, 215–226.

Kenyon, B. M., *et al*. (1996). A model of angiogenesis in the mouse cornea. *Investigative Ophthalmology and Vision Science*, **37**, 1625–1632.

Kihara, K., *et al*. (1990). Inhibitory effect of testosterone on gap junctional intercellular communication of human transitional cell carcinoma cell lines. *Cancer Research*, **50**, 2848–2852.

Koh, J., *et al*. (1998). Stimulation of human tumour xenograft growth by local estrogen biosynthesis in stromal cells. *AntiCancer Research*, **18**, 2375–2380.

Laird, D. W. (1996). The life cycle of a connexin: gap junction formation, removal, and degradation. *Journal of Bioenergetics and Biomembranes*, **28**, 311–318.

Lawrence, T. S., *et al*. (1978). Transmission of hormonal stimulation by cell-to-cell communication. *Nature*, **272**, 501–502.

Lee, Y. J. and Streuli, C. H. (1999). Extracellular matrix selectively modulates the response of mammary epithelial cells to different soluble signaling ligands. *Journal of Biological Chemistry*, **274**, 22401–22408.

Leung, H. Y., *et al*. (1997). Keratinocyte growth factor expression in hormone insensitive prostate cancer. *Oncogene*, **15**, 1115–1120.

Madhukar, B. V., *et al*. (1989). Altered regulation of intercellular communication by epidermal growth factor, transforming growth factor-beta and peptide hormones in normal human keratinocytes. *Carcinogenesis*, **10**, 13–20.

Masson, R., *et al*. (1998). In vivo evidence that the stromelysin-3 metalloproteinase contributes in a paracrine manner to epithelial cell malignancy. *Journal of Cell Biology*, **140**, 1535–1541.

Matsumoto-Taniura, N., *et al*. (1999). Prostaglandin production in mouse mammary tumour cells confers invasive growth potential by inducing hepatocyte growth factor in stromal fibroblasts. *British Journal of Cancer*, **81**, 194–202.

Miller, B. E., *et al*. (1986). Metabolic cooperation between mouse mammary tumour subpopulations in three-dimensional collagen gel cultures. *Cancer Research*, **46**, 89–93.

Miller, F. R. (1981). Comparison of metastasis of mammary tumours growing in the mammary fatpad versus the subcutis. *Invasion and Metastasis*, **1**, 220–226.

Miller, F. R., *et al*. (1981). Preferential growth of mammary tumours in intact mammary fatpads. *Cancer Research*, **41**, 3863–3867.

Miller, F. R., *et al*. (1989). Growth regulation of mouse mammary tumor cells in collagen gel cultures by diffusible factors produced by normal mammary gland epithelium and stromal fibroblasts. *Cancer Research*, **49**, 6091–6097.

Mor, G., *et al*. (1998). Macrophages, estrogen and the microenvironment of breast cancer. *Journal of Steroid Biochemistry and Molecular Biology*, **67**, 403–411.

Nielsen, B. S., *et al*. (1996). Messenger RNA for urokinase plasminogen activator is expressed in myofibroblasts adjacent to cancer cells in human breast cancer. *Laboratory Investigations*, **74**, 168–177.

Noe, V., *et al*. (1999). Extracellular regulation of cancer invasion: the E-cadherin–catenin and other pathways. *Biochemical. Society Symposia*, **65**, 43–62.

Noel, A., *et al*. (1998). Inhibition of stromal matrix metalloproteases: effects on breast-tumour promotion by fibroblasts. *International Journal of Cancer*, **76**, 267–273.

Ogmundsdottir, H. M., *et al*. (1993). Effects of lymphocytes and fibroblasts on the growth of human mammary carcinoma cells studied in short-term primary cultures. *In Vitro Cellular and Developmental Biology in Animals*, **29A**, 936–942.

Price, J. E., *et al*. (1990). Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice. *Cancer Research*, **50**, 717–721.

Rajotte, D., *et al*. (1998). Molecular heterogeneity of the vascular endothelium revealed by *in vivo* phage display. *Journal of Clinical Investigation*, **102**, 430–437.

Rasmussen, A. A. and Cullen, K. J. (1998). Paracrine/autocrine regulation of breast cancer by the insulin-like growth factors. *Breast Cancer Research and Treatment*, **47**, 219–233.

Sakakura, T., *et al*. (1979). Persistence of responsiveness of adult mouse mammary gland to induction by embryonic mesenchyme. *Developmental Biology*, **72**, 201–210.

Schaapveld, R. Q., *et al*. (1997). Impaired mammary gland development and function in mice lacking LAR receptor-like tyrosine phosphatase activity. *Developmental Biology*, **188**, 134–146.

Schedin, P. J., *et al*. (1995). Treatment with chemopreventive agents, difluoromethylornithine and retinyl acetate, results in altered mammary extracellular matrix. *Carcinogenesis*, **16**, 1787–1794.

Sethi, T., *et al*. (1999). Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: a mechanism for small cell lung cancer growth and drug resistance *in vivo*. *Nature Medicine*, **5**, 662–668.

Sidky, Y. A. and Auerbach, R. (1975). Lymphocyte-induced angiogenesis: a quantitative and sensitive assay of the graft-vs.-host reaction. *Journal of Experimental Medicine*, **141**, 1084–1100.

Singer, C., *et al*. (1995). Malignant breast epithelium selects for insulin-like growth factor II expression in breast stroma: evidence for paracrine function. *Cancer Research*, **55**, 2448–2454.

Singh, A., *et al*. (1995). IL-6sR: release from MCF-7 breast cancer cells and role in regulating peripheral oestrogen synthesis. *Journal of Endocrinology*, **147**, R9–12.

Spanakis, E. and Brouty-Boye, D. (1997). Discrimination of fibroblast subtypes by multivariate analysis of gene expression. *International Journal of Cancer*, **71**, 402–409.

Stein, L. S., *et al*. (1991). Rat ovarian granulosa cell culture: a model system for the study of cell–cell communication during multistep transformation. *Cancer Research*, **51**, 696–706.

Uhm, J. H., *et al*. (1999). Vitronectin, a glioma-derived extracellular matrix protein, protects tumor cells from apoptotic death. *Clinical Cancer Research*, **5**, 1587–1594.

Vanhamme, L., *et al*. (1989). Inhibition of gap-junctional intercellular communication between epithelial cells transformed by the activated H-ras-1 oncogene. *Experimental Cell Research*, **180**, 287–301.

van Roozendaal, K.E.P., *et al*. (1996). Differential regulation of breast tumor cell proliferation by stromal fibroblasts of various tissue sources. *International Journal of Cancer*, **65**, 120–125.

Watanabe, J., *et al*. (1999). All-*trans*-retinoic acid enhances gap junctional intercellular communication among renal epithelial cells *in vitro* treated with renal carcinogens, *European Journal of Cancer*, **35**, 1003–1008.

Wiesen, J. F., *et al*. (1998). Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. *Development*, **126**, 335–344.

Woodward, T. L., *et al*. (1998). Deficient epithelial–fibroblast heterocellular gap junction communication can be overcome by co-culture with an intermediate cell type but not by E-cadherin transgene expression. *Journal of Cell Science*, **111**, 3529–3539.

Yee, D., *et al*. (1991). The insulin-like growth factors, their receptors, and their binding proteins in human breast cancer. *Cancer Treatment Research*, **53**, 93–106.

Zeng, L., *et al*. (1999). Protein tyrosine phosphatase alpha (PTPalpha) and contactin form a novel neuronal receptor complex linked to the intracellular tyrosine kinase fyn. *Journal of Cell Biology*, **147**, 707–714.

Zeng, Z. S. and Guillem, J. G. (1996). Colocalisation of matrix metalloproteinase-9-mRNA and protein in human colorectal cancer stromal cells. *British Journal of Cancer*, **74**, 1161–1167.

Zheng, W. Q., *et al*. (1999). A comparison of the pattern of cathepsin-D expression in fibroadenoma, fibrocystic disease, preinvasive and invasive ductal breast carcinoma. *Pathology*, **31**, 247–251.

Zondag, G. C. and Moolenaar, W. H. (1997). Receptor protein tyrosine phosphatases: involvement in cell–cell interaction and signaling. *Biochimie*, **79**, 477–483.

FURTHER READING

Bruzzone, R., *et al*. (1996). Connections with connexins: the molecular basis of direct intercellular communication. *European Journal of Biochemistry*, **238**, 1–27.

Jones, J. L., *et al*. (1992). Alteration of stromal protein and integrin expression in breast – a marker of premalignant change? *Journal of Pathology*, **167**, 399–406.

Loewenstein, W. R. (1979). Junctional intercellular communication and the control of growth. *Biochimica Biophysica Acta*, **560**, 1–65.

Miller, F. R. and Heppner, G. H. (1987) Interaction of mammary tumour subpopulations. In: Medina, D., *et al*. (eds), *Cellular and Molecular Biology of Mammary Cancer*, 141–162 (Plenum Publishing, New York).

Miller, F. R. (1993) Immune mechanisms in the sequential steps of metastasis. *CRC Critical Reviews in Oncogenesis*, **4**, 293–311.

Pienta, K. J., *et al*. (1993). The tissue matrix and the regulation of gene expression in cancer cells. In: Bittar, E. E. and Heppner, G. H. (eds), *Advances in Molecular and Cell Biology*. 131–156 (JAI Press, Greenwich, CT).

Robinson, G. W., *et al*. (1999). Regulation of mammary gland development by tissue interaction. *Journal of Mammary Gland Biology and Neoplasia*, **4**, 9–19.

Weisberg, E., *et al*. (1997). Role of focal adhesion proteins in signal transduction and oncogenesis. *CRC Critical Reviews in Oncogenesis*, **8**, 343–358.

Chapter 67

Models for Drug Development and Drug Resistance

Ting-Chao Chou
Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Tai-Tsung Chang
Kaohsiung Medical University, Kaohsiung, Taiwan

CONTENTS

- The Drug Development Process
- Criteria for Selecting Candidate Lead Compounds
- Models for New Drug Screening and Testing
- The High-throughput Strategy
- Genetic Approaches for Drug Development
- Resistance to Anticancer Drugs
- Multidrug Resistance (MDR)
- Expression of *MDR1* Gene in Health and Disease
- Therapeutic Application of the *MDR1* Gene
- Apoptosis and MDR
- Reversal of Anticancer MDR
- Assessment of MDR Reversing Agents
- Multi-hit and Multi-target Approach Against MDR: Drug Combinations
- Quantitation of Synergism and Antagonism
- Experimental Design for Drug Combinations
- Selectivity of Synergism and Antagonism
- Rational Approach to the Clinical Protocol Design for Drug Combinations

THE DRUG DEVELOPMENT PROCESS

There are two basic elements for anticancer drug development: efficacy and selectivity. The search for highly efficacious compounds may be achieved by a rational approach or an empirical approach. After the selection of a candidate compound with *in vitro* and preliminary *in vivo* testing, extensive preclinical pharmacological and toxicological studies need to be performed to assess the possible therapeutic potential. The toxicological profile and selectivity of effect play a crucial role in determining the success or failure of a drug development.

In the United States, the Food and Drug Administration (FDA) plays an important role in each stage of evaluation, monitoring and approval from investigational new drug (IND) application for phase I to phase III clinical trials, to new drug application (NDA) and to postmarketing surveillance (Food and Drug Administration, 1995). The approximate time period required for each step of the process is given in **Table 1**.

The evaluation of new anticancer agents starts from *in vitro* experimentation, usually with small quantities of compounds. Preliminary testing in rodents allows the selection of the lead candidate compound(s) for further preclinical pharmacological and toxicological studies (**Table 1**) using mice, rats and dogs. At this stage, scale-up synthesisis required and compliance with the regulatory guidelines for good manufacturing practice (GMP) and good laboratory practice (GLP) should be enforced. In clinical trials, the guidelines for good clinical practice (GCP) need to be followed.

An IND application to the FDA consists of the following: an introductory background statement; general investigational plan; the clinical trial protocol including investigator data, facility data and institutional review board (IRB) data; chemistry, manufacturing and control data; environmental assessment; pharmacological and toxicological data; and previous human experience. The following information is usually part of the protocol: study design, criteria for patient eligibility, recruitment

Table 1 The new drug development process

Preliminary testing	Preclinical research and development	Clinical studies	NDA review	Postmarketing surveillance
Time range Vary	Range: 1–3 years Average: 18 months	Range: 2–10 years Average: 5 years	Range: 2 months–7 years Average: 24 months	Adverse reaction reporting
		Phase 1 →		
Discovery and evaluation		Phase 2 →		
Design		Phase 3 →		Survey
Targeting	Scale-up Synthesis and purification			Sampling
Modelling	Animal testing: Pharmacokinetics	Accelerated development and review →	→	Testing Inspections
Synthesis Mechanism	Toxicology Short term →			Sponsor answers any question from review
In vitro and preclinical *in vivo* testing	Long term →	→		
		↑ IND Submission to FDA 30-day safety review	↑ NDA Submission to FDA	↑ NDA Approval by FDA

plan, pretreatment evaluation, treatment/intervention plan, grading toxicities/side effects, criteria for therapeutic response/outcome assessment, criteria for removal from study, biostatistics, subject registration and randomization procedures, quality assurance, protection of human subjects, informed consent procedures and references.

When an IND application is submitted to the FDA, there will be a 30-day safety review period before phase I clinical trials can be started (**Table 1**). The phase I trial is mainly a dose-finding study to explore the safety, the tolerated doses, pharmacokinetics and toxicities. It usually enlists 20–100 patients and takes several months to complete the study. The phase II trial uses up to several hundred patients with specified tumour type(s). In this phase, the aims are to determine therapeutic effects, toxicities and any unexpected complications. The study may take up to 2 years to complete. The phase III trial is an expanded study enlisting hundreds to several thousand patients and usually involves multiple medical centres. Efficacy and toxicity are evaluated for a range of tumours under specified regimens of treatment. The phase III trial may take 2–4 years to complete (**Table 2**). When safety and therapeutic efficacy are established, the new drug application (NDA) can be submitted to the FDA for review and approval. Postmarketing surveillance is an important aspect for the manufacturers and FDA and involves adverse reaction reporting, sampling, testing and inspections. (See also chapter on *Translational Research*.)

The rational approach of drug development includes selecting a molecular target and the development of molecular entities of drugs. Owing to the similarity of biological and molecular make-up between tumour and normal cells, the strategies for developing agents for cancer therapy include (1) inactivation or blocking the molecular process that leads to carcinogenic transformation, (2) exploiting the differences between normal and tumour cells at the molecular, biochemical and cellular levels and gaining a selective advantage towards therapeutic aims, (3) eliminating or inhibiting tumour cells with cytotoxic agents, cytodifferentiating agents or reversing agents of multidrug resistance (MDR) and (4) approaches of combination therapy with therapeutic agents with different modalities of treatments.

Table 2 Clinical trials in humans

Phase	No. of patients	Length	Purpose	Approximate percentage of drugs successfully tested
I	20–100	Several months	Dose-finding Safety Toxicity	70
II	Up to several hundred	Months to 2 years	Short-term safety and effectiveness	33
III	Hundreds, up to several thousand	2–4 years	Safety, dosage and effectiveness	25–30

In this chapter we will concentrate on the methods and models of new drug development and overcoming the drug resistance that would affect the therapeutic end results of cancer treatment.

CRITERIA FOR SELECTING CANDIDATE LEAD COMPOUNDS

It is usually difficult to select a candidate compound for further development. The decision is of great importance since drug development takes a long period of time to test, to evaluate and to meet regulatory requirements and it might easily cost as much as 50 million dollars to market a new drug.

For every 10 000 compounds screened, approximately 500 may show antitumour activity *in vitro*. Among these, 100 may be worthy of testing in animals. Of these 100 compounds, 10 may meet the criteria to perform extended pharmacological and toxicological studies to obtain enough information to warrant clinical trials in humans. Past experience (Food and Drug Administration, 1995) suggests that for the phase I trial only 70% of drugs are successfully tested. For the phase II and phase III trials only about 33% and 25% of drugs are successfully tested, respectively (**Table 2**).

There are several criteria for selecting a candidate compound, including potency, pharmacological properties, toxicity, feasibility of formulation, and degree of difficulty of synthesizing or producing a sufficient quantity.

MODELS FOR NEW DRUG SCREENING AND TESTING

The US National Cancer Institute (NCI) has played a very important role in antitumour screening and testing for the past four decades. Its approach and policy have evolved through the decades and created models that have a long-lasting impact (Driscoll, 1984; Chabner, 1990; Johnson, 1990; Sikic, 1991).

Historical Models and Their Evolution

The NCI established the Cancer Chemotherapy National Service Center (CCNSC) to support research efforts of investigators in the discovery of new agents for cancer treatment. The main models used then were murine leukaemias, e.g. L1210 and P388, as index tumours for *in vivo* screening. This was followed by additional tumour models used to follow up the initial lead.

In 1985, the NCI changed its cancer drug-screening programme by (1) emphasizing primary *in vitro* screening using an organ-site oriented approach, (2) using human solid tumour cell lines, (3) using small amounts of compounds for the primary screening, and (4) a renewed commitment to use screening as a method for exploring natural products from diverse sources (e.g. plants, marine organisms and microbes)

During 1990–1992, the NCI Cancer Drug Screen was established. The main feature is *in vitro* exposure (48 h) of a panel of 60 human tumour cell lines at five different concentrations with 10-fold serial dilution. Crude extracts are usually evaluated at a single concentration. The SRB (sulforhodamine B) protein stain assay (see below) was used to generate 60 dose–effect curves for each compound. It should be noted from the mass-action law principle that 1000-fold dilution with five or more dose densities would be too broad to have smooth dose–effect curves, although it would provide some efficacy indications. This programme also provides: the development of the ‘mean graph’ (Paull *et al*., 1989) to interpret the relative sensitivities of each of the tumour cell lines to an individual agent and the development of computer-based ‘Compare Program’ (Paull *et al*., 1989) for recognition of these patterns of effectiveness. By comparing the correlation with a representative seed compound, such as taxol, vinblastine or adriamycin, potential clues to understanding the mechanism of antitumour action may be obtained.

For example, vinblastine, taxol and numerous compounds that target microtubules showed similar patterns in the ‘mean graph’ and a strong correlation in the ‘Compare Program,’ although inhibitors and promoters of tubulin polymerization were both present.

Molecular Targeting and Cell-based Drug Screen

In early 1990, the NCI launched a major effort to characterize and define molecular targets with the 60 human tumour cell lines (Chabner 1990; Johnson, 1990; Sikic, 1991). The data suggest that a combination of biological patterns of effectiveness with information on specific molecular targets will enhance the recognition of novel agents. So far, the molecular targets that have been characterized in the screen include *mdr-1*, *H*-, *N*- and *K-ras*, *bcl-2*, *SKC*, *p53*, *MDM2*, G_1 and G_2 checkpoints, DT diaphorase, thymidylate synthase, etc. The length of the list is increasing.

The tumour cell lines used for experiments are usually obtained from various sources: (1) generated by individual laboratories from experimental animals whether they are spontaneous tumours or carcinogen-induced tumours, and from human tumour specimens; and (2) cell lines generated by transgene procedures. Large collections of animal and human tumour cell lines are available commercially from the cell bank of the American Type Culture Collection (ATCC) (Rockville, MD, USA), but they are mainly unmodified cells. Drug-resistant cells are usually generated either *in vivo* in animals or *in vitro* in tissue culture by continuous exposure with gradually increasing sublethal doses or concentrations of an antitumour agent. Depending on the drug and tumour type, the time required for generating a drug-resistant tumour cell line ranges from several months to several years. If the selecting agent is from a natural source, frequently MDR cell lines can be obtained. So far, the panel of 60 human tumour cell lines used by the NCI are not MDR cell lines.

A Model for Solid Tumours In Vitro: Hollow-fibre Assay

The traditional xenograft models of human tumours in nude mice or scid mice are rather expensive, time consuming and require 20–500 mg of compounds, depending on the potency. In the newly adopted hollow-fibre assay system (Casciari, *et al*., 1994) human tumour cell lines are implanted within a polyvinylidene hollow fibre. The fibres are heat-sealed before implantation into i.p. or s.c. compartments in the same nude mouse. Three fibres are simultaneously inserted into both compartments in the same mouse. Thus, antitumour effect of an agent can be simultaneously evaluated against three individual tumours growing in two separate physiological compartments of the same animal. The hollow fibres are removed after 1 week to examine the tumour cell viability using SRB or XTT assays (Skudiero *et al*., 1998; Skehan *et al*., 1990).

Cytotoxic effects of an antitumour agent on the growth of various tumour cell lines *in vitro* are evaluated under specified conditions. Typically, the cells are cultured at an initial density of 5×10^4 cells/mL. They are maintained in a 5% CO_2-humidified atmosphere at 37°C in RPM I medium 1640 (or other medium) containing penicillin (100 units/ml), streptomycin (l00 μg/mL) and l0% heat-inactivated fetal bovine serum. For cell suspension cultures, cytotoxicity assays are performed by using the 2,3-bis(2-methoxy-4-nitro-5-sulfophenyl)-5-(carboxanilide)-2*H*-tetrazolium hydroxide (XTT) microculture tetrazolium method (Scudiero *et al*., 1998) in duplicate in 96-well microtitre plates. XTT is prepared at 1 mg/ml in prewarmed (37°C) medium without serum. Phenazine methosulfate (PMS) and fresh XTT are mixed to obtain 0.025 mmol/L PMS–XTT solution (27 μL of the stock 5 mmol/L PMS is added per 5 mL of 1 mg/mL XTT). After a 72 h of incubation, 50 μL of assay aliquots are added to each well of the 200 μL cell culture. After incubation at 37°C for 4 hr, the absorbance at 450 nm and 630 nm is measured with a microplate reader. For the monolayer cell cultures, they are determined in 96-well microtitre plates by the SRB method described by Skehan *et al*. (1990) for measuring the cellular protein content. Cultures are fixed with trichloroacetic acid and then stained for 30 min with 0.4% SRB dissolved in 1% acetic acid. Unbound dye is removed by acetic acid washes, and the protein-bound dye is extracted with an unbuffered Tris (tris(hydroxymethyl) aminomethane) base for determination of the absorbance at 570 nm in a 96-well microtitre plate reader. The experiments are usually carried out in duplicate or triplicate. Each run entails 5–6 concentrations of the drug being tested. The IC_{50} values are determined by the median-effect equation and plotted (Chou, 1991a; Chou and Talalay, 1984) using computer software.

The Knockout Models for Multi-drug Resistance

To assist in the understanding of the mechanism of action and physiological functions of the transporter proteins, the generation of gene knockout models is often very useful. Borst and co-workers have obtained mice with single and double knockout in *mdr1a* and *mdr1b* genes (Schinkel *et al*., 1997). They found that the recognition of the P-glycoprotein (Pgp, see below) as an important component of the blood–brain barrier and its involvement in the absorption and elimination of several clinically used drugs. The fact that the multidrug resistance protein (MRP) is ubiquitously expressed, although at very low levels, in the liver and some other organs suggests that it exerts a common function. The development knockout mice for MRP and other transporters would greatly improve the understanding of their mechanisms and physiological functions.

Several physiological functions of MRP have been suggested. These include protection against environmentally present heavy metal oxyanions, modulation of the activity of ion channels, transport of leukotriene C_4

and glutathione, glucuronate and sulfate conjugates and transport of GSH through a co-transport mechanism (Rappa *et al*., 1999). The importance of MRP in the protection of normal tissue from the toxicity of anticancer agent such as etoposide had been established. A total block of MRP has been found to be compatible with life, suggesting that MRP inhibitors can be safely used for treating cancer patients (Rappa *et al*., 1999).

Recently, the transporter was found in drug-resistant cell lines with low level Pgp and MRP, but with high levels of drug transport, particularly for mitoxantrone. This was designated BCRP, MXR and ABCP by different groups of investigators (Kuska, 1999; Ross *et al*., 1999). Today, by searching a computerized database looking for sequences similar to a signature ATP-binding motif found in the known ABC genes, at least 50 human ABC genes have been found. Because the expression level of BCRP/MXR/ABCP is so low in adult tissues, it might be an excellent candidate to target for an inhibitor and enhance the effect of chemotherapy without affecting normal cells (Higgins, 1995; Kuska, 1999). (See also chapter on *Drug Resistance and Reversal*.)

THE HIGH-THROUGHPUT STRATEGY

The large number of compounds from biodiversified sources and combinatorial chemistry libraries require high-efficiency assays. In recent years, high-throughput assay systems have been developed from 96-well microplates for optical, colorimetric and fluorescence readers to microplates consisting of as large as a 1536-well format for multimode analysers. The availability of robotic systems is replacing labour-intensive manual work with increased workload and efficiency. The efficiency of drug development is also facilitated by the identification of small molecules that bind with high-affinity acceptor molecules (e.g. cell-surface receptors, enzymes, antibodies) so to mimic or block their interaction with natural ligands. Assuming that peptides can be used as antitumour agents, on the basis of a ‘one-bead, one peptide’ approach (Lam *et al*., 1991), one can greatly enhance the production and rapid evaluation of random libraries of millions of peptides so that receptor-binding ligands of high affinity can be rapidly identified and sequenced. By using the split synthesis approach, the first cycle consisted of distributing a pool of resin beads into separate reaction vessels each with a single amino acid, allowing the coupling reaction to go to completion, and then repooling the beads. This cycle was repeated several times to extend the peptide chain. In this fashion, each bead should contain only a single peptide species. To develop a rapid approach for screening the library to find beads containing peptides able to bind to any particular acceptor molecule, acceptor molecules were coupled to an enzyme (alkaline phosphatase) or to fluorescein and added in soluble form to the peptide bead library.

Typically, a few beads were intensely stained and were visible to the naked eye and easily seen with a low-power dissecting microscope (bead diameter 100–200 μm against a background of countless nonactive beads. With the aid of tiny forceps coupled to a micromanipulator, the intensely stained beads could be remove for analysis. Each bead was washed with 8 mol/L guanidine hydrochloride to remove the acceptor complex and then the peptide sequence contained on the bead was determined by placing it on a glass filter which was inserted into a peptide microsequencer. A library containing several million beads could be screened in about 10 Petri dishes in half a day. Subsequently, it could be washed with 8 mol/L guanidine hydrochloride and subsequently reused for screening new acceptors. This micromethod is likely to be extended to other applications, such as polynucleotides or desoxypolynucleotides.

GENETIC APPROACHES FOR DRUG DEVELOPMENT

The process of therapeutic drug development evolved from the experience-based herbal medicine to evidence-based pharmaceuticals. In the modern era, it has further progressed to gene-based biotechnological products and gene therapy. The combinatorial chemistry and platform technology aided by computer modelling of chemical molecules and chiral properties and structural folding hold new promise and challenges in the new millennium.

Large genetic databases now include portions of all of the human genome. A variety of technologies make it possible to catalogue the expression of large numbers of genes in specific organs and tissues and to compare healthy and disease states, and drug-sensitive and drug-resistant conditions. Disease- or pathology-associated and mutated genes isolated through biotechnologies are now numbered in the thousands. In addition, the scientific success in positional cloning has been impressive, and the number of positionally cloned disease genes is increasing (Darby, 1997). With the rapid rate of accumulation of genetic information, new drug-screening models and targets will facilitate the discovery of new and better drugs.

Expression profiling technologies and positional cloning are the two main approaches to the drug targeting (Scangos, 1997). Expression profiling technologies include subtraction hybridization, computerized comparisons of gene libraries from different tissues, serial analysis of gene expression and differential displays. Differential expressions have been correlated with disease states, but it has been difficult to determine which differentially expressed genes are cause rather than effect, and which, if targeted by a drug, could reverse the disease process. The positional cloning approaches have identified and cloned genes involved in human diseases, including cancer,

Alzheimer's disease, obesity, diabetes and asthma. The problem is that although these genes provide insight into the cause of disease, they are limited in what they can deliver for therapeutic intervention. For example, positionally cloned genes often encode proteins of unknown biochemical or physiological function, leading to a situation where it is difficult or impossible to set up screening assays for inhibitors. More important, positionally cloned genes often contribute to a disease phenotype by virtue of being defective. Since drugs most often are inhibitors of their targets, these genes most often are inappropriate screening targets. However, one may be able to use the identified gene to understand the biochemistry and regulatory process of the pathway and, thereby, identify an optional point of intervention (Skangos, 1997). An important question for drug development is what genes and proteins exist in the target cells which, when inhibited by a drug, will reverse the effects of the disease gene and restore homeostatic balance.

RESISTANCE TO ANTICANCER DRUGS

Innate or acquired resistance in chemotherapy remains a serious obstacle towards the chemotherapeutic treatment of cancer and infectious diseases. Various mechanisms of drug resistance exist which include drug metabolism (decrease in drug activation or increase in drug inactivation), transport (decrease in drug influx or increased drug efflux), drug target alteration (increased target levels, altered function of target or decreased affinity of drug for target), DNA repair (increased repair of DNA damage or mutations in mismatch repair genes) and apoptosis (altered function and regulation that change the ability of tumour cells to execute programmed cell death). Among these mechanisms, the best characterized mechanisms of MDR involves P-glycoprotein 170 (Pgp) (Kartner and Ling, 1989; Juliano and Ling, 1996; Ueda *et al*., 1999).

MULTIDRUG RESISTANCE (MDR)

The original concept of MDR was introduced in 1970 by Biedler and Riehm of the Sloan-Kettering Institute in New York (Biedler and Riehm, 1970). They found that cells resistant to drugs that had no structural or mechanistic similarities showed cross-resistance. In some cases the resistant cells expressing Pgp show collateral sensitivity to unrelated compounds and become more sensitive, when compared with the parent (non-MDR) cells (Chou *et al*., 1998). However, there is apparently no unifying theory for a simple explanation of the MDR problem (Laing and Tew, 1997).

There are at least three types of MDR that have been characterized: (i) classical Pgp-mediated MDR, in which the causal relationship between resistance and overexpression of Pgp and resistance has been clearly demonstrated *in vitro*; (ii) atypical MDR, which involves drugs that interact with DNA topoisomerase that has decreased enzyme levels and/or a mutated structure, and (iii) non-Pgp-mediated MDR, in which resistant cells do not show elevated Pgp levels of expression, but instead, with increased level of MRP, lung resistance protein (LRP) or the expression of other members of the ABC transporter gene family (Gottesman and Pastan, 1993; Kuska, 1999).

Despite an overwhelming number of papers on MDR_1 expression and Pgp activity in tissue culture systems of multidrug resistance, the relationship of MDR_1 with the clinical observation of multidrug resistance remains weak. Although a strong correlation has been obtained in haematological malignancies, it is premature to generalize the association in solid tumours (Kellen, 1994; Antoney and Kayl, 1999; Robert, 1999). The fact that MDR_1 is just one of many polypeptides in the ABC superfamily related to drug resistance may weaken the correlation since the contribution of multiple factors such as Pgp, MRP, apoptosis antagonist, Bcl-2, Bcl-X_L and compartmentalization were not simultaneously analysed.

EXPRESSION OF *MDR1* GENE IN HEALTH AND DISEASE

The identification of Pgp as an ATP-dependent pump, which could confer resistance to a variety of hydrophobic compounds cytotoxic to cancer cells, raised the possibility of normal function of Pgp (Kellen, 1994; Antoney and Kaye, 1999; Bates *et al*., 1999; Robert, 1999). Monoclonal antibody labelling of the tissue distribution of Pgp revealed the plasma membrane localization on the positive cells. For the epithelial cells of jejunum, ileum and colon, the major location was on the mucosa. For the brush border of proximal tubule cells of the kidney and the biliary face of the hepatocytes of the liver, and also placenta, the Pgp expression was most evident. Although *mdr* knockout mice are viable, mice in which *mdr* genes have been intentionally inactivated have major alterations in pharmacokinetics and tissue distribution of substrates for Pgp. The physiological role of Pgp in health and disease has therapeutic implications for pharmacological inhibition of Pgp. The importance of Pgp in MDR in cancer suggests that the efficacy of administered chemotherapeutic agents can be enhanced and MDR can be reversed by Pgp blockers. An alternative procedure to the pharmacological approach to circumvent Pgp-mediated MDR in cancer cells is to prevent the biosynthesis of transporter by selectively affecting its transcription. These include the use of antisense, ribozymes and triplex-forming oligonucleotides targeted against different regions of the *MDR1* gene. In either cases, exploitation for selectivity against

tumour over normal tissues is an important determinant for beneficial therapeutic effects.

THERAPEUTIC APPLICATION OF THE *MDR1* GENE

Cloning and sequencing (**Figure 1**) of *MDR1* cDNA provide the tool to confirm that the phenotype of MDR can be conferred on drug-sensitive cells by expression of Pgp. The possibility of using genetic materials for gene therapy has promoted the development of *MDR1* gene transfer and expression methodologies (Aran *et al*., 1999). One of the goals is to protect drug-sensitive cells (such as bone marrow progenitor cells) from cytotoxic effects in cancer chemotherapy. *In vitro* and animal models have been established. However, it remains to be seen whether the genetic advantage gained by cells expressing transgenic *MDR1* will be sufficient to fulfil the expected goals. Clinical trials have already been performed to determine the stability of expression of the *MDR1* gene and its ability to confer chemoprotection on bone marrow cells. To date, the success of these studies has been limited by the inefficiencies of delivery of genes into human bone marrow cells. As a dominant selectable marker, the *MDR1* gene still has to prove its efficacy *in vivo* to ascertain stable expression of non-selectable genes of interest by drug selection in the appropriate disease animal model (Aran *et al*., 1999).

Since the *MDR1* gene confers multidrug resistance *in vitro* and *in vivo*, *mdr1* transgenic mice could be serve as an *in vivo* model for new anticancer drug evaluation. Unfortunately, these *mdr1* transgenic mice have not lived up to the original expectation because after several generations of inbreeding, the level of expression of human *MDR1* became too low (Aran *et al*., 1999).

APOPTOSIS AND MDR

Programmed cell death (apoptosis) is mediated by degradation of DNA and a number of cellular proteins. The discovery of the Bcl family of proteins by Tsujimoto and Croce (1986) as apoptosis antagonists suggested a novel mechanism of MDR and a new approach to therapeutic interventions.

The *bcl-2* (B-cell leukaemia/lymphoma 2) gene products are one subfamily of proteins that determine how particular cells die in response to particular physiological or pharmacological environments. Bcl-2 in widely expressed in human tumours, and by inhibiting cell death it may allow tumour cells to accumulate the mutations which in turn may cause the cell to become invasive. Studies of the programmed cell death in the nematode *Caenorhabditis elegans* have provided many important leads to the understanding of apoptosis. Recent findings in the JNK pathway, cell death signalling via Fas and TNF receptors, Bcl-X_L for protecting cancer cells *in vitro*

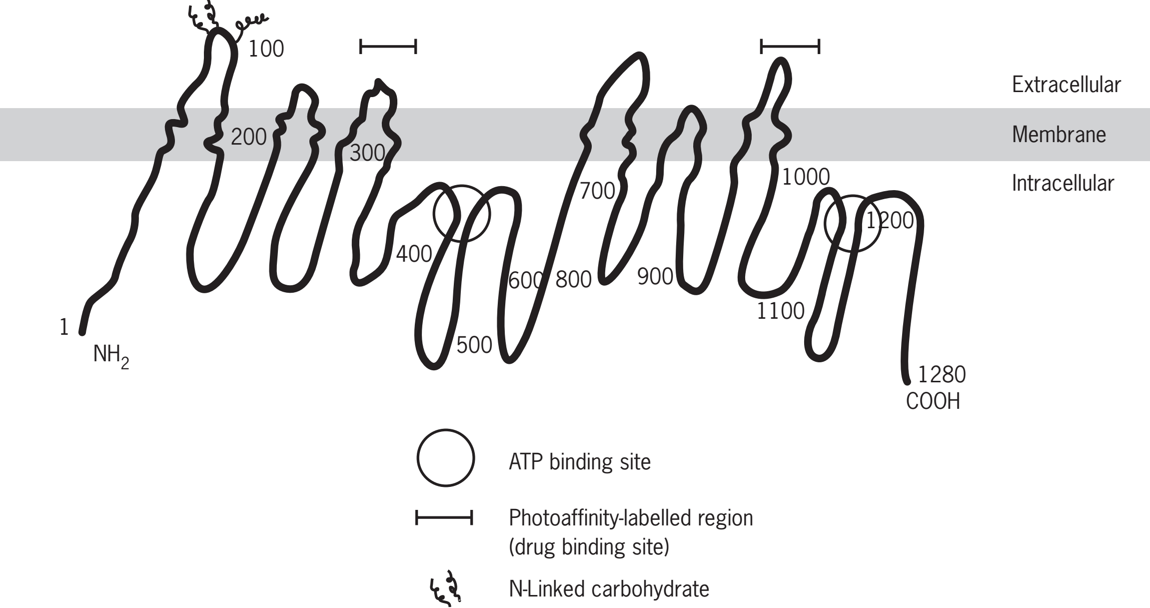


Figure 1 Model of human P-glycoprotein (Pgp). A two-dimensional drawing of Pgp with numbers referring to amino acid residues in the linear sequence. The ATP-binding domains and the putative drug-binding domains are shown for each half of the molecule. Pgp is a single polypeptide consisting of 12 transmembrane and helical segments. Putative N-linked carbohydrates, located in the first extracellular loop, are indicated in the diagram. (From Lam *et al*., 1991, Gottesman and Pastan, 1993, Kellen, 1994 and Robert, 1999.)

against drug-induced apoptosis and the enhancement of chemosensitivity of cancer cell apoptosis by over-expression of Bcl-X_s have all added to the complexity in this field (Clynes *et al.*, 1998).

REVERSAL OF ANTICANCER MDR

A plethora of agents have been developed that modify, modulate or reverse the MDR phenotype (Chou *et al.*, 1998). There are diverse classes of compounds that interact with Pgp (**Table 3**). Many anticancer agents from natural sources, such as vinblastine, vincristine, doxorubicin, paclitaxel, actinomycin D and their analogues, are MDR substrates. Other natural products, such as colchicine, rhodamine 123, erythromycin, ardeemin and gypsetin also interact with Pgp (Kellen, 1994). Steroid hormones such as aldosterone, cortisol and dexamethasone, calcium blockers such as verapamil and dilitiazem, quinolines such as quinidine, quinine and quinicrine, immosuppressants such as cyclosporin A, rapamycin and FK506, peptides such as gramicidin D and valinomycin, and calmodulin antagonists such as trifluoperazine and chlorpromazine are also modulators of Pgp. These resistance modifiers have undergone structural modification to reduce undesirable side effects, such as from verapamil to dexverapamil, and from cyclosporin A to SDZ PSC 883, (**Table 4**). Of these, verapamil, trifluoperazine, cyclosporin A, nifedipine and verapamil have been subjected to phase I/II clinical trials (Volm, 1998).

Table 3 Diverse classes of compounds that interact with Pgp

Antineoplastic drugs	Modulators/chemosensitizers
Vinblastine, vincristine	Verapamil, diltiazem, bepridil
Doxorubicin, daunomycin, mitoxantrone	Azidopine, nitrendipine, nifedipine
Bisantrene	
Etoposide, teniposide	
Topotecan	Quinidine
Mitomycin C	Cyclosporin SDZ PSC-833
Paclitaxel, taxotere	FK506
Mithramycin	Rapamycin
Actinomycin D	Terfenadine
	Trifluoperazine
Other agents	Quinine, quinicrine
Colchicine	Reserpine
Rhodamine 123	Forskolin
Ethidium bromide	
Erythromycin	
Ketoconazole	**Peptides**
Ardeemin	Gramicidin D
Diketopiperazine	Valinomycin
Gypsetin	
Steroid hormones	**Photoaffinity drug analogues**
Aldosterone	Azidopine
Cortisol	Iodoarylazidoprazosin
Tamoxifen	*N*-(*p*-Azido-3-iodophenethyl) spiperone
Megestro	
Dexamethasone	Tamoxifen aziridine

Table 4 Resistance modifiers

First generation	Second generation
Verapamil	Dexverapamil
Cyclosporin A	SDZ PSC 833
Calmodulin inhibitors	SDZ 280-446
Quinine	XR 9051
Progesterone	GF 120918
Tamoxifen	S-9788
Amiodarone	Ardeemin
Trifluoperazine	

ASSESSMENT OF MDR REVERSING AGENTS

Evaluation of a new MDR-reversing agent consists of many important aspects. Using 5-*N*-acetylardeemin (NAA) as an example (Chou *et al.*, 1998), these include the following:

1. *Reversing efficacy*: whether it completely or partially reverses drug resistance to a certain drug in a given MDR cell line. For CCRF-CEM and CCRF-CEM/VBL_{100} cells, the IC_{50} values for VBL are 0.000 83 and 0.632 μmol/L, respectively. In the presence of 15 μmol/L of NAA (IC_{50} values are 29.8 and 20.8 μmol/L, respectively), the IC_{50} for VBL became 0.000 21 and 0.000 11/μmol/L, respectively. Thus, NAA completely reversed the VBL drug resistance in CCRE-CEM/VBL_{100}. cells. The analogue of NAA 5-*N*-acetyl-8-demethylardeemin (NADMA) showed a similar effect to NAA (**Figure 2**).
2. *Improvement of therapeutic efficacy*: how much it improves the therapeutic effect *in vivo* in B6D2F_1 mice bearing P388/0 and P388/DX leukaemia. For P388/0, 2 mg/kg Q D×4 i.p. of doxorubicin increased the life span by 178%. In combination with NAA, 70 mg/kg, the increase in life span became 189%. However, in doxorubicin-resistant P388/DX tumour-bearing mice the increases in life span were 106% and 151% in the absence and presence of NAA, respectively. In nude mice bearing MX-1 mammary carcinoma xenograft, treatment with doxorubicin 1.5 mg/kg Q D×5 i.p. reduced the tumour size by 44%, whereas with doxorubicin and NAA (50 mg/kg) combination, the tumour size was reduced by 75%; two of eight mice were tumour free but one mouse died of toxicity.
3. *Toxicity toward the host*: whether it has little or no cytotoxicity or side effects when compared with other MDR-reversing agents. For NAA there was no apparent toxicity other than the reversible body weight drop at

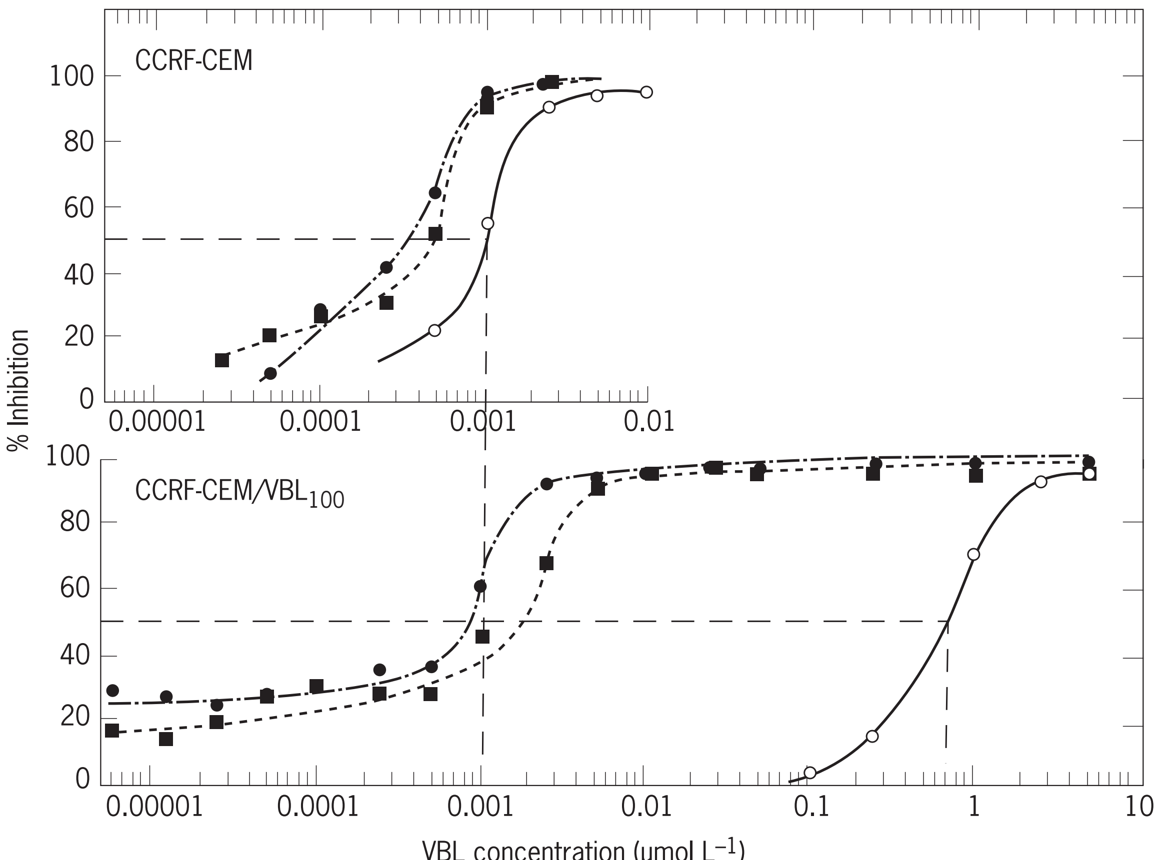


Figure 2 Dose–effect curves for vinblastine (VBL) against the cell growth of human leukaemic cells in the absence and presence of MDR reversing agents, with (○) varying [VBL] alone; (●) varying [VBL] with constant [NAA] (15 μmol/L); (■) varying [VBL] with constant [NADMA] (15 μmol/L). Top, CCRF-CEM lymphoblastic leukaemic cells; bottom, CCRF-CEM/VBL_{100} leukaemic cells 760-fold resistant to VBL. The concentrations of the ardeemins used were about half of their IC_{50} values toward CCRF-CEM cells. (From Chou *et al.*, 1998, *Proceedings of the National Academy of Sciences of the USA*, **95**, 8369–8374.)

150 mg/kg Q D×8 i.p. By contrast, verapamil at 150 mg/kg Q D×3 i.p. led to a 40% lethality towards the mice.

4. *Preferential synergism in MDR over the parent tumour cells*: in our earlier studies, we observed that NAA is slightly more potent in inhibiting CCRF-CEM/VBL_{100} cells than in inhibiting the parent CCRF-CEM cells (i.e. collateral sensitivity) (Chou *et al.*, 1998), and we conducted classical drug combination studies for VBL and NNA. As shown in **Figure 3**, VBL + NAA showed much stronger synergism in CCRF-CEM/VBL_{100} cells than in CCRF-CEM cells. These results suggest an extra therapeutic value of NAA in combination with VBL, a classical MDR substrate.
5. *Increase intracellular accumulation of drugs*: whether and how much an MDR-reversing agent increases intracellular accumulation of cancer chemotheraputic agents. Intracellular accumulation of [^{3}H]VBL in CCRF-CEM cells was increased twofold by 30 μmol/L of verapamil and threefold by 30 μmol/L of NAA.
6. *Efflux kinetics*: whether the MDR-reversing agent alters the elimination (efflux) kinetics of a drug under investigation. For the efflux kinetics for preloaded [^{3}H]VBL in CCRF-CEM cells, drug elimination was biphasic with an initial half-life of 12 min and β-phase half-life of 105 min. In the presence of 30 μmol/L of verapamil, the half-lives for the two phases were prolonged to 18 and 210 min, respectively. In the presence of 30 μmol/L of NAA, these half-lives were extended to 29 and 380 min, respectively (**Figure 4**). Thus, NAA not only resulted in a higher intracellular concentration of vinblastine but also prolonged the retention half-life of the preloaded vinblastine.
7. *Selective binding*: whether an MDR-reversing agent binds specifically to a particular MDR pump protein. Evidence that NAA interacts directly with the transporter, Pgp, was obtained by examining its effect on the inhibition of photoaffinity labelling of [^{3}H]azidopine, a photoactivatable substrate of Pgp (Sikic, 1991). The membrane functions from CCRF-CEM parent cells showed little or no [^{3}H]azidopine binding, whereas the binding to the membrane fractions isolated from

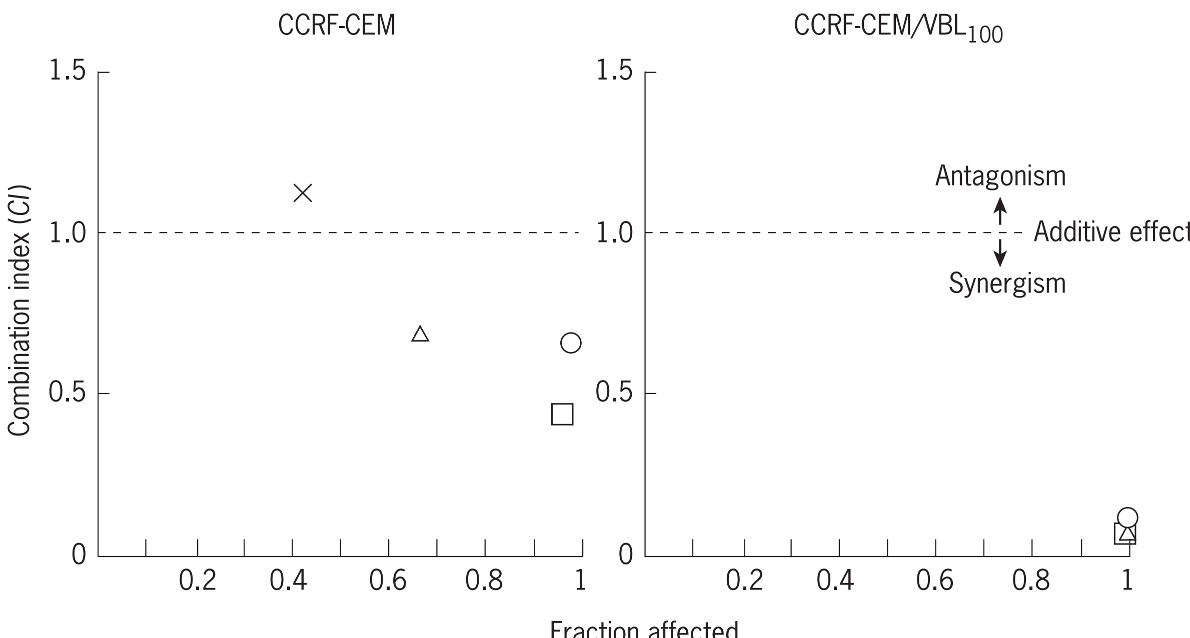


Figure 3 Synergism between vinblastine (VBL) and MDR reversing agent, 5-*N*-acetylardeemin (NAA), in combination against the growth of CCRF-CEM and its subline resistant to VBL, CCRF-CEM/VBL_{100}. $CI < 1$, $CI = 1$, and $CI > 1$ indicate synergism, additive effect and antagonism, respectively (Chou and Talalay, 1984; Chou 1991a).

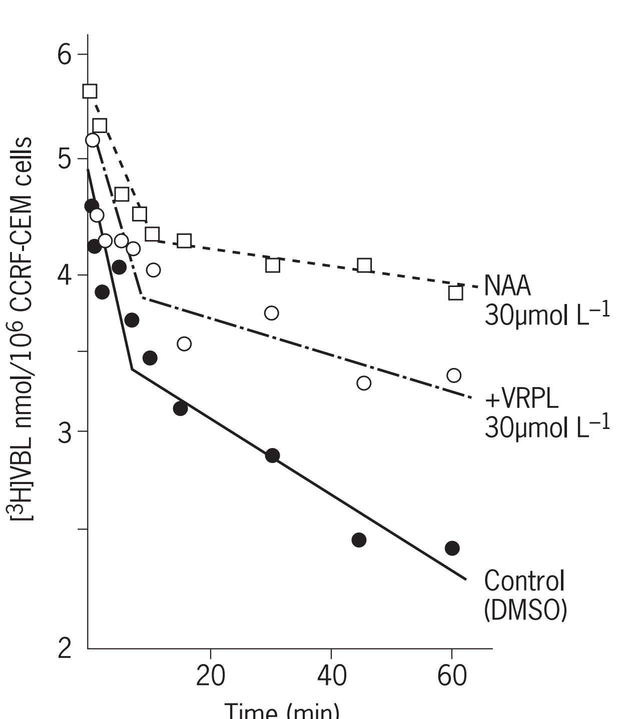


Figure 4 Reduced rates of efflux of preloaded [^{3}H]vinblastine (VBL) in CCRF-CEM cells by the MDR reversing agents 5-*N*-acetylardeemin (NAA) and verapamil (VRPL). [^{3}H]VBL was preloaded for 30 min at 37°C and, washed twice at 4°C with fresh medium. NAA or VRPL was added at 0 min and incubated at 37°C for various time periods. (From Chou *et al*., 1998, *Proceedings of the National Academy of Sciences of the USA*, **95**, 8369–8374.)

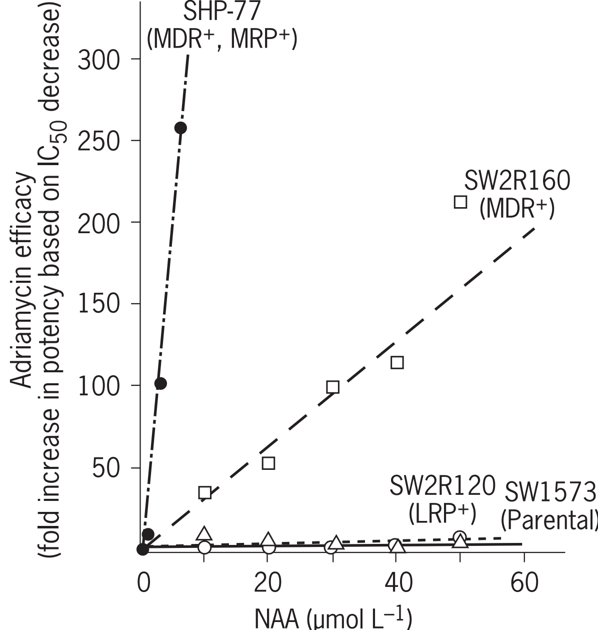


Figure 5 Enhancement of doxorubicin (DX) cytotoxicity by NAA. Six concentrations of DX were tested for cytotoxicity in (●) SHP77 ($MDR^+ + MRP^+$); (△) SW2R120 (LRP^+); (□) SW2R160 (MDR^+); and (○), SW1573 (wild-type) cells in the presence and absence of NAA. The concentrations required for 50% inhibition of cell growth (IC_{50} in μmol/L) were determined by the median-effect plot using computer software (Chou and Hayball, 1996). Fold of increase in efficacy of DX is based on fold of decrease in IC_{50}. NAA alone showed little or no cytotoxicity in these cell lines.

CCRF-CEM/VBL_{100} cells expressing Pgp were competitively displaced.

8. *Selective effects of the MDR-reversing agent*: whether the reversing agent sensitizes the cells of various subtypes of MDR such as Pgp, MRP and LRP. Enhancement of doxorubicin cytotoxicity by NAA (1–50 μmol/L) in human cancer cell lines expressing different MDR phenotypes was studied in culture (**Figure 5**). For SHP-77 (MDR-Pgp^+ + MRP^+) cells, the IC_{50} of doxorubicin was reduced 260-fold to 0.009 μmol/L by 6 μmol/L of NAA. In contrast, for SW1573 (parent), SW2R120 (LRP^+) and SW2R160 (MDR-Pgp^+) cells, the doxorubicin IC_{50} values were reduced three-, seven-, and 50-fold, respectively, by 20 μmol/L of NAA. Therefore, NAA markedly enhanced the efficacy of doxorubicin against SHP-77 (MDR-Pgp^+ + MRP^+) and SW2R160 (MDR-Pgp^+) cells, but had little effect in enhancing the efficacy of doxorubicin against SW2R120 (LRP^+) or SW1573 (parental cells).

MULTI-HIT AND MULTI-TARGET APPROACH AGAINST MDR: DRUG COMBINATIONS

As indicated above, there are many possible mechanisms that confer drug resistance. For multidrug resistance, there are mdr_1, mdr_2, MRP_1, MRP_2, MRP_6, LRP, etc. In addition, the ATP-binding cascade (ABC) superfamily has numerous genes (Higgins, 1995; Kuska, 1999) that may be relevant to drug resistance. It is reasonable to expect that a single agent alone is insufficient to overcome multifactorial multidrug resistance. Therefore, multidrug combination is a logical way to combat drug resistance.

This approach can be divided into the following categories: (1) combining drugs with different mechanisms of action; (2) combining MDR substrates with non-MDR substrates; and (3) combining MDR substrates belonging to different subclasses.

The outcomes of drug combination can be synergistic, antagonistic or additive. There is no simple way to predict the outcome quantitatively even at the *in vitro* level, and in animal models or clinics it will be much more complex.

The goal of drug combination is to (1) increase therapeutic efficacy by synergism, thus lowering the dose(s) of the drugs, (2) reduce toxicity or side effects against the host by reducing the dose; (3) minimize or reverse the drug resistance or delay the development of drug resistance; increased therapeutic efficacy leads to a decrease in the length of therapy, which in turn decreases the risk of developing resistance; and (4) achieve selective synergism against the target and/or selective antagonism toward the host, hence increasing the therapeutic index.

There are general rules that can be generated from past experience with drug combination studies. First, drugs with different mechanisms of actions or different modes of effect usually generate more synergism than those with similar modes of action. Second, based on the mass-action principle, synergism is usually more prominent at high dose or high effect levels than at low dose or low effect levels. Third, one should not attempt to predict or to interpret synergism or antagonism based on individual drug mechanisms alone since it tends to be oversimplified. Synergism or antagonism can be due to drug transport, metabolism, etc., unrelated to mechanism *per se*. For a simple case such as inhibition of tumour cell growth *in vitro*, it can be the combined results of cytocidal and cytostatic effects, and nobody knows from a live cell to a dead cell how many steps there are involved that can be the target step(s) for the component drugs.

Although the main focus for the pharmaceutical industries and academic drug discovery programmes is to search for individual new chemical entities or new therapeutic modes of action, in the treatment of neoplastic disease and infectious disease such as AIDS, monotherapy has been proved inadequate and combination therapy should be used (Chou, 1991a). The main problem, however, is that most clinical trials are still at the stage of trial-and-error in which there is no direct evidence that a particular combination may give a beneficial synergistic effect even *in vitro*. In addition, there is a great deal of confusion regarding data analysis and interpretation owing to a lack of a clear definition of 'synergism' (Chou, 1998).

Since synergism is an effect that is more than an additive effect and antagonism is less than an additive effect, a clear definition of 'additive effect' is of crucial importance in any meaningful quantitation of synergism or antagonism. So far, only two methods, namely the isobologram (Chou and Talalay, 1983) and the combination index method of Chou and Talalay (1984), involve a vigorously derived equation for defining what 'additive effect' is.

QUANTITATION OF SYNERGISM AND ANTAGONISM

The median-effect equation of Chou (Chou, 1976; Chou and Talalay, 1984) and the median-effect plot of are derived on the basis of the mass-action law principle:

$$(f_a/f_u) = (D/D_m)^m \quad [1]$$

where D is the dosage or concentration of a drug, f_a and f_u are the fraction affected and unaffected, respectively, by the drug, D_m is the the median-effect dose, signifying the 'potency' of the drug, and, the exponent m signifies the 'shape' or sigmoidicity of the dose–effect curve (i.e. $m = 1$, hyperbolic; $m > 1$, sigmoidal; $m < 1$, negative (flat) sigmoidal). The linear correlation coefficient of the median–effect plot, r signifies the 'conformity' of the data to the

method of analysis. Rearrangement of eqn. [1] gives

$$f_a = 1/[1 + (D_m/D)^m] \quad [2]$$

and

$$D = D_m[f_a/(1 - f_a)]^{1/m} \quad [3]$$

Thus, as the m and D_m values are known, the dose and effect can be interchanged. The logarithmic form of eqn [1] gives a linear equation:

$$\log(f_a/f_u) = m \log D - m \log D_m \quad [4]$$

Therefore, Chou's median-effect plot of log D against log (f_a/f_u) yields an x-intercept of log D_m and thus the D_m value and the slope of the plot (i.e. the m value) can be easily obtained (Chou and Talalay, 1984).

The median-effect principle has been extended to the analysis of the dose–effect relationship for multiple drugs. The multidrug effect equation was derived from the inhibition of the enzyme kinetic models via mathematical induction and reduction. The resulting general equation was then used to define the combination index (*CI*) equation of Chou and Talalay (1981, 1984):

$$CI = (D)_1/(Dx)_1 + (D)_2/(Dx)_2 \quad [5]$$

for a two-drug combination, where the denominators are the doses of each drug alone that produce x% effect and the numerators are the dose of the two drugs combined that also produce x% effect. Equation [5] allows the determination of the *CI* value, where $CI < 1$, $CI = 1$ and $CI > 1$ indicate synergism, additive effect and antagonism, respectively (Chou and Talalay, 1984). Computer software has been developed for automated analysis (Chou, 1991b; Chou and Hayball, 1996; Chou *et al*., 2000). This mass-action law-based method has gained popularity in the past two decades. Based on the *Citation Index* (Institute for Scientific Information, Philadelphia, Pa, 1982–1999), the median-effect principle of Chou and the combination-index equation of Chou-Talalay and its computer software have been used in scientific papers published in over 270 different biomedical journals.

EXPERIMENTAL DESIGN FOR DRUG COMBINATIONS

Because each drug may have different 'potencies' and may have different 'shapes' of dose–effect curves, it is essential to have a dose–effect curve (dose–effect relationship) for each drug before contemplating drug combination analysis. A typical old-fashioned experimental design is to vary the dose of each drug alone for the dose–effect curve parameters, and then use a 'non-constant ratio' design of combination (e.g. keeping the dose of one drug constant while varying the doses of the second drug). In this design, the *CI* value at each combination of data points can be calculated (Chou and Talalay, 1984; Chou, 1991a). However, the most efficient design is to vary the dose of each drug alone and then make a mixture of the drugs, and then serially dilute it. Thus, the combination is at a 'constant ratio' of combination. In effect, the mixture is treated as the third drug in the two-drug combination. In this design, not only can the *CI* values at each combination of data points be calculated, but we can also simulate the *CI* values at all effect levels. In addition, the classical isobologram can be automatically generated. By contrast, in a one-constant ratio design, only the normalized isobologram can be constructed (i.e. the dose (D) is normalized by Dx for x% inhibition) (Chou and Talalay, 1984; Chou, 1991a; Chou *et al*., 2000).

SELECTIVITY OF SYNERGISM AND ANTAGONISM

Most researchers have paid attention to the synergistic efficacy against the target (e.g. tumour cells, HIV), but neglected the possibility of synergism against the host which produces side effects or toxicity. In a series of classical studies by Chang *et al* (1985, 1987), 4-hydroperoxycyclophosphamide (4-HC) was combined with another anticancer agent, etoposide (VP-16), for purging leukaemic cells in autologous bone marrow transplantation. At high dose–high effect levels, the combination showed a synergistic effect against the proliferation of human promyelocytic leukaemic cells (HL-60), but showed antagonism towards the bone marrow progenitor cells as measured by CFU-GM and BFU-E (Chang *et al*., 1985). In another study, 3′-azidothymidine (AZT) and interferon-alpha$_{2A}$ (IFN) showed strong synergism against HIV proliferation, but showed only moderate synergism against bone marrow progenitor cells (Berman *et al*., 1989).

RATIONAL APPROACH TO THE CLINICAL PROTOCOL DESIGN FOR DRUG COMBINATIONS

With a quantitative method for synergism and antagonism readily available, we no longer take the risk of trial-and-error methods for conducting drug combination clinical studies. Although *in vitro* data under simple and defined conditions may not be fully applicable for *in vivo* or clinical situations, the quantitative method at least provides directions as to answering the following questions

(Chou, 1991a, 1998):

1. *Selection of candidate combination*: among many antineoplastic drugs with different mechanisms of action, which pair, triplet or quadruplet of drug combination has better synergism than others?
2. *Pattern of synergism/antagonism*: if synergism is observed, does synergism occur similarly at different effect levels or different dose levels?
3. *Combination ratio*: does an equimolar ratio of combination yield a greater synergistic effect than other combination ratios?
4. *Schedule dependency*: do two drugs that are administered simultaneously or in a sequential order or reverse order produce similar degrees of synergism?

All of the above questions are important and relevant to the rational approach to clinical drug combination protocol design. For a more detailed review and discussion see Chou *et al*. (1994, 1997) and Chou (1998).

REFERENCES

Antoney, A. and Kaye, S. B. (1999). Drug resistance: the clinical perspective. In: Brown, R. and Bouger-Brown, U. (eds), *Cytotoxic Drug Resistance Mechanisms*. 1–16. (Humana Press, Totowa, NJ).

Aran, J. M., *et al*. (1999). Therapeutic strategies involving the multidrug resistance phenotype: the MDR1 gene as target, chemoprotectant, and selectable marker in gene therapy. *Advances in Pharmacology*, **46**, 1–42.

Bates, S. E., *et al*. (1999). Measuring MDR-1 by quantitative RT-PCR. In: Brown, R. and Bouger-Brown, U. (eds), *Cytoxic Drug Resistance Mechanisms*. 63–82 (Humana Press, Totowa, NJ).

Biedler, J. L. and Riehm, H. (1970) Cellular resistance to actinomycin D in Chinese hamster cells *in vitro*: cross resistance, radioautographic and cytogenetic studies. *Cancer Research*, **30**, 1174–1184.

Berman, E., *et al*. (1989). Synergistic cytotoxic effect of azidothymidine and recombinant interferon alpha on normal human bone marrow progenitor cells. *Blood*, **74**, 1281–1286.

Casciari, J. J., *et al*. (1994). Growth and chemotherapeutic response of cells in a hollow-fiber *in-vitro* solid tumor model. *Journal of the National Cancer Institute*, **86**, 1846–1852.

Chabner, B. A. (1990). In defense of cell-line screening. *Journal of the National Cancer Institute*, **82**, 1083–1085.

Chang, T. T., *et al*. (1985). Synergistic effect of 4-hydroperoxycyclophosphamide and etoposide on HL-60 myelogenous leukemia cell line demonstrated by computer analysis. *Cancer Research*, **45**, 2434–3439.

Chang, T. T., *et al*. (1987) Comparative cytotoxicity of various drug combinations for human leukemic cells and normal hematopoietic precursors. *Cancer Research*, **47**, 119–122.

Chou, T.-C. (1976). Derivation and properties of Michaelis–Menten type and Hill type equations for reference ligands. *Journal of Theoretical Biology*, **39**, 253–276.

Chou, T.-C. and Talalay, P. (1981). Generalized equations for the analysis of inhibitors of Michaelis–Menten and higher order kinetic systems with two or more mutually exclusive and nonexclusive inhibitors. *European Journal of Biochemistry*, **115**, 207–216.

Chou, T.-C. and Talalay, P. (1983). Analysis of combined drug effects: a new look at a very old problem. *Trends in Pharmacological Science*, **4**, 450–454.

Chou, T.-C. and Talalay, P. (1984). Quantitative analysis of dose–effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Advances in Enzyme Regulation*, **22**, 27–55.

Chou, T.-C. (1991a). The median-effect principle and the combination index for quantitation of synergism and antagonism. In: Chou, T.-C. and Rideout, D. C. (eds), *Synergism and Antagonism in Chemotherapy*. 61–102 (Academic Press, San Diego).

Chou, J. (1991b). Quantitation of synergism and antagonism of two or more drugs by computerized analysis. In: Chou, T.-C. and Rideout, R. C. (eds), *Synergism and Antagonism in Chemotherapy*. 223–244 (Academic Press, New York).

Chou, T.-C., *et al*. (1994). Computerized quantitation of synergism and antagonism of taxol, topotecan and cisplatin against teratocarcinoma cell growth: a rational approach to clinical protocol design. *Journal of the National Cancer Institute*, **86**, 1517–1524.

Chou, T.-C. and Hayball, M. (1996). *CalcuSyn: quantitation of synergism and antagonism software for Windows*. (Biosoft, Cambridge).

Chou, T.-C., *et al*. (1997). Chemotherapeutic synergism, potentiation and antagonism. In: Dulbecco, R. (ed), *Encyclopedia of Human Biology*, 2nd edn. Vol. 2, 675–683 (Academic Press, San Diego).

Chou, T.-C., *et al*. (l998). Reversal of anticancer multi-drug resistance by the ardeemins. *Proceedings of the National Academy of Sciences of the USA*, **95**, 8369–8374.

Chou, T.-C. (1998). Drug combinations: from laboratory to practice. *Journal of Laboratory and Clinical Medicine*, **132**, 6–8.

Clynes, M., *et al*. (1998). Recent developments in drug resistance and apoptosis research. *Critical Reviews in Oncology and Hematology*, **28**, 181–205.

Darby, G. (1997). In: Cox, T. M. and Sinclair, J. (eds), *Molecular Biology in Medicine*. 271–283 (Blackwell Science, Oxford).

Driscoll, J. S. (1984). The preclinical new drug research program of the National Cancer Institute. *Cancer Treatment Reports*, **68**, 63–76.

Food and Drug Administration (1995). *An FDA Consumer Special Report From Test Tube to Patient: New Drug Development in the United States.* DHHS Publication No. (FDA) 95–3168. (Superintendent of Documents, Government Printing Office, Washington, DC).

Gottesman, M. M. and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual Reviews in Biochemistry*, **62**, 385–427.

Higgins, C. F. (1995). The ABC of channel regulation. *Cell*, **82**, 693–696.

Johnson, R. K. (1990). Screening methods in antineoplastic drug discovery. *Journal of the National Cancer Institute*, **82**, 1082–1083

Juliano, R. L. and Ling, V. (1996) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochimica Biophysica Acta*, **455**, 152–162.

Kartner, N. and Ling, V. (1989). Multidrug resistance in cancer. *Scientific American*, March 44–51.

Kellen, J. A. (1994), Multidrug resistance. In: Kellen, J. A. (ed.), *Reversal of Multidrug Resistance in Cancer*. Ch. 1, 1–19 (CRC Press, Boca Raton, FL).

Kuska, B. (l999). As easy as ABC: scientists fish out another drug resistance gene. *Journal of the National Cancer Institute*, **91**, 402–404.

Laing, N. M. and Tew, K.-D. (1997). Drug resistance to chemotherapy: mechanisms. In: Bertino, R. J. (ed.), *Encyclopedia of Cancer*, Vol. 1. 560–570 (Academic Press, New York).

Lam, K. S., *et al*. (1991). A new type of synthetic peptide library for identifying ligand-binding activity. *Nature*, **354**, 82–84

Paull, K. D., *et al*. (1989). Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *Journal of the National Cancer Institute*, **81**, 1088–1092.

Rappa, G., *et al*. (1999). New insight into the biology and pharmacology of the multidrug resistance protein (MRP) from gene knockout models. *Biochemical Pharmacology*, **58**, 557–562.

Robert, J. (1999). Multidrug resistance in oncology: diagnostic and therapeutic approaches. *European Journal of Clinical Investigation*, **29**, 536–545.

Ross, D. D., *et al*. (1999). Atypical multidrug resistance: breast cancer resistance protein messenger RNA expression in mitoxantrone-selected cell lines. *Journal of the National Cancer Institute*, **91**, 429–433.

Scangos, G. (1997). Drug discovery in the postgenomic era. *Nature Biotechnology*, **15**, 1220–2211.

Schinkel, A. M., *et al*. (1997). Normal viability and altered pharmacokinetics in mice lacking mdrl-type (drug-transporting) P-glycoproteins. *Proceedings of the Natural Academy of Sciences of the USA*, **94**, 4028–4033.

Scudiero, D. A., *et al*. (1988). Evaluation of a soluble tetrazolium/formazan assay for cell growth and drug sensitivity in culture using human and other tumor cell lines. *Cancer Research*, **48**, 4827–4833.

Sikic, B. I. (1991). Anticancer drug discovery. *Journal of the National Cancer Institute*, **83**, 738–740.

Skehan, P., *et al*. (1990) New colorimetric cytotoxicity assay for anticancer drug screening. *Journal of the National Cancer Institute*, **82**, 1107–1112.

Tsujimoto, Y. and Croce, C. M. (1986). Analysis of the structure, transcripts and protein products of bcl-2: the gene involved in human follicular lymphoma. *Proceedings of the National Academy of Sciences of the USA*, **83**, 5214–5218.

Ueda, K., *et al*. (1999). Recent progress in P-glycoprotein research. *Anti-Cancer Drug Design*, **14**, 115–121.

Volm, M. (1998). Multidrug resistance and its reversal. *Anticancer Research*, **28**, 2905–2918.

FURTHER READING

Ambudkar, S. V., *et al*. (1999). Biochemical, cellular and pharmacological aspects of the multidrug transporter. *Annual Reviews in Pharmacology and Toxicology*, **39**, 361–398.

Aszalos, A. and Ross, D. D. (1999). Biochemical and clinical aspects of efflux pump related resistance to anti-cancer drugs. *Anticancer Research*, **18**, 2937–2944.

Bradshaw, D. M. and Arceci, R. J. (1998). Clinical relevance of transmembrane drug efflux as a mechanism of multidrug resistance. *Journal of Clinical Oncology*, **16**, 3674–3690.

Harvey, A. L. (1999). Medicine from nature: are natural products still relevant to drug discovery? *Trends in Pharmaceutical Science,* **20**, 196–198.

Kessel, D. (ed.) (1989). *Resistance to Antineoplastic Drugs*. (CRC Press, Boca Raton, FL).

Safa, A. R. (1992). Photoaffinity labelling of P-glycoprotein in multidrug resistant cells. *Cancer Investigations*, **10**, 295–305.

Safa, A. R. (1996). Multidrug resistance. In: Schilsky, R. L., *et al*. (eds), *Priniciple of Antineoplastic Drug Development and Pharmacology*. 457–486 (Marcel Dekker, New York).

Vossebeld, P. J. M. and Sunneveld, P. (1999). Reversal of multidrug resistance in hematological malignancies. *Blood Reviews*, **13**, 67–68.

Chapter 68

Models for Immunotherapy and Cancer Vaccines

Volker Schirrmacher, Victor Umansky, Matthias Lindner, Susanne Müerköster and Marian Rocha
German Cancer Research Center, Heidelberg, Germany

CONTENTS

- Biological Background
- Principles of Model Establishment
- General Applications
- Biological Limitations
- Technical Considerations
- Perspectives

BIOLOGICAL BACKGROUND

Why Immunotherapy?

The greatest limitation to the successful treatment of cancer is the ability of tumours to form distant metastases. These are often located in vital organs and are either inaccessable or refractory to conventional forms of cancer treatment such as surgery, chemotherapy or radiotherapy. A further limitation of conventional cancer treatment is the necessity to balance the toxicity and side effects with the dose-dependent beneficial antitumour effects of these modalities. The ideal form of cancer therapy would be selective for the tumour and thereby nontoxic to the host, able to circumvent heterogeneity within a given tumour and be able to reach tumour metastases in all anatomical sites. High expectations have arisen that the immune system can be exploited for immunotherapeutic anticancer treatment because of its exquisite specificity. This is associated primarily with T lymphocytes and antibodies, but also with professional antigen-presenting cells (APCs), such as dendritic cells (DCs), and with cytotoxic effector cells such as NK cells, macrophages and others. (See also chapters on *Antibodies and Recombinant Cytokines* and *Genetic and Cellular Vaccines*.)

Biology of Metastasis Formation

For the design of immunological interventions to be successful in the treatment of metastases it seems essential to be aware of basic concepts and aspects of metastasis formation which have been known for many years (Fidler and Radinsky, 1990; Kerbel, 1990; Nicolson, 1993). These are, for example, clonal selection of metastatic cells, heterogeneity of metastatic subpopulations, organ specificity of metastases and the role of angiogenesis. Proteolysis and alterations in adhesive functions are the most obvious and thus one of the most thoroughly investigated processes. Various proteases and precursors and their inhibitors exhibit a complex interplay of a highly regulated nature. The same holds true for adhesive and 'de-adhesive' tumour–host interactions, involving adhesion molecules from the immunoglobulin superfamily, integrins, cadherins, selectins and hyaluronic acid receptors as well as their ligands. The control of the metastatic cascade by these extracellularly acting molecules is delicately balanced. Slight changes could affect any balance and consequently promote metastasis formation.

Molecular Targets of Metastases

Upon ligand binding, cell surface receptors transmit signals into the cell, activating multiple metabolic and signal transduction pathways which can promote cell growth, cell differentiation or programmed cell death (apoptosis) (Hart *et al*., 1989; Fodstad, 1993; Khazaie *et al*., 1993). Metastatic cancer cells may become more resistant to apoptosis than their non-metastatic counterparts (Glinsky, 1998). Although the nature of the stimulatory molecules is highly variable, the principles involved in propagating the signals seem to be ubiquitous. Thus, such different stimulators as antigens, cytokines, hormones and growth factors (Khazaie *et al*., 1993) use similar transduction principles including ligand induced receptor internalization of receptor–ligand complexes, activation of G proteins, oligomerization, increases in calcium uptake, activation of tyrosine kinases or other kinases and phosphatases, protein

phosphorylation or dephosphorylation and the triggering of oncogenes and transcription factors. The signalling molecules interact with each other through universal 'zippers' such as SH2 and SH3 domains. Many of the currently described metastases-related molecules are confined to distinct cellular compartments such as the extracellular space, the cell membrane, the cytoplasmic signalling network and the nuclear regulatory system. Promising new therapeutic approaches try to use recombinant soluble metastases-related molecules which interfere with ligand interactions.

Tumour-associated Antigens (TAAs) and the Immune System

In recent years we have also seen an increase in biological therapies which incorporate new insights from immunology and molecular biology (Eibl *et al*., 1997, 1998). The process of selection and expansion of mutant clones that occurs during progression of human malignant disease leads to a wide range of random mutational, recombinatorial, deletional and regulatory gene changes. Neoantigens may be produced (1) by up-regulation of lineage-specific markers (e.g. differentiation antigens) and/or oncofetal antigens to immunologically 'visible' levels and (2) by the occurrence of nonlethal mutations in widely expressed proteins. Thus, tumours may express both unique TAAs and cross-protective antigens (Coggin *et al*., 1998). Together, such new TAAs make a strong case for the use of autologous tumour cell vaccines (rather than defined antigenic peptides) to ensure the recognition of unknown but potentially targetable TAAs generated by random mutational processes (Ellem *et al*., 1998).

The immune system is a complex system with several functional components and powerful information-processing capabilities. It employs a multilevel defence against invaders through nonspecific (innate) and specific (acquired) immune mechanisms. The immune system has learned through evolution to distinguish between foreign antigens (e.g. bacteria, viruses, etc.) and the body's own cells or molecules. The lymphocyte is the main type of immune cell participating in the immune response and responsible for specificity, diversity, memory and adaptivity. Other cells called phagocytic cells – neutrophils, eosinophils, basophils, monocytes – are accessory immune cells whose primary function is to provide facilities to eliminate antigens. Different populations of lymphocyte recirculate at primary and secondary lymphoid organs and are carefully controlled to ensure that appropriate populations of B and T cells (naive, effector and memory) are recruited into different locations. This differential migration of lymphocyte subpopulations at different locations (organs) of the body is called trafficking or homing.

The key features of the immune system which are of great interest to the field of information processing (Dasgupta, 1999) are as follows:

- *recognition* of patterns (antigen) and generation of selective responses;
- *feature extraction*: each APC, by processing and presenting antigenic peptides, serves as a filter and a lens: a filter that destroys molecular noise, and a lens that focuses the attention to the lymphocyte receptors;
- *diversity generation* in lymphocyte receptors by genetic combinatory processes involving basic elements;
- *learning* by experience and associations mediated by changes in specific lymphocyte concentrations;
- *memory*: using short- and long-term memory mechanisms, the immune system keeps an ideal balance between economy and performance in conserving a minimal but sufficient memory of the past (Ahmed and Gray, 1996);
- *self-regulation*: immune responses are self-regulatory although there is no central organ for this; the regulation can be either local or systemic, depending on the route and property of the antigenic challenge;
- *threshold mechanisms*: immune responses take place above a certain matching threshold (strength of chemical binding);
- *co-stimulation*: activation of T and B cells is closely regulated through second signal co-stimulation; the second signal (often from helper T cells) helps to ensure tolerance and judge between dangerous and harmless invaders or false alarm.

In order for immunological approaches to the treatment of cancer to live up to their expected potential, it seems essential to investigate systematically in appropriate model systems the parameters required for optimum stimulation of effector activities and for selective accumulation of effector cells and/or molecules at the sites of metastases. Because of the complex nature of metastasis formation, of tumour–host interactions, of antitumour immune responses and of tumour immune escape mechanisms (Schirrmacher *et al*., 1982; Ellem *et al*., 1998) it is very important that such experimental model systems fulfil certain criteria, such as: (1) reproducible formation of metastases at distinct organ sites, (2) protocols for reproducible induction of antitumour immune responses and (3) reproducible sensitivity of the tumour system to immune manipulation and immunotherapy.

PRINCIPLES OF MODEL ESTABLISHMENT

A prerequisite for basic research in cancer metastases is the availability of a tumour line with the capacity to metastasize. In order to find out which properties enable a metastasizing cancer cell to metastasize, it is advantageous to compare it with a related cancer line which has either no or a very reduced metastatic capacity. Many of the molecules that were found to play a role in metastasis were

identified in such a way. For studies on immune recognition by T lymphocytes such tumour lines should be typed for cell surface expression of major histocompatibility complex (MHC) molecules of class I or class II. These molecules are used to present intracellular peptides to immune CD8 or immune CD4 T lymphocytes, respectively. Further molecules of importance are TAAs, cell adhesion molecules (CAM) and differentiation antigens (DA). These molecules may play important roles in cell–cell interactions, cell–matrix interactions and cellular signalling. Degradative enzymes are important for tumour cell invasiveness and are regulated by respective inhibitors.

Establishment of Related Tumour Sublines with Different Metastatic Capacity

The model system we have been studying most intensively is the Eb/ESb/ESb-MP mouse tumour (Schirrmacher, 1982). ESb cells represent a spontaneous highly metastatic variant of the chemically induced T cell lymphoma L5178 Y (Eb) of DBA/2 mice and arose most likely following fusion of Eb cells with a host macrophage. The ESb-MP subline is an adhesion variant of ESb which grows *in vitro* attached to plastic whereas Esb cells grow in suspension. *In vivo*, ESb-MP cells grow progressively but show a less aggressive phenotype, metastasizing more slowly than ESb and involving multiple organs. The following sublines of ESb were also established: ESbL, a line reisolated from a single liver metastasis after spontaneous metastasis of ESb from an intradermal site; ESb-TA$^-$ cells, an immune escape variant resistant to syngeneic ESb-TA$^+$ specific cytotoxic T lymphocytes. The phenotype and characteristics of the tumour lines of this model system are summarized in **Table 1**.

Gene Tagging

We decided to use the bacterial *lacZ* gene to tag genetically our tumour cells and at the same time to use this foreign gene as a model tumour-associated antigen (TAA). ESbL-lacZ cells represent a gene transfectant of ESbL retrovirally transduced with *lacZ* vector BAG, selected for high expression via fluorescence-activated cell sorter (FACScan) and cloned (Krüger *et al*., 1994a, b). Since its description almost 40 years ago, the *lacZ* gene has become particularly useful in phage and bacterial genetics and in molecular DNA cloning. In the early 1980s the *lacZ* gene together with the chromogenic substrate 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-Gal) was also introduced into cell biology, developmental biology and finally into metastasis research. Lineage studies or detection of individual tumour cells in the context of normal tissue could then be attempted since the tag was transmitted to subsequent generations upon cell division. X-Gal

Table 1 Phenotype and characteristics of tumour lines used

	Eb	**ESb**	**ESb-MP**
Aetiology	Chemically induced T lymphoma (L5178Y)	Spontaneous metastasis variant	Adhesion variant
Phenotype[a]			
MHC class I, II	I^+, II^-	I^+, II^-	I^+, II^-
TA	Eb-type (K^d)	ESb-type (K^d)	ESb-type (K^d)
CAM	LFA-1	ICAM-1	LFA-1, ICAM-1, L1
DA	Thy1, CD8, CD45R	Lyt1, CD2, CD45R	Mac-1, CD2, CD45R
Invasiveness[b]	Low	High	High
Degradative enzymes	MTSP-1	MTSP-1, u-PA, heparanase	MTSP-1, u-PA, heparanase
Malignancy	Low	High	Low
Primary tumour (size)	1.5 cm	0.8 cm	2.0 cm
50% survival time (t_{50})	28 days	10 days	42 days
Dissemination	—	Fast	Slow
Involved organs	—	Liver, bone marrow, lung, spleen	Liver, bone marrow, lung, spleen, kidney, spinal column
Sublines			
Liver selected[c]		ESb-L	
LacZ transfected[d]	Eb-lacZ	ESb-L-lacZ	
Immune escape[e]		ESb-TA$^-$	

[a]MHC = major histocompatibility molecules; TA = tumour antigen; CAM = cell adhesion molecules; DA = differentiation antigens.
[b]Through endothelium and subendothelial extracellular matrix, matrigel, fibrin gel, collagen gel and lung cubes; MTSP-1 = murine T cell related serine proteinase; u-PA = plasminogen activator of urokinase type.
[c]After spontaneous metastasis from intradermal site, reisolated from a single liver metastasis.
[d]Retrovirally transduced with *lacZ*, *neo*^{R} vector, selected via G418 and FACS and cloned.
[e]Immune escape variant ESb816, resistant to ESb-TA$^+$ specific cytotoxic T lymphocyte (CTL).

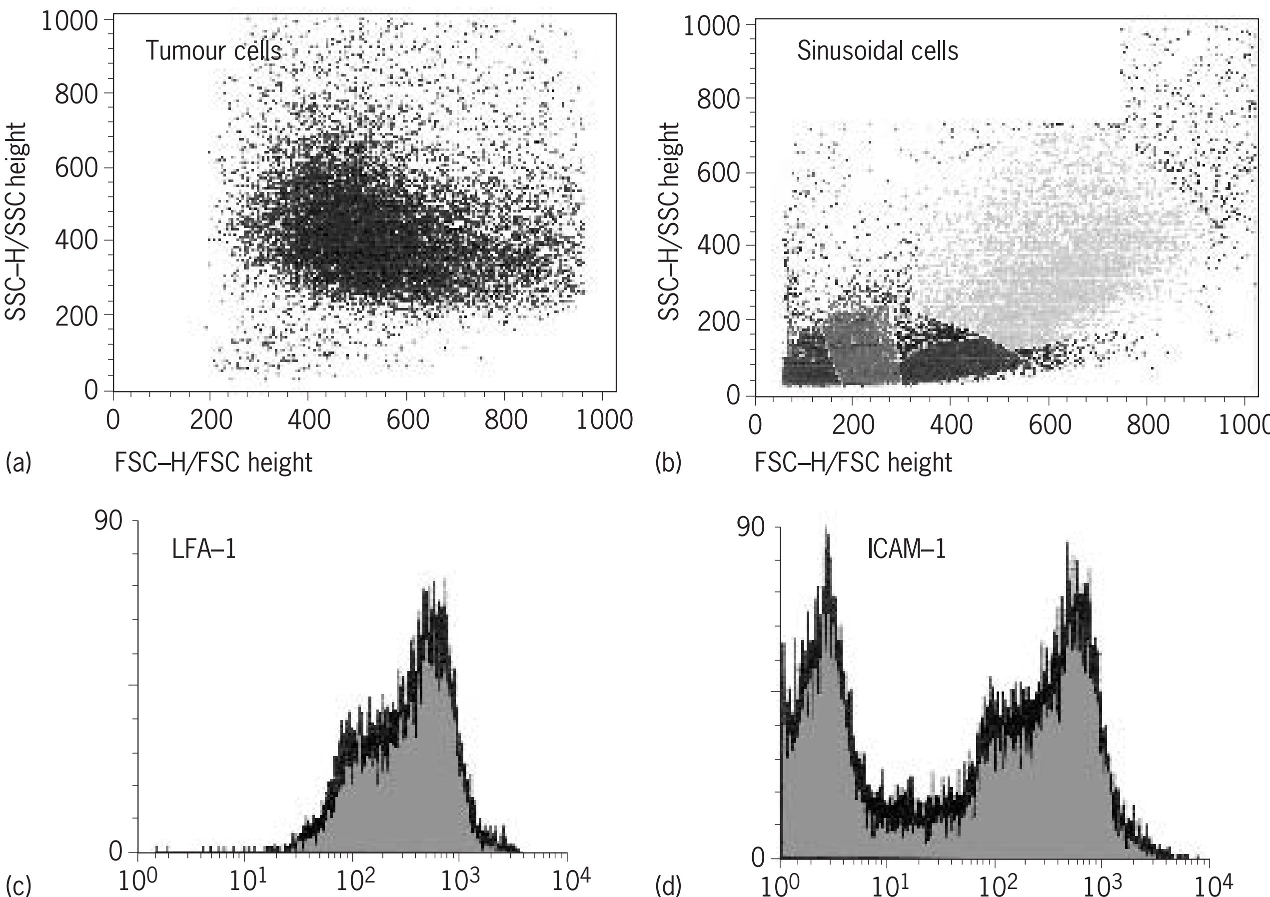


Figure 1 Flow cytometric analysis of FDG-loaded isolated liver sinusoidal cells and tumour cells. Tumour cells expressing the *lacZ* gene were identified as green fluorescence-positive cells (a) and sinusoidal cells as green fluorescence-negative cells (b). (c, d) monoparametric frequency distribution of red fluorescence (FL2) in tumour and sinusoidal cells. An example is given with LFA-1 expression on tumour cells (c) and ICAM-1 in sinusoidal cells (d).

staining is usually performed on fixed cells or tissues with the consequence that stained cells are not viable. This limitation was overcome by the development of another substrate for β-galactosidase, namely fluorescein-di-β-D-galactopyranoside (FDG), making the detection of viable lacZ expressing cells possible by flow cytometry. This added a new dimension to cancer research as it allowed not only the histological description of early metastases by X-Gal staining but also the reisolation of FDG-stained tumour cells and purification from host cell populations making *ex vivo* characterization and molecular analysis of metastatic cells possible. With these new technologies, quantitative and qualitative studies permitted new insights into individual steps of the metastatic cascade, in modulated host environments and in evaluating the effectiveness of cancer therapies (Krüger *et al*., 1994a, b, 1999).

Visualization of Micrometastases

Micrometastases are secondary tumour cell colonies below the detection threshold of conventional technology in the clinic. They are important targets for adjuvant therapies. In primary operated breast cancer, for instance, there is an uncertainty whether micrometastastes are left behind which are a potential risk factor for tumour recurrence. Because of this uncertainty adjuvant chemotherapy is given to most breast-cancer patients. If one were able to visualize micrometastases, not every patient would need to receive such additional therapy which has considerable side effects.

Detection of micrometastases and small numbers of dormant tumour cells has only recently been possible in this model when using the genetically tagged tumour lines. X-Gal staining allowed single ESb-lacZ tumour cells to be detected in infiltrated organs such as lymph nodes, bone marrow or liver (Krüger *et al*., 1994a). Despite expression of the *lacZ* gene, the tumour cells were still tumorigenic, metastatic, unchanged in phenotype and therefore comparable to the parental ESb cells. After spontaneous metastasis, whole-organ staining revealed metastatic foci at the surface of the liver, while X-Gal staining of frozen tissue sections revealed micrometastasis in the form of

clusters and diffusely disseminated single cells (Krüger *et al*., 1994b). It was also possible to reisolate cells from metastases and to quantify them in metastasized organs via loading with the substrate FDG, which allowed tumour cell fluorescence staining and quantitative FACScan analysis. In a typical experiment, 12 days after intradermal tumour cell inoculation, 55% of the reisolated cells from the liver and 13% of the cells from the spleen were tumour cells.

The visualization and quantification of micrometastases gives important information about the type and degree of cellular spreading. For instance, it may be focal or diffuse, a distinction that may influence the patient's prognosis. **Figure 1** shows an example of flow cytometric analysis of tumour cells which had metastasized to the liver. Tumor cells (a) could be distinguished from normal sinusoidal cells (b) by green fluorescence following hypotonic FDG loading. Further staining with monoclonal antibodies allowed the detection of LFA-1 adhesion molecules on the tumour cells (c) and of ICAM-1 cell adhesion molecules on sinusoidal cells (d).

Visualization of Macrometastases

With nuclear magnetic resonance (NMR) imaging, it was possible in mice to detect the formation of focal metastasis in the liver and kidneys noninvasively and without contrast agents. The NMR studies were performed at high magnetic field strength (7.0 T) using a Bruker AM 300 spectrometer and a superconducting magnet with a 15-cm diameter vertical bore. During the measurements, mice were anaesthesized with chloral hydrate for 1–2 h in the case of ^{31}P-NMR spectroscopy of primary tumours or with isofluorane inhalation narcosis for 2–3 h in the case of ^{1}H-NMR imaging and kept at a temperature of about 32 °C. Metastases with a diameter $>$0.3 mm could usually be detected by day 21–23 after ESb-MP inoculation. This noninvasive imaging procedure in live tumour-bearing mice enabled us to follow the effect of immunotherapy (Fichtner *et al*., 1997).

GENERAL APPLICATIONS

Model Vaccination Studies Against Tumour Challenge

The field of human cancer vaccines is currently undergoing a revival. In a special issue of *Seminars in Oncology* (Cancer Vaccines, *Seminars Oncology*, **25**, No. 6, December 1998), some of the underlying principles are summarized as follows: (1) human cancers are weakly immunogenic or not immunogenic at all; (2) cell-mediated immune responses, mainly those of T-lymphocytes, are critical to tumour rejection; (3) almost nothing is known about the nature of the rejection antigens on human tumours; (4) it is not clear whether these antigens are different for each case of human cancer or whether there are clinically useful antigens that are shared by all or most tumours of a given histological type; and (5) the effectiveness of vaccines is greatly limited by excessive tumour burden. Examples of the most advanced clinical studies with human cancer vaccines are presented in this special issue. Most cancer vaccines use live tumour cells or lysates thereof. Some others use tumour-derived proteins, peptides or gangliosides. The use of DNA or RNA vaccines in general is relatively new. Its application to cancer was tested in our own laboratory.

Vaccination with DNA or RNA Encoding a Model Tumour Antigen and a Cytokine

Three different vaccination sites were compared for efficiency of immunization with naked DNA or self-replicating RNA (Förg *et al*., 1998; Schirrmacher *et al*., 2000a). Using the bacterial *lacZ* gene as a model, three injection sites of the mouse (skeletal muscle, dermis of the abdominal skin and ear pinna) were used for inoculation. Gene expression varied in expression time with muscle tissue showing the longest expression. Expression time, however, did not correlate with immune response intensity. The ear pinna was by far the most effective and muscle the least effective priming site for specific humoral and cytotoxic T cell-mediated immune responses (Förg *et al*., 1998). *Lac-Z* DNA immunization was also compared with cell based vaccination with *lacZ* transfected tumour cells, in which case again the pinna was the best site for inducing strong immune responses. Tumour-specific T cell responses could also be well induced in the pinna, leading to cytotoxic T lymphocyte induction and protective antitumour immunity (Schirrmacher *et al*., 2000a). Gene expression following pinna inoculation of either nonreplicating DNA plasmids or self-replicating RNA (Schirrmacher *et al*., 2000a) was similar, lasting for 2–3 weeks. Higher antibody responses were obtained with RNA than with DNA. β-Gal peptide-specific CTL memory responses to *lacZ* DNA or RNA were similar. They lasted for more than 6 weeks whereas the respective responses induced by *lacZ* tumour cells lasted for only 2 weeks.

The extent of protective antitumour immunity depended not only on the gene dose used for vaccination but also on the aggressiveness of the *lacZ* transfected tumour line used for challenge (Schirrmacher *et al*., 2000a). In comparison with *lacZ*-transfected tumour cells as vaccines, polynucleotide vaccination also demonstrated superiority with regard to cross-protection. Protective antitumour immunity could be strongly increased upon co-inoculation of *lacZ* DNA with IL-2 DNA or IL-12 RNA. These new model studies demonstrate efficient protective antitumour immunity after intra-pinna *lacZ* TAA polynucleotide vaccination and show additional immunomodulatory

effects by co-administration of cytokine polynucleotides (Schirrmacher *et al*., 2000a).

Vaccination with Live Tumour Cells; Establishment of Tumour Dormancy and Immune Memory

With the low metastatic Eb lymphoma, also genetically tagged with *lacZ*, we were able to visualize the rapid migration of tumour cells from a primary inoculation site such as the skin or the ear pinna to the lymphatic circulatory system and to the bone marrow (Khazaie *et al*., 1994). Live proliferation-competent and irradiated proliferation-incompetent Eb-*lacZ* cells were compared for their potency to induce after intra-pinna inoculation systemic antitumour immunity in syngeneic DBA/2 mice. Inoculation of nontumorigenic doses of live tumour cells led to long-lasting specific and systemic T cell-mediated antitumour responses requiring both CD4 and CD8 T lymphocytes. Irradiated cells as vaccines offered only limited short-term protection which could be marginally improved by local secretion of interleukin-4 following respective gene transfer (Khazaie *et al*., 1994). The more effective protection offered by vaccination with live tumour cells correlated with rapid migration and persistence of small numbers of tumour cells in the bone marrow of host animals. In contrast, irradiated Eb-*lacZ* cells had a short persistence. These observations indicated that in the course of vaccination with live tumour cells, a fraction of these cells escaped destruction by host mechanisms and persisted in a dormant state in the bone marrow for long periods of time. Persistence of dormant cells in the bone marrow correlated with the duration of antitumour immunity (Khazaie *et al*., 1994). We recently provided evidence for an important role of CD8 immune T cells in the control of tumour dormancy in the bone-marrow compartment. The bone-marrow residing tumour cells expressed the proliferation associated Ki-67 antigen and expanded upon CD8 T cell depletion. Upon transplantation of bone marrow from tumour-dormant to naive recipients, the co-existence of low numbers of tumour cells and CD8 immune memory T cells provided transient protection against further tumour cell challenge (Müller *et al*., 1998). The bone marrow thus appears to be a privileged site where protective immune T cells keep tumour cells in a dormant state. The use of *lacZ* tagging thus allowed us to study in detail tumour dormancy and the role of tumour–host interactions in this phenomenon.

Intra-pinna inoculation of live aggressive ESbL tumour cells also led to the induction of specific long-term systemic protective immunity rather than tumour growth. The ear pinna appeared to be the only site in syngeneic mice where these tumour cells which express a distinct TAA recognizable by specific CTL cannot grow. Induction of anti-ESbL immunity could be shown *in vivo* by rejection of a second tumour cell challenge or *in vitro* by killing of ESbL tumour cells in a ^{51}Cr-release assay using *in vitro* re-stimulated spleen cells from syngeneic immunized mice as effector cells. Both CD4 and CD8 T cells were required for induction of CTL (Schirrmacher *et al*., 1992) and protective immunity. Since we had a tumour model at hand in which we could compare experimental situations of tumour resistance (e.g. after pinna inoculation) with tumour susceptibility (e.g. after subcutaneous inoculation), we decided to examine the activation of cytokine genes directly *ex vivo* and to compare cytokine profiles with tumour resistance or susceptibility. The results revealed a superiority of the ear pinna over a subcutaneous inoculation site for induction of a Th1-type cytokine response (Jurianz *et al*., 1998).

Vaccination with Fusion Products of Tumour Cells and Dendritic Cells

If tumour cells express a tumour-associated antigen but no costimulatory molecules for activation of T lymphocytes, the interaction of such tumour cells with respective specific T cells could induce anergy rather than T cell activation. That means that such T cells will be shut off from responding to that same antigen upon further contact. This problem is particularly relevant for tumours arising from cells which are not professional antigen-presenting cells and are therefore devoid of costimulatory molecules. Carcinomas are such an example. They are derived from epithelia and represent the majority of human cancers. Dendritic cells are the most competent professional antigen-presenting cells and are the only cells able to prime nonantigen-experienced specific T lymphocytes. A concept has therefore been developed to fuse carcinoma cells with dendritic cells in order to introduce into the cancer cells all the properties which are necessary for costimulation of T cells.

Using *lacZ*-tagged cells, we also performed immunization experiments with carcinoma cells in the mouse. For this purpose we used the DA3 line of a chemically induced mammary carcinoma of BALB/c mice which we transfected with *lacZ* to generate the clone III-25 which was used in all further experiments. Whereas the immunogenicity of the parental carcinoma line was very low, possibly linked with the low expression of MHC I molecules, the *lacZ* transfectant showed a slight increase in immunogenicity and a reduced tumour growth *in vivo* due to the immune recognition of β-galactosidase as model TAA. Also in this model the ear pinna appeared to be the best site for immune response induction. Preimmunizing mice with fusion products of tumour cells and DCs led to the induction of systemic protective immunity against the parental DA3 cell line which was revealed by delayed tumour growth after challenge. In case of using fusion products with the *lacZ* transfectant, a very clear-cut protection was achieved in the majority of tumour-challenged animals. The immunization led to β-galactosidase-specific

memory CTLs which still existed 10 weeks after immunization. Immune spleen cells from such preimmunized mice were able to transfer protective antitumour immunity into nude mice preinoculated with respective tumour cells. These new and not yet published results reveal the special immunotherapeutic efficiency of fusion products of tumour cells and DCs in situations of low tumour immunogenicity (Lindner, 1998).

Model Immunotherapy Studies Against Established Metastases

Active Specific Immunotherapy (ASI) with Cancer Vaccines for Treatment of Micrometastases

Major contributions to the understanding of the principles of therapeutic tumour vaccination and ASI were made in the L10 guinea pig hepatocarcinoma (Key, 1981) and in the ESb-mouse lymphoma model. A number of similarities existed between the optimum protocols developed in these two metastasizing tumours and the therapeutic results obtained. Whereas in the L10 tumour postoperative immunotherapy with irradiated autologous tumour cells admixed with BCG protected against lymph node and lung metastases, in the ESb tumour model postoperative treatment with irradiated Newcastle disease virus (NDV)-infected autologous tumour cells protected against haematogenous metastases (Heicappell *et al*., 1986). Important variables for optimum therapeutic effects were the time of operation of the primary tumour, their residual disseminated tumour burden and the dose of virus or BCG added to a standard dose of 10^7 irradiated tumour cells. The site of vaccination also appeared to be important: In the L10 guinea pig model, three vaccinations injected intradermally protected against lymphatic spread, whereas in the ESb mouse model, intraperitoneal administration of the vaccine had the strongest protective effect against visceral metastastes. In both models, postoperative vaccination with inactivated viable tumour cells without viral or bacterial adjuvants had no therapeutic effect (Key and Hanna, 1981, Heicappell *et al*., 1986).

In both tumour models it was shown that vaccination leads to a local inflammatory reaction which extends to the draining lymph nodes and eventually leads to a systemic immunity and inflammatory reaction at sites of metastases. In the guinea pig, the interconnections between local vaccine reactions and systemic immunity were evaluated by means of removal of the local reaction sites of the skin or of the draining lymph nodes at different times and testing the effects on postoperative survival. The stronger the local reaction, the better was the extent of systemic immunity induced. The combination of autologous tumour cells and BCG in the vaccine had a synergistic effect both on the intensity of the local reaction and on the protective immunity. In both models it was documented that the viability of the inactivated tumour cells was of decisive importance and that the cells had to be frozen under controlled conditions otherwise their immunogenicity was greatly decreased. In the Esb–tumour system, we found that simple shock freezing/thawing of the tumour cells completely destroyed their immunogenicity with regard to stimulating tumour-specific CTL in mixed lymphocyte–tumour cell (MLTC) cultures (Schirrmacher *et al*., 1993). This defect could not be restored by the addition of cytokines such as IL-1, IL-2 or interferon-γ. The stronger immunogenicity of intact tumour cells compared with membrane preparations could be due to accessory adhesive and signalling molecules.

In a comparative study we found that vaccination with either L10/BCG or with L10-NDV vaccine induced similar protective antitumour immunity (Bier *et al*., 1989). A combination of the two types of vaccine did not lead to better protective immune effects. BCG induces long-lasting ulcers in animals and patients whereas NDV does not. Another advantage of NDV may be its more precise localization at the tumour cell surface where immune cell–tumour cell interactions take place. When we analysed the immune responses of animals inoculated in the pinna with either live ESb or live ESb-NDV cells we found three levels at which NDV exerted an amplifying effect on the immune response: (1) an increased CD4 T cell-mediated helper response, (2) an increased frequency of tumour-specific CTL precursors and (3) an increased stimulatory capacity of NDV-modified ESb cells for activating tumour-specific CTL precursors to mature CTL *in vitro*. Advantages of NDV are its good cell-binding properties, its selective replication in tumour cell cytoplasm, which is independent of cell proliferation, and its relative safety (Schirrmacher *et al*., 1998). Most important for its use as an adjuvant in human cancer vaccines are its ability to introduce T cell costimulatory activity, to prevent anergy induction and to induce locally chemokines (e.g. RANTES, IP10) and cytokines (e.g. interferon-α, T-β) that affect T cell recruitment and activation. A further development consists in attachment, via NDV-derived haemagglutinin–neuraminidase (HN) membrane anchoring molecules, of universal defined bispecific reagents such as T cell activating anti-CD28 antibodies (Haas *et al*., 1998).

Clinical ASI Studies

The present status of our clinical studies with the autologous NDV-modified high-quality live tumour cell vaccine ATV-NDV which have been ongoing since 1988 has recently been updated. These include Kaplan–Meier probability plots of postoperative survival of cancer patients from five different phase II vaccination studies. Comparisons are made between groups of patients with similar prognostic parameters. In one study (glioblastoma), the vaccinated group (ASI) was compared with a historical control from the same clinic. In the other four studies

(colorectal carcinoma Dukes' C (Ockert 1996); primary operated mammary carcinoma (Ahlert, 1997); melanoma and ovarian carcinoma (Ahlert *et al*., 1997)), both groups of patients received ASI treatment simultaneously. The ATV-NDV vaccine, prepared after purification starting with 10^7 cells, differed with respect to tumour cell number and cell viability (Ahlert *et al*., 1997). Therefore, retrospectively, groups were formed and compared with similar prognostic parameters that differed only with respect to postulated quality parameters of the vaccine. In all five studies, postoperative treatments with high-quality ATV-NDV vaccine showed a strong trend for improval of survival in comparison with patients treated with low-quality vaccine or in comparison with historical controls (Ockert *et al*., 1996). In addition, there were various signs of systemic effects which support the notion of effectivity. A prospective randomized, controlled phase III study with the ATV-NDV vaccine in primary operated breast-cancer patients in the medium risk category has recently been started. The design of this study is shown in **Figure 2**.

Adoptive Cellular Immunotherapy (ADI) with Immune Cells for Treatment of Macrometastases

We established in the ESb/ESb-MP lymphoma model a cellular cancer therapy with unique efficiency even in late-stage advanced metastasized cancer. *In situ*-activated tumour immune T cells, induced in allogeneic, tumour-resistant, MHC-identical but superantigen-different donor mice (B10.D2), could transfer strong graft versus leukaemia (GvL) effects accompanied by only mild graft versus host (GvH) reactivity. Systemic immune cell transfer into 5 Gy-irradiated DBA/2 mice bearing up to 4 week established syngeneic tumours and macro-metastases led to massive infiltration of the tumour tissues by CD4 and CD8 donor T lymphocytes. Upon interactions of these T cells with host antigen-presenting cells, primary tumours (>1 cm diameter) were encapsulated and rejected from the skin and liver metastases were eradicated. The effects of this effective ADI therapy could be followed in individual animals by ^{31}P-NMR spectroscopy of primary tumours and by ^{1}H-NMR imaging of metastases. An approximately 25 000-fold excess of metastatic tumour cells could be rejected, as revealed quantitatively by FACSscan analysis of *lacZ* gene-transfected tumour cells.

This T cell-mediated therapy caused recruitment and activation of a distinct host macrophage population into the liver bearing the lymphocyte adhesion molecule sialoadhesin (Sn) (Umansky *et al*., 1995a). We demonstrated *in vitro* that these Sn^+ macrophages were capable of processing and presenting human C-reactive protein via MHC class II to a CD4 T helper clone specific for an identified MHC II associated C-reactive protein-derived peptide (Umansky *et al*., 1995a). Furthermore, Sn^+ macrophages which were isolated from the liver of ADI-treated animals could directly restimulate *in vitro* CD8 T lymphocytes primed against the tumour cells, suggesting that they could process and present exogenous TAA to T lymphocytes (Rocha *et al*., 1997a). Moreover, using as *in vitro* model the antigen β-galactosidase, for which a dominant MHC class I L^d restricted peptide epitope is known to be recognized by specific CD8 CTLs, we demonstrated that purified Sn macrophages can process

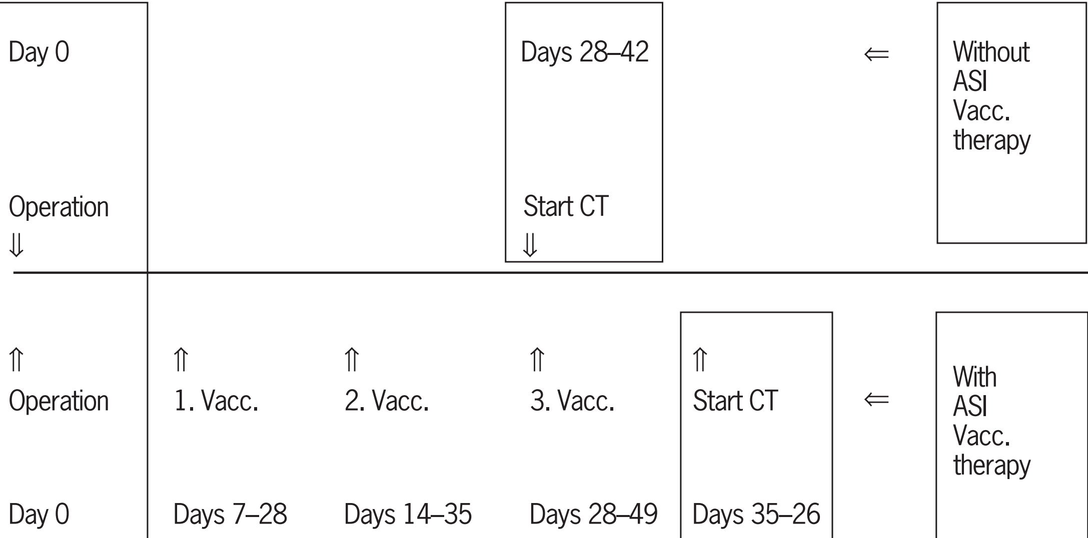


Figure 2 Protocol of a phase II/III clinical study to test for efficacy of active specific immunotherapy (ASI) of the autologous virus-modified tumour vaccine ATV-NDV in primary operated breast-cancer patients. It is a multicentre, prospective, double-blind, placebo-controlled, randomized parallel-group study (protocol No. ASI A-01). CT = chemotherapy.

exogenous β-galactosidase and stimulate peptide specific class I MHC restricted CTL responses (Muerköster *et al.*, 1999). Sn macrophages which are significantly increased in the liver after ADI played an essential role in GvL reactivity (Muerköster *et al.*, 1999). They appeared to be able to process tumour-derived proteins via the MHC class I and also via the MHC class II pathway and present peptide epitopes to CD8 and to CD4 immune T cells, respectively. The synergistic interactions demonstrated before between immune CD4 and CD8 T cells during ADI could thus occur in the observed clusters (Rocha *et al.*, 1997a) in the liver between such T cells and host Sn^+ macrophages.

Breaking Tolerance via GvL Reactivity

A major goal in tumour immunotherapy consists in breaking potential tumour-specific T cell unresponsiveness (tolerance) that may explain tumour growth in cancer patients. In the above GvL/GvH/ADI model we were able to demonstrate that immunological tolerance to a tumour-associated viral superantigen (SAG) is overcome by transfer of allogeneic T cells expressing SAG-reactive Vβ6 T cell receptor chains (Schirrmacher *et al.*, 2000b). DBA/2 mice and the derived tumour ESb-MP express vSAG-7 (Mls^a), an endogenous viral superantigen which is absent in B10.D2 mice. In DBA/2 mice, virtually no Vβ6 SAG-reactive cells can be detected owing to central tolerance established by negative selection in the thymus. In contrast, in B10.D2 mice, about 10% of peripheral CD4 and CD8 T cells express Vβ6 T cell receptor chains. Upon contact with SAG-expressing lymphoma cells, these Vβ6 T cells became activated rather than tolerated (as reported previously) and developed SAG-specific cytotoxic T lymphocyte activity and secreted IL-2 and interferon-γ. The grafted T cells infiltrated liver metastases, formed close contact and clusters (Muerköster *et al.*, 1998) with SAG-expressing tumour cells and caused significant GvL effects. Selection for tumour resistance among the progeny from a cross between SAG-negative donor and SAG-positive recipient strains revealed a strict correlation between loss of the endogenous SAG tolerogen, rescue of Vβ6 T cells from SAG-mediated deletion and leukaemia resistance. These findings suggest that immune responses to superantigens can be exploited to break tolerance and augment immune responses to tumours (Schirrmacher *et al.*, 2000b).

BIOLOGICAL LIMITATIONS

Specific cancer immunotherapy approaches also have their limits. Tumour genomic instability, heterogeneity and plasticity allow the selection of immune escape variants. Major restrictions to the effectivity of specific immunotherapy relate to complex phenomena such as (1) tumour epitope loss, (2) immunosuppressive tumour products and (3) defective tumour capillaries. It is beyond the scope of this chapter to go into more detail on tumour immune escape mechanisms and readers are therefore referred to literature reviews (Eibl *et al.*, 1997, 1998; Coggin *et al.*, 1998; Ellem *et al.*, 1998; Glinsky, 1998). The capillary endothelium of the intratumour neovasculature is frequently defective, particularly in presenting the homing signals that facilitate the massive diapedesis of tumour-reactive T cells into tumour tissue (Ellem *et al.*, 1998).

TECHNICAL CONSIDERATIONS

Dissection of Tumour Microenvironment of Liver Metastases for Studies on Tumour–Host Interactions

The interaction between tumour and host cells determines to a large extent the outcome, namely tumour growth and progression towards metastases or tumour arrest, dormancy or rejection. Most of the studies published so far on interactions of tumour cells and host cells were made *in vitro* and dealt with aspects such as cell adhesion, proliferation, invasiveness, cytotoxicity or cytokine production. Cells which are taken out from their respective tissues may suffer from mechanical stress during cell isolation. When subjected to tissue culture they are in an entirely different situation from *in vivo*. Often they lack cell–cell contact or contact with surrounding extracellular matrix. Since the microenvironment in tissue culture differs in many respects from that *in vivo*, new approaches for *in vivo* studies or tumour–host cell interactions are of utmost importance in cancer research. To elucidate the metastatic phenotype, approaches have been made to relate, for instance, cell surface molecules expressed on the tumour cell lines from tissue culture to their propensity to generate metastases *in vivo*. Several authors have reported that certain steps of the metastatic cascade are rate limiting (Fidler and Radinsky, 1990; Kerbel, 1990; Nicolson, 1993). To produce metastases, tumour cells must complete each of the sequential steps in the pathogenesis of cancer metastasis. Each discrete step appears to depend on the interaction between tumour cells and multiple host factors (i.e. the microenvironment of the tumour) and to be regulated by transient or permanent changes in DNA, RNA or proteins of multiple genes. With this background, the need for comprehensive *in vivo*/*ex vivo* studies on tumour–host interactions and their kinetics in relevant model systems becomes obvious. (See also chapter on *Modelling Tumour Tissue Interactions*.)

We have established a new method allowing the *ex vivo* isolation of tumour and host cells (tumour microenvironment) at any time point during the metastatic process and without any further *in vitro* culture. We chose as a model the ESbL lymphoma, transduced with the *lacZ* gene

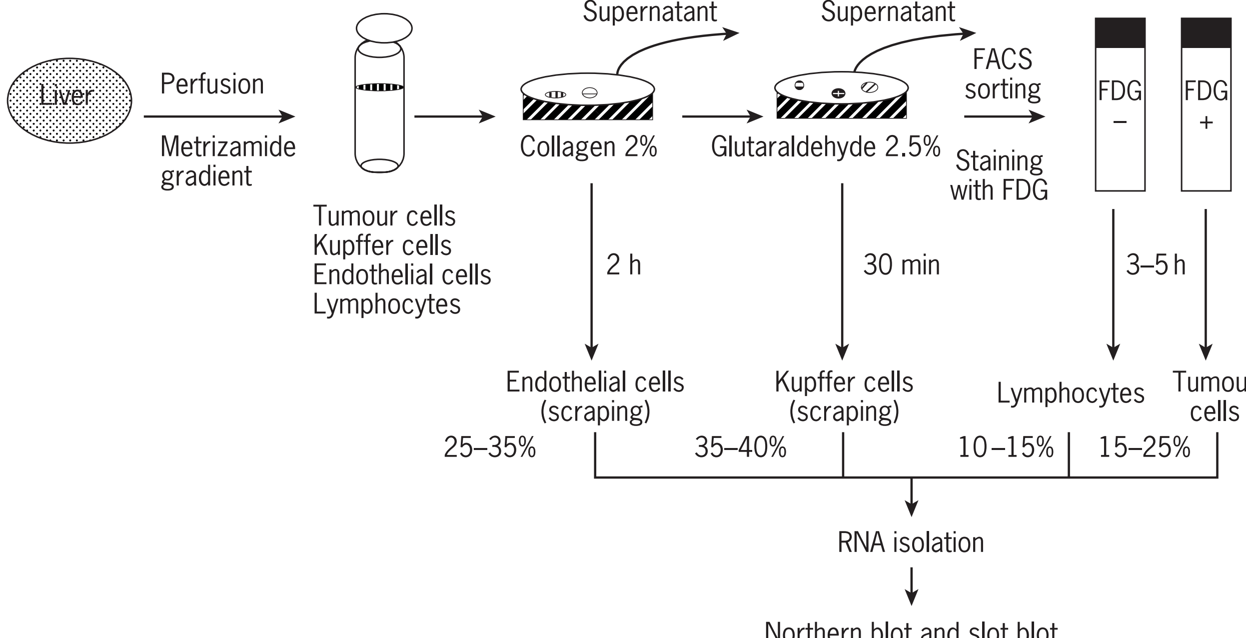


Figure 3 Method of *ex vivo* separation of cells from a liver metastasis for preparation of total RNA. Sinusoidal cells and tumour cells from livers were fractionated into subpopulations by differential adhesion (endothelial and Kupffer cells) and by FACS (lymphocytes and tumour cells). Numbers indicate the range in percentage of cells obtained in each purification step relative to the total number of live recovered cells.

(**Figure 3**). In this ESbL-*lacZ* tumour model, intradermal tumour growth and liver metastasis development followed three distinct phases: a first exponential growth phase, a transient plateau phase and a second expansion phase (Krüger *et al*., 1994a, b). The plateau phase was characterized by a constant tumour diameter, correlating with a constant small amount of metastasis in the liver. This phase was followed by an aggressive second expansion phase leading to macroscopic metastases in multiple visceral organs and to death of the animals within a few days (Krüger *et al*., 1994b). These different phases of the metastatic process provide a good model for the investigation of the influence of tumour–host interactions during metastasis and of means for its modulation. The plateau phase allows one to study mechanisms of immunoresistance and tumour control. The second tumour expansion phase allows the study of mechanisms of breakdown of immunoresistance and perhaps molecules of importance in metastatic progression.

Isolation of Tumour and Sinusoidal Cells from Liver Metastases

Livers from anaesthetized tumour-bearing mice were washed *in situ* by perfusion through the portal vein and tissue digestion was carried out during perfusion with pronase and collagenase. Tumour and sinusoidal cells were separated by a metrizamide gradient as described (Rocha *et al*., 1996b).

Isolation of Liver Endothelial Cells

Liver endothelial cells were isolated by co-culturing liver sinusoidal and tumour cells on collagen-coated plastic Petri dishes. The purity of endothelial cell populations was evaluated as described (Rocha *et al*., 1995).

Isolation of Kupffer Cells from Livers

Kupffer cells were isolated after differential adhesion on 2.5% glutaraldehyde Petri dishes (Umansky *et al*., 1995b). The purity of Kupffer cell population was evaluated as described (Umansky *et al*., 1995b).

Isolation of Lymphocytes and Tumour Cells from Liver Metastases

Separation of tumour cells from lymphoid cells was performed as described (Krüger *et al*., 1994b). A window for sorting was defined in the FDG-positive cells (tumour cells) excluding dead cells and debris with propidium iodide. The flow rate was 3000–5000 cells s^{-1} and sorted cells were collected in sterile tubes containing RPMI with 20% FCS.

Flow Cytometry Analysis of Tumour and Host Cells

Antibody Staining of Tumour and Sinusoidal Cells

As a standard protocol, 1×10^6 cells were incubated at 4 °C for 10 min with the first antibodies following a described protocol (Rocha *et al*., 1996a). To test the purity of the separated populations, the following rat anti-mouse monoclonal antibodies were used as culture supernatants: anti-CD4 (clone GK 1.5); E-selectin (clone 21KC10), specific for endothelial cells; anti-macrophage antibody (F4/80). After washing, the cells were incubated with the second antibody ($F(ab')_2$ goat anti-rat, mouse Ig absorbed, (R)-phycoerythrin conjugated) and analysed.

FDG Staining and FACS Analysis of Tumour and Host Cells

After antibody staining, liver metastases were quantified at the single cell level by loading and staining with FDG as described (Rocha *et al*., 1996a). Samples were simulaneously measured for FSC and integrated side scatter (SSC) as well as green (FL1) and red (FL2 and FL3) fluorescence (expressed as logarithm of the integrated fluorescence light). Recordings were made only on propidium iodide-negative (viable) cells of the red (FL3) fluorescence, excluding aggregates whose FSCs were out of range. *Ex vivo* expression of cell surface molecules was analysed by histograms of red fluorescence (FL2) distribution plotted as number of cells (y-axis) against fluorescence intensity (x-axis) for the different tumour and sinusoidal cell populations (Rocha *et al*., 1996a).

Gene Expression Studies: RNA Extraction, Hybridization and Densitometric Quantitation

RNA extraction was performed by the chloroform–phenol technique. Extracted RNA was precipitated with propan-2-ol and the pellet washed in ethanol, dried under vacuum and resuspended in 0.1% diethyl pyrocarbonate solution. The quantity of RNA was determined by measuring the absorbance at 260 nm. Total isolated RNA was denatured and spotted on to a nitrocellulose filter using a dot blot apparatus and fixed by UV cross-linking under vacuum. The following cDNA probes were used for hybridization: MHC class II, α and β chain, ICAM-1 and integrin $\alpha4$ chain. cDNA inserts were labelled with ^{32}P to a specific activity of about 2×10^8 cpm mg^{-1} DNA with an oligo-labelling kit. After hybridization, the filters were washed three times for 30 min with sodium chloride (3 mol L^{-1})–trisodium citrate (0.3 mol L^{-1}) solution and sodium dodecyl sulfate at 68 °C. Membranes were exposed to O-MAT films at −70 °C. Expression of the mRNA was quantified by densitometry of autoradiograms using the Adobe Photoshop program and the SCAN analysis program from Macintosh with each sample measurement calculated from the ratio of the average areas between the specific mRNA transcripts and the β-actin mRNA transcripts (Rocha *et al*., 1996b). An application of this technology revealed *in situ* downregulation of VLA-4 integrin cell surface expression during lymphoma growth and liver metastases (Rocha *et al*., 1997b).

PERSPECTIVES

We have tried in this review to demonstrate how the use of model systems can lead to the development of new strategies for the immunotherapy of metastases. We have primarily focused on our own research relating (1) to the virus-modified autologous live cell vaccine ATV-NDV developed for ASI-treatment and (2) to cell-based ADI studies in a GvL/GvH animal model system. Both ASI- and ADI-type immunotherapies are being evaluated in clinical trials. Positive results have been reported recently from a randomized controlled ASI study (Vermorken *et al*., 1999) and also from ADI studies in particular when using allogeneic donor lymphocyte infusions (DLI) to obtain graft versus leukaemia (GvL) effects (Slavin *et al*., 1996; Collins *et al*., 1997). ADI with autologous lymphokine-activated killer (LAK) cells or with tumour-infiltrating lymphocytes (TIL) was reported to have an overall response rate of about 25% in metastatic melanoma (Rosenberg *et al*., 1993) and renal cell carcinoma patients (Osband *et al*., 1990). Gene-modified cell therapies (Gage, 1998) are also being developed for gene therapy of cancer.

Apart from ASI and ADI immunotherapies, a variety of other interesting new strategies have been developed which could not be discussed because of space limitations. With regard to antibody-based therapies, including apoptosis-inducing monoclonal antibodies (mAbs), recombinant antibody-fusion proteins, bispecific antibodies and antibodies engineered for tumour targeting and pro-drug activation, we refer to our review from 1996 (see Further Reading). Some mAbs have already been proved clinically effective for adjuvant immunotherapy (Riethmüller *et al*., 1994).

New techniques are being developed for the identification and cloning of genes for T cell targets and the identification of T cell epitopes (Bona *et al*., 1998) and mimotopes (Chen, 1999) by sequencing peptides eluted from HLA molecules of tumour cells or from peptide libraries (Chen, 1991). Based on such specific information about genes and peptides of TAAs, new gene and/or peptide vaccination strategies are likely to be developed. When combined with our increasing ability to manipulate the immune system via cytokines, costimulatory molecules, delivery systems and new adjuvants, these advances

provide hope for effective immunotherapy and vaccination against cancer.

Two other developments open further exciting perspectives for the future: DCs (Steinmann, 1991) and heat-shock proteins (HSPs) (Srivastava, 1997). Since DCs are a unique leukocyte population which control the primary immune response, it is to be expected that many new technologies will be developed for dendritic cell-based vaccines to be used for immunotherapy of cancer. This can involve the loading with peptides or tumour lysates (Nestlé *et al*., 1998), transfection with tumour-associated DNA or RNA (Boczkowski *et al*., 1996) or fusion of DCs with carcinoma cells (Gong *et al*., 1997). HSPs of different molecular mass have been demonstrated to carry representative peptides of tumour cells and are highly efficient for the induction of protective immunity based on oligoclonal T cell priming. It can be expected that standard protocols will allow the preparation of HSPs from individual cancer patients for autologous vaccination purposes (Srivastava, 1997).

Another future perspective relates to the new technologies of MHC-peptide tetramers for the direct visualization of TAA-specific T cells (Crawford *et al*., 1998; Romero *et al*., 1998) and of quantitative ELISPOT assays for cytokine-producing T cells. Both assays provide new diagnostic tools to evaluate the immune status of cancer patients and their response to therapy.

All these new approaches for immunotherapy can be expected to have an effect on microdisseminated carcinoma cells. As the disease progresses towards macrometastatic disease, the risk increases that the tumour may develop immune escape variants and immunosuppression which could be formidable barriers to effective cancer immunotherapy. Perhaps combined T cell and antibody-based immunotherapy approaches may outsmart even such tumour cell variants (Abken *et al*., 1998). Anyway, it is extremely important that clinical experiments are well thought out, are started early enough and are carefully designed. Immune monitoring should include measurements of the immune response and characterization of HLA expression on tumours. Only in this way will immunotherapy of metastases eventually receive the attention that it undoubtedly deserves.

REFERENCES

Abken, H., *et al*. (1988). Can combined T-cell and antibody-based immunotherapy outsmart tumor cells? *Immunology Today*, **19**, 1–5.

Ahlert, T., *et al*. (1997). Tumor cell number and viability as quality and efficacy parameters of autologous virus modified cancer vaccines. *Journal of Clinical Oncology*, **15**, 1354–1366.

Ahmed, R. and Gray, D. (1996). Immunological memory and protective immunity: understanding their relation. *Science*, **272**, 54–60.

Bier, H., *et al*. (1989). Postoperative active-specific immunotherapy of lymph node micrometastasis in a Guinea pig tumor model. *Otorhinopharyngology*, **51**, 197–205.

Boczkowski, D., (1996). Dendritic cells pulsed with RNA are potent antigen-presenting cells *in vivo* and *in vitro*. *Journal of Experimental Medicine*, **184**, 465.

Bona, C. A., *et al*. (1998). Towards development of T-cell vaccines. *Immunology Today*, **19**, 126–132.

Chen, L. (1999). Mimotopes of cytolytic T lymphocytes in cancer immunotherapy. *Current Opinions Immunology*, **11**, 219–222.

Coggin, I. H., *et al*. (1998). Tumors express both unique TSTA and crossprotective 44 kDA oncofetal antigen. *Immunology Today*, **19**, 405–408.

Collins, R. H., *et al*. (1997). Donor leukocyte infusions in 140 patients with relapsed malignancy after allogeneic bone marrow transplantation. *Journal of Clinical Oncology*, **15**, 433–444.

Crawford, F., *et al*. (1998). Detection of antigen-specific T cells with multivalent soluble class II MHC covalent peptide complexes. *Immunity*, **8**, 675–682.

Dasgupta, D. (1999). *Artificial Immune Systems and Their Applications*. (Springer, Berlin).

Eibl, M. M., *et al*. (1997). *Symposium in Immunology VI. Tumor Immunology*. (Springer, Berlin).

Eibl, M. M., *et al*. (1998). *Symposium in Immunology VII. Vaccination*. (Springer, Berlin).

Ellem, K. A., *et al*. (1998). The labyrinthine ways of cancer immunotherapy – T cell, tumor cell encounter: "How do I lose Thee? Let me count the ways". *Advances in Cancer Research*, **75**, 203–249.

Fichtner, K. P., *et al*. (1997). *In vivo* ^{1}H-NMR microimaging with respiratory triggering for monitoring adoptive immunotherapy of metastatic mouse lymphoma. *Magnetic Resonance in Medicine*, **38**, 440–455.

Fidler, I. J. and Radinsky, R. (1990). Genetic control of cancer metastasis. *Journal of the National Cancer Institute*, **82**, 166–168.

Fodstad, O. (1993). Metastatic ability of cancer cells: pheno- and genotypic characteristics and role of the micro-environment. In: Iversen, O. H. (ed.), *New Frontiers in Cancer Causation*. (349–358) (Taylor and Francis, Washington, DC).

Förg, P., *et al*. (1998). Superiority of the ear pinna over muscle tissue as site for DNA vaccination. *Gene Therapy*, **5**, 789–797.

Gage, F. H. (1998). Cell therapy. *Nature*, **393**, 18–24.

Glinsky, G. V. (1998). *Failure of Apoptosis and Cancer Metastasis*. (Springer, Berlin).

Gong, J., *et al*. (1997). Induction of antitumor activity by immunization with fusions of dendritic cells and carcinoma cells. *Nature Medicine*, **3**, 558.

Haas, C., *et al*. (1998). Bispecific antibodies increase T cell stimulatory capacity *in vitro* of human autologous virus modified tumor vaccine. *Clinical Cancer Research*, **4**, 721–730.

Hart, I. R., *et al*. (1989). Molecular aspects of the metastatic cascade. *Biochimica Biophysica Acta*, **989**, 65–84.

Heicappell, R., *et al*. (1986). Prevention of metastatic spread by postoperative immunotherapy with virally modified autologous tumor cells. *International Journal of Cancer*, **37**, 569–577.

Jurianz, K., *et al*. (1998). Superiority of the ear pinna over a subcutaneous tumor inoculation site for induction of a TH1 type cytokine response. *Cancer Immunology and Immunotherapy*, **45**, 327–333.

Kerbel, R. S. (1990). Growth dominance of the metastatic cancer cell: cellular and molecular aspects. *Advances in Cancer Research*, **55**, 87–132.

Key, M. E., *et al*. (1981). Mechanism of action of BCG-tumor cell vaccines in the generation of systemic tumor immunity. *Journal of the National Cancer Institute*, **67**, 863–869.

Khazaie, K., *et al*. (1993). EGF receptor in neoplasia and metastasis. *Cancer and Metastasis Reviews*, **12**, 255–274.

Khazaie, K., *et al*. (1994). Persistance of dormant tumor cells in the bone marrow of tumor cell-vaccinated mice correlates with long-term immunological protection. *Proceedings of the National Academy of Sciences of the USA*, **91**, 7430–7434.

Krüger, A., *et al*. (1994a). Scattered micrometastasis visualized at the single cell level: detection and re-isolation of lacZ labeled metastasized lymphoma cells. *International Journal of Cancer*, **58**, 275–284.

Krüger, A., *et al*. (1994b). Pattern and load of spontaneous liver metastasis dependent on host immune status studied with a lacZ transduced lymphoma. *Blood*, **84**, 3166–3174.

Krüger, A., *et al*. (1999). The bacterial lacZ gene: an important tool for metastasis research and evaluation of new cancer therapies. *Cancer and Metastasis Reviews*, **17**, 285–294.

Lindner, M. (1998). Etablierung eines murinen Karzinommodells zur Untersuchung des immuntherapeutischen Potentials von Fusionsprodukten aus Tumor- und Dendritischen Zellen. *PhD Thesis*, University of Heidelberg.

Müerköster, S., *et al*. (1998). GvL reactivity involves cluster formation between superantigen-reactive donor T-lymphocytes and host macrophages. *Clinical Cancer Research*, **4**, 3095–3106.

Müerköster, S., *et al*. (1999). Sialoadhesin-positive host macrophages play an essential role in graft versus leukemia (GvL) reactivity in mice. *Blood*, **93**, 4375–4386.

Müller, M., *et al*. (1998). Eb-lacZ tumor dormancy in bone marrow and lymph nodes: active control of proliferative tumor cells by $CD8^+$ immune T cells. *Cancer Research*, **58**, 5439–5446.

Nestlé, F. O., *et al*. (1998). Vaccination of melanoma patients with peptide or tumor lysate-pulsed dendritic cells. *Nature Medicine*, **4**, 328.

Nicolson, G. L. (1993). Cancer progression and growth: relationship of paracrine and autocrine growth mechanisms to organ preference of metastasis. *Experimental Cell Research*, **204**, 171–180.

Ockert, D., *et al*. (1996). Newcastle disease virus infected intact autologous tumor cell vaccine for adjuvant active specific immunotherapy of resected colorectal carcinoma. *Clinical Cancer Research*, **2**, 21–28.

Osband, M. E., *et al*. (1990). Effect of autolymphocyte therapy on survival and quality of life in patients with metastatic renal cell carcinoma. *The Lancet*, **335**, 994–998.

Riethmüller, G., *et al*. (1994). Randomized trial of monoclonal antibody for adjuvant therapy of resected Duke's C colorectal carcinoma. *The Lancet*, **343**, 1177–1183.

Rocha, M., *et al*. (1995). Liver endothelial cells participate in T cell dependent host resistance to lymphoma metastasis by production of nitric oxide *in vivo*. *International Journal of Cancer*, **63**, 405–411.

Rocha, M., *et al*. (1996a). Dynamic expression changes *in vivo* of adhesion and costimulatory molecules determine load and pattern of lymphoma liver metastasis. *Clinical Cancer Research*, **2**, 811–820.

Rocha, M., *et al*. (1996b). Dissection of tumor and host cells from a metastasized organ for testing gene expression directly *ex vivo*. *British Journal of Cancer*, **74**, 1216–1222.

Rocha, M., *et al*. (1997a). Differences between graft-versus-leukemia (GvL) and graft-versus-host (GvH) reactivity. I. Interaction of donor immune T cells with tumor and/or host cells. *Blood*, **6**, 2189–2202.

Rocha, M., *et al*. (1997b). *In situ* downregulation of VLA-4 integrin cell surface expression during lymphoma growth and liver metastasis. *International Journal of Oncology*, **10**, 457–464.

Romero, P., *et al*. (1998). *Ex vivo* staining of metastatic lymph nodes by class I major histocompatibility complex tetramers reveals high numbers of antigen-experienced tumor-specific cytolytic T lymphocytes. *Journal of Experimental Medicine*, **188**, 1641–1650.

Rosenberg, S. A., *et al*. (1993). Prospective randomized trial of high-dose interleukin-2 alone or in conjuction with lymphokine-activated killer cells for the treatment of patients with advanced cancer. *Journal of the National Cancer Institute*, **85**, 622–632.

Schirrmacher, V. and von Hoegen, P. (1993). Importance of tumor cell membrane integrity and viability for CTL activation by cancer vaccines. *Vaccine Research*, **2**, 183–196.

Schirrmacher, V., *et al*. (1982). Antigenic variation in cancer metastasis: immune escape versus immune control. *Cancer and Metastastics Reviews*, **19**, 241–274.

Schirrmacher, V., *et al*. (1992). Tumor-specific CTL response requiring interactions of four different cell types and recognition of MHC class I and class II restricted tumor antigens. *Immunological and Cellular Biology*, **71**, 311–326.

Schirrmacher, V., *et al*. (1998). Human tumor cell modification by virus infection: an efficient and safe way to produce cancer vaccine with pleiotropic immune stimulatory properties when using Newcastle disease virus. *Gene Therapy*, **6**, 63–73.

Schirrmacher, V., *et al*. (2000a). Intra-pinna anti-tumor vaccination with self-replicating infectious RNA or with DNA encoding a model tumor antigen and a cytokine. *Gene Therapy*, **7**, 1137–1147.

Schirrmacher, V., *et al*. (2000b). Breaking tolerance to a tumor-associated viral super-antigen as a basis for graft versus leukemia (GvL) reactivity. *International Journal of Cancer*, **87**, 695–706.

Slavin, S., *et al*. (1986). Allogeneic cell therapy with donor peripheral blood cells and recombinant human interleukin-2 to treat leukemia relapse post allogeneic bone marrow transplantation. *Blood*, **87**, 2195–2204.

Srivastava, P. K. (1997). *Methods*, **12**, 165–171.

Steinmann, R. M. (1991). The dendritic cell system and its role in immunogenicity. *Annual Reviews of Immunology*, **9**, 271.

Umansky, V., *et al*. (1995a). A role for sialoadhesin-positive tissue macrophages in host resistance to lymphoma metastasis *in vivo*. *Immunology*, **87**, 303–309.

Umansky, V., *et al*. (1995b). *In situ* activated macrophages are involved in host resistance to lymphoma metastasis by production of nitric oxide. *International Journal of Oncology*, **7**, 33–40.

Vermorken, J. B., *et al*. (1999). Active specific immunotherapy for stage II and stage III human colon cancer: a randomized trial. *The Lancet*, **353**, 345–350 .

FURTHER READING

Farzaneth, F., *et al*. (1998). Gene therapy of cancer. *Immunology Today*, **19**, 294–296.

Finke, J., *et al*. (1999). Where have all the T cells gone? Mechanisms of immune evasion by tumors. *Immunology Today*, **20**, 158–160.

Hellström, K. E., *et al*. (1996). Can costimulated tumor immunity be therapeutically efficacious? *Immunological Reviews*, **145**, 123–145.

Herberman, R. B., *et al*. (1987). *Immune Responses to Metastases*, Vol. II (CRC Press, Boca Raton, FL).

Ockert, D., *et al*. (1999). Advances in cancer immunotherapy. *Immunology Today*, **20**, 63–65.

Pardoll, D. M. (1995). Paracrine cytokine adjuvants in cancer immunotherapy. *Annual Reviews in Immunology*, **13**, 399–416.

Schirrmacher, V. (1985). Cancer metastasis: experimental approaches, theoretical concepts and impacts for treatment strategies. *Advances in Cancer Research*, **43**, 1–73.

Schirrmacher, V., *et al*. (1998). Immunization with virus modified tumor cells. *Seminars in Oncology*, **25**, 677–696.

Schirrmacher, V., *et al*. (1995). Effective immune rejection of advanced metastasized cancer. *International Journal of Oncology*, **6**, 505–521.

Schirrmacher, V., *et al*. (1996). Immunotherapy of metastases. In: Günthert, U., *et al*. (eds), *Current Topics in Microbiology and Immunology (CTMI), Vol. 213/III*: *Attempts to Understand Metastasis Formation III*. 189–216 (Springer, Berlin).

Chapter 69

Gene Therapy Models

Justin C. Roth, Steven P. Zielske, Punit D. Wadhwa, Christopher B. Ballas, Janice E. Bowman, Jane S. Reese and Stanton L. Gerson

University Hospitals of Cleveland and Case Western Reserve University, Cleveland, OH, USA

CONTENTS

- Suicide Gene Therapy
- Tumour-Suppressor Gene Therapy
- Oncolytic Viruses
- Antiangiogenic Gene Therapy
- Cytokine Gene Therapy
- Gene-Modified Dendritic Cells
- Drug Resistance Gene Therapy
- Acknowledgement

SUICIDE GENE THERAPY

Suicide genes are negative selectable markers encoding for enzymes that convert prodrugs into compounds that are toxic to the genetically modified cell. This approach, designed to alter the drug-sensitivity profile of tumours, is a two-step process: in the first step the gene for the foreign enzyme is delivered and targeted within a viral vector or using a naked DNA fragment to the tumour where it is to be expressed. Thereafter the prodrug is administered, which is converted to the active cytotoxic compound by the foreign enzyme expressed by the tumour cells. An inherent limitation of most viral-based cancer gene-therapy protocols is the inability to transduce the entire population of a tumour mass; this may be overcome (at least partly) by the ability of gene-modified cancer cells to mediate the killing of adjacent untransduced cells (the so-called bystander effect, discussed below). (See also chapter on *Genetic Prodrug Activation Therapy*.)

The best characterized 'suicide gene' is the herpes simplex virus thymidine kinase (HSV-*tk*) gene, which mediates phosphorylation of the prodrug ganciclovir 1000 times more efficiently than its mammalian counterpart. The phosphorylated forms of ganciclovir bring about DNA chain termination and cause single-strand breaks, leading to cell death. Similarly, the *Escherichia coli* gene cytosine deaminase (CD), when expressed in mammalian cells, can convert the nontoxic compound 5-fluorocytosine (5FC) to the cytotoxic compound 5-fluorouracil (5FU).

Both the *CD* and HSV-*tk* genes have been validated in multiple *in vitro* systems, wherein tumour cell lines tested demonstrate dose-dependent inhibition of proliferation upon exposure to the prodrugs 5FC and ganciclovir, respectively. In their *in vivo* model, Bentires-Alj *et al*. (2000) used either parental or *CD* transfected DHD/K12 colorectal cancer cells to establish peritoneal carcinomatosis in BDIX rats. Following 3 weeks of intraperitoneal 5FC injections, rats treated with the *CD* transfected cell line demonstrated total regression of their peritoneal tumours, as opposed to rats treated with parental DHD/K12 cells, who showed extensive disseminated peritoneal carcinomatosis. Six of eight rats treated with the DHD/K12-CD cells were alive at 360 days, as opposed to all control rats, which died between days 45 and 71.

The 'bystander effect' is a phenomenon defined by the ability of gene-transduced tumour cells to mediate cytotoxicity to adjacent untransduced tumour cells. Although there is no single explanation for this phenomenon, a number of alternative hypotheses have been generated. These include direct drug transfer, immune-mediated responses and intercellular induction of apoptotic signals. First is the proposal for cell–cell transfer of ganciclovir metabolites through intercellular gap junctions made up of proteins called 'connexins;' other mechanisms include phagocytosis of 'apoptotic vesicles' containing phosphorylated ganciclovir metabolites by untransduced tumour cells, or the disruption of tumour vasculature leading to a haemorrhagic tumour necrosis mediated by the release of cytokines such as TNF-α from transduced cells.

Mesnil *et al*. (1996) established the importance of intercellular gap junctions (connexins) with their studies conducted on HeLa cancer cells, which express very low basal levels of connexins. Upon transfection of the HSV-*tk* gene, the HeLa cells became highly sensitive to the cytotoxic effect of ganciclovir; however, they failed to mediate killing of untransfected HeLa cells in 'mixing'

co-culture studies, i.e. no *in vitro* bystander effect was observed. When HeLa cells were transfected with the gene encoding for the gap junction protein connexin 43, significant killing of HSV-tk^{-} cells occurred with the presence of only 10% HSV-tk^{+} cells in the co-culture, suggesting a powerful *in vitro* bystander effect.

An immune response may also occur. In an *in vivo* model of CNS glioma in Lewis rats, Dewey *et al*. (1999) were able to inhibit effectively the growth of established syngeneic gliomas using intratumoural injections of an adenoviral vector expressing the HSV-*tk* gene followed by systemic ganciclovir treatment, with 100% survival at 3 months and no evidence of residual glioma assessed by weekly magnetic resonance imaging. However, when survivors were killed at 100 days, evidence of diffuse perivascular inflammatory infiltrates and secondary demyelination was observed, along with persistent transgene (HSV-*tk*) expression in both cerebral hemispheres. (See also chapter on *Modelling Tumour–Tissue Interactions*.)

To improve the efficacy of the suicide gene therapy system, Black *et al*. (1996) used random sequence mutagenesis to create HSV-*tk* mutants, which significantly increased the sensitivity of transduced cells to acyclovir and ganciclovir. One of these mutants, the HSV-*tk* 30 (Qiao *et al*., 2000), has been tested in various human cancer cell lines, and renders these cells 9–500-fold more sensitive to ganciclovir than the wild-type (wt) HSV-*tk*, and appears to be a promising candidate for evaluation in gene therapy trials.

TUMOUR-SUPPRESSOR GENE THERAPY

Most of the known mutations in cancer lead to either gain of function of transforming oncogenes or loss of function of tumour-suppressor genes. Despite the presence of multiple genetic mutations in cancer cells, correction of a single critical lesion is often sufficient to abrogate the malignant phenotype. The *p53* tumour-suppressor gene is an attractive target for gene therapy replacement strategies, because of the central role it plays in determining the cellular response to genomic stress and DNA damaging agents. Depending upon the degree of DNA damage, p53 can either mediate p21-induced cell cycle arrest at the $G_1 \rightarrow S$ interphase, or induce apoptosis through its influence on genes such as *bax* and *fas*. p53 also may play an anti-angiogenic role through its regulatory effects on VEG-F and thrombospondin.

Multiple models (both *in vitro* and *in vivo*) validate the rationale for p53 gene therapy in a broad spectrum of tumour types, either alone or in combination with radiation or chemotherapeutic agents. Most investigators currently favour replication-defective (E1 deleted) adenoviruses because they do not integrate into the host chromosome, can infect a wide range of tissue types and can be generated in high titres. Although most strategies have focused on cell lines or established *in vivo* tumours with a mutant p53 status, some groups have demonstrated efficacy even in the presence of tumours with a wt p53 status, when used in conjunction with radiation therapy.

Kock *et al*. (1996) used adenoviral-mediated *p53* gene transfer (Ad *p53*) into p53 mutant glioma cell lines and demonstrated dose-dependent inhibition of proliferation (by [^{3}H]thymidine incorporation assays), accompanied by significant DNA fragmentation (apoptosis) and induction of p21. Similarly, p53 mutant glioma xenografts were established in the flanks of nude mice, and a significant reduction in tumour volume, retardation of tumour progression and improved survival were seen in mice treated intratumourally with Ad *p53* compared with a control vector. Similar results were observed using stereotactic injections of Ad *p53* in an intracerebral orthotopic rat glioma model. Ad *p53* did not, however, exert the same inhibitory effect on U87 glioma xenografts (characterized by a wt *p53* status). However, when used in combination with radiation, there was significant reduction in tumour volume and improved survival compared with mice treated with Ad *p53* or radiation alone; this radiosensitizing effect was associated with increased apoptotic cell death determined by the TUNEL assay (Lang *et al*., 1998).

Nielsen *et al*. (1998) demonstrated significant synergy between Ad *p53* gene therapy and paclitaxel (which stabilizes microtubules and mediates a $G_2 \rightarrow M$ arrest) in xenograft models of ovarian cancer ($p53^{null}$SK-OV-3), prostate cancer ($p53^{mut}$DU-145) and breast cancer ($p53^{mut}$MDA-MB-231). In all these models, a significant reduction in tumour burden (60–90%) was seen in animals treated with the combined modality as opposed to Ad *p53* or paclitaxel alone.

The combination of Ad *p53* gene therapy with immunotherapy based on Ad IL-2 has revealed impressive results in the murine mammary adenocarcinoma (PyMT) model with regression of 65% of established tumours (Putzer *et al*., 1998). Upon rechallenge with tumour cells, 50% of mice demonstrated long-term tumour-free survival, suggesting the development of antitumour immunity. Injection of mice with either vector alone resulted only in a delay in tumour growth, and no evidence of an antitumour cytolytic T lymphocyte response, as was evident in mice treated with both vectors. (See also chapter on *Models for Immunotherapy and Cancer Vaccines*.)

An inherent limitation of current gene transfer methods is the lack of regulation of gene expression following transfer. Some groups have recently described the use of trans-splicing ribozymes simultaneously to repair mutant *p53* transcripts and restore translation of wt p53 protein in the Saos-2 osteosarcoma and the SW480 colon carcinoma models (Watanabe and Sullenger, 2000). (See also chapter on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Cells*.)

Based on the above-reviewed data, and a plethora of other preclinical data, *p53* gene therapy is currently being investigated in the setting of a number of phase I and phase

II trials in non-small cell lung and squamous cell head and neck cancers. Although most of these trials are focused on safety and dissemination issues, and the efficacy of gene transfer into tumour tissue, some are noteworthy for responses in a subset of treated patients. Ultimately, this therapy will succeed if there are data to suggest a better response rate, potentiation of other agents and selectivity of efficacy in cancer cells. (See also chapter on *Regulation of the Cell Cycle*.)

ONCOLYTIC VIRUSES

Enhanced understanding of the biological processes involved in viral life cycles and virus–host cell interactions has led to novel strategies for gene therapy. Significant achievements have been attained in the development of both nonviral agents and replication-incompetent viral vectors for gene replacement, drug-resistance gene delivery and prodrug-activating enzymes. However, limitations in gene expression levels, viral titre and target cell specificity have limited the efficacy of these agents in the setting of cancer gene therapy. Novel attempts to overcome limitations in cancer gene therapy involve the use of natural or recombinant viruses that selectively replicate in human tumour cells. The ability of these viruses to replicate in cancer cells leads to increased gene expression via transcription from nascent viral genomes and subsequent infection of neighbouring cells. Although virus replication can directly lyse cells as a result of the viral load exceeding the cellular capacity, induction of inflammatory cytokines, interferon response pathways, CTL activation and accumulation of cytotoxic proteins can enhance their oncolytic activity. Currently, the most promising oncolytic viral agents include reovirus (Type 3 Dearing), and recombinant versions of vaccinia viruses, adenoviruses and herpes simplex viruses (HSV-1). These viruses are able to replicate selectively in cancer cells as a result of constitutively active signalling cascades, inactivated tumour-suppressor gene functions, or elevated pools of polymerase substrates in highly proliferative cells.

Strategies employing the use of adenoviruses as tumour-selective oncolytic agents have targeted mutations in the early region 1A or early region 1B genes (E1A and E1B) located in the E1 segment of the adenovirus genome. E1A proteins interact with pRB, releasing the E2F transcription factors needed for expression of genes involved in entry into S phase and transactivation of the viral E2 promoter. E1A proteins have also been shown to associate with p300/CBP to promote E2F-mediated transcription. The E1B gene products E1B55K and E1B19K act to suppress apoptotic pathways in the infected cell. E1B55K inhibits p53-mediated cell cycle arrest or apoptosis, and serves as a transporter of viral mRNAs from the nucleus to the cytoplasm, while the 19-kDa E1B gene product mimics the antiapoptotic activity of Bcl-2. Since almost half of all human cancers are defective in p53 tumour-suppressor function, Bischoff *et al*. (1996) used an adenovirus with an E1B55K gene deletion (*dl1520*), subsequently designated ONYX-015, to target these malignancies. A screen of several human cell lines demonstrated that the mutant virus elicited its cytotoxic effects only to those cells lacking functional p53. In addition, the ONYX-015 virus caused significant size reduction and complete regression in a majority of human cervical carcinoma ($p53^{-/-}$) xenografts in nude mice, while having no effect on the tumour volume of human glioblastoma multiforme ($p53^{+/+}$) xenografts. The specificity of the ONYX-015 virus for p53 null cells has been questioned, but the efficacy of this virus as an oncolytic agent has now been demonstrated clinically. In addition, treatment of squamous cell carcinomas of the head and neck with combined intratumoural ONYX-015 injection, cisplatin and 5FU have improved response rates in phase II clinical trials (Khuri *et al*., 2000). Additional adenovirus mutants have recently been constructed that lack E1B55K and deliver prodrug-converting enzymes, or target pRB-deficient tumour cells.

HSV-1 is a double-stranded DNA virus and, typical of all herpes viruses, has the capacity for both lytic and latent life cycles. HSV-1 latency is established in neurons, but these viruses can be reactivated to cause encephalitis and lesions at the primary site of infection. The first replication-competent HSV-1 mutant (*dlsptk*) consisted of a 360-bp deletion within the thymidine kinase (*tk*) gene (UL23) (Martuza *et al*., 1991). The *tk* deletion makes virus replication dependent on elevated thymidine triphosphate pools, which are found in highly proliferative cells. The dlspk HSV-1 mutant was shown to inhibit tumour growth and prolong survival in nude mice containing subcutaneous and subrenal xenografts of U87 human gliomas, but HSV-1 *tk* mutants are not devoid of the capacity for neurovirulence and removal of *tk* eliminates the ability to treat with antiviral agents, such as ganciclovir. Other HSV-1 viruses have been safety-modified by deleting both copies of the *γ34.5* gene (*RL1*) that is responsible for neurovirulence (Chou *et al*., 1990). Mineta *et al*. (1994) created an HSV mutant containing a *lacZ* insertion in place of the ribonucleotide reductase gene (*UL39*), making virus replication depend on cellular proliferation, allowing reporter gene assessment, and *tk* to be maintained for sensitivity to antiviral agents. The same group subsequently generated an HSV-1 mutant (G207) that included both *RL1* deletions and the *LacZ* insertion into UL39 to combine the advantages of the single mutants. The G207 virus was also shown to reduce U87 human glioma xenograft growth and prolong survival of mice bearing these tumours. Toxicity studies in HSV-sensitive owl monkeys have further demonstrated the lack of neurovirulence of the G207 virus (Hunter *et al*., 1999). Recently, the G207 virus and an HSV-1 virus (1716) lacking both copies of *RL1* have been evaluated for dose-escalation

safety trials in patients with incurable malignant brain tumours.

Reovirus is being investigated as an oncolytic agent as a result of its selective replication in cells with activated Ras signalling pathways. This selectivity is due to the fact that activated Ras signalling leads to dephosphorylation and inactivation of the double-stranded RNA-activated protein kinase (PKR) (Strong *et al*., 1998). PKR serves as an intracellular antiviral agent by its ability to detect double-stranded RNA, unique to specific viral genomes and transcripts. Activated PKR phosphorylates eIF-2 alpha, a component of the cell translation machinery, preventing its ability to initiate translation on viral transcripts. Ras-activated signalling leads to inactivation of PKR, allowing primary viral products to accumulate and carry out the lytic life cycle of the virus, but the mechanism by which Ras inactivates PKR is unknown. To date, reovirus has been found to be capable of infecting a wide range of human tumour cell lines (Norman and Lee, 2000) and induce regression of both SCID tumour xenografts, and tumours in immunocompetent mice (Coffey *et al*., 1998). The large proportion of human tumours having constitutively active signalling cascades upstream of Ras, or activating mutations in Ras itself, exemplifies the potential of reovirus as an oncolytic agent. Furthermore, the wide distribution of the virus receptor (sialic acid) over a variety of cell types and the nonpathogenic nature of reovirus infections makes it a promising therapeutic candidate.

Vaccinia virus belongs to the Poxviridae family of complex DNA viruses that encode all replication factors necessary for their cytoplasmic life cycle. The ability of the virus to tolerate large insertions of foreign DNA ($\leq$25 kb) and infect a wide variety of cell types has led to the development of vaccinia expression vectors. As with HSV-1 *tk* mutants, vaccinia virus *tk* mutants require elevated thymidine triphosphate pools for replication and are less pathogenic than the wild-type virus. Puhlmann *et al*. (2000) demonstrated the enhanced ability of *tk*-deleted vaccinia viruses to selectively replicate in tumour cells and inserted a luciferase reporter that allowed them to detect significant expression levels in tumour cells compared to normal tissue. Recently, prodrug-activating enzymes, including purine nucleoside phosphorylase and cytosine deaminase, have been engineered into the *tk*-deleted viruses to improve tumour cell-specific killing and to provide a mechanism to inhibit viral replication in pathogenic situations. Both of these recombinant viruses demonstrated significant oncolytic activity and enhanced tumour regression following addition of prodrug but the efficiency of tumour cell-killing mediated by the *tk*-deleted virus replication required significant decreases in the amount of virus used in order to see the effect of prodrug conversion. Although these vaccinia agents have been used to demonstrate impressive oncolytic activity, the extent of viral pathogenicity will need to be reduced before considering these viruses for clinical use.

Table 1 Cancer-directed gene therapy strategies

Suicide gene delivery
HSV-tk: Converts prodrugs, such as acyclovir and ganciclovir, into nucleoside analogues that inhibit DNA replication
Cytosine deaminase: Converts 5-fluorocytosine into the cytotoxic compound 5-fluorouracil

Tumour-suppressor gene delivery
p53: Mediates cell cycle arrest and/or apoptosis in p53-deficient tumour cells
pRB: Binds E2F family of $G_1 \rightarrow S$ entry transcription factors
p21: Prevents cyclin D/cyclin-dependent-kinase mediated phosphorylation (inactivation) of pRB

Oncolytic viruses
Adenovirus: E1B55K-deletion mutants selectively replicate/kill p53-deficient malignancies
HSV-1: Thymidine kinase (UL23) and ribonucleotide reductase (UL39) gene deletion mutants require elevated nucleotide pools, found in highly proliferative cells, for viral replication. Deletion of γ34.5 (RL1) genes prevents neurovirulence
Reovirus: Malignancies with constitutively active Ras signalling pathways have inactive double-stranded RNA-activated protein kinase (PKR), allowing eIF-2 alpha to initiate translation of viral proteins
Vaccinia virus: Thymidine kinase gene deletion mutants require elevated nucleotide pools for viral replication (see HSV-1)

Viruses with tumour-selective replication hold great promise as oncolytic agents; however, further understanding of host–virus interactions and the aberrant pathways associated with specific malignancies will be essential for improving upon the specificity and efficacy of these agents. Future improvements will undoubtedly include the use of viruses that target tumour-specific receptors and/or contain tumour-specific promoters. (See also chapters on *Human DNA Tumour Viruses* and *RNA Viruses*.)

To summarize the above sections, cancer-directed gene therapy strategies are outlined in **Table 1**.

ANTIANGIOGENIC GENE THERAPY

Survival and growth of a tumour in living organisms are dependent on two factors, its ability to escape immune surveillance and its capacity to maintain a continuous supply of oxygen and nutrients. Angiogenesis, the process in which new blood vessels grow from existing vessels, must occur within solid tumours in order for tumour cells to acquire the oxygen and nutrients that they need for proliferation. Tumour growth may be limited to 1–2 mm^2 by the diffusion of nutrients without recruitment of new blood vessels. In addition, neovascularization of a solid tumour appears to be required before formation of metastasis can occur. Angiogenesis is a multistep process which

includes recruitment and activation of nearby endothelial cells, degradation of the vascular basement membrane, proliferation and migration toward the angiogenic stimulus and formation of new capillaries by linkage to the preexisting capillary network. Either tumour cells or accessory cells within the tumour stroma can trigger angiogenesis by the secretion of angiogenic factors such as basic fibroblast growth factor (bFGF) and vascular endothelial growth factor (VEGF). These factors, along with platelet-derived growth factor (PDGF), transforming growth factor beta (TGF-β) and angiopoietins, are also involved in maturation and stabilization of newly formed vessels. The angiogenic switch during tumorigenesis appears to be caused by the disruption of a complex balance between angiogenesis stimulators and inhibitors, both of which fluctuate over the course of tumour development. (See also chapters on *Angiogenesis* and *Angiogenesis Models*.)

Delivery of antiangiogenic agents to tumours by means of gene therapy has become an attractive new approach since it has been difficult to mass produce stable recombinant forms of endogenous angiogenesis inhibitors; for a detailed review, see Feldman and Libutti (2000). Antiangiogenic gene therapy of cancer is currently designed either to target the tumour directly, by increasing local concentrations of antiangiogenic agents within the tumour, or by a systemic delivery approach whereby the patient's normal tissues would function as a 'factory' to increase circulating levels of a particular agent. While the tumour-directed gene therapy approach avoids systemic toxicity of the antiangiogenic gene product, its effectiveness may be limited by blood supply within the tumour for efficient distribution of the vector. Furthermore, systemic gene delivery seems to be an attractive method for delivery of antiangiogenic agents owing to their apparent lack of toxicity.

Antiangiogenesis gene therapy strategies to date include the use of endogenous angiogenesis inhibitors, or those that target the production or action of endogenous proangiogenic agents. VEGF is a known tumour-derived cytokine that promotes angiogenesis. Hence much attention has been focused on inhibiting transcription or translation of VEGF and its receptors through the use of neutralizing antibodies. Vitaliti *et al*. (2000) demonstrated that single-chain antibody fragments directed against VEGF effectively blocked angiogenesis in the chorioallantoic membrane of chick embryos. In addition, these antibody fragments were effective in reducing the growth of subcutaneous tumours in nude mice.

The majority of antiangiogenic therapies have focused on rapidly growing, highly vascularized tumours or tumour cell lines. Interestingly, angiostatin, a fragment of the plasminogen protein, has been shown to inhibit significantly the growth of tumours in mice which were derived from both slow and rapidly growing human bladder carcinoma cell lines (Beecken *et al*., 2001). Thus, the responses of slowly growing, poorly vascularized and rapidly growing, highly vascularized tumours to this angiogenesis inhibitor appear to be similar.

Endostatin, a protein fragment derived from the C-terminal noncollagenous domain of the basement membrane constituent Type XVIII collagen, has been shown to inhibit tumour growth *in vivo*. The potential of the systemic gene therapy approach was evaluated in murine VEGF implant angiogenesis and orthotopically implanted human colon/liver metastasis xenograft models after adenovirus-mediated gene delivery of endostatin (Chen *et al.,* 2000). In the VEGF implant model, mice were treated with the adenoviral construct by tail-vein injection, and 13 days later Teflon chambers containing $VEGF_{165}$ were implanted subcutaneously. Angiogenesis was then measured by analysis of vascularized tissue that formed around the implants. Mice in the human colon/liver xenograft model were also pretreated with the adenovirus, and colon tumour fragments were orthotopically transplanted 13 days after treatment. Circulating endostatin levels increased and lasted longer in both models in response to treatment. Furthermore, partial inhibition of VEGF-induced angiogenesis was shown in the VEGF implant angiogenesis model, and complete inhibition of tumour growth was achieved in 25% of mice in the human colon/liver metastasis xenograft murine model.

In order to augment the inhibitory effects of antiangiogenic agents on tumour growth, a new combination therapy approach has emerged using simultaneous treatment with two different agents. In a murine leukaemia model, Scappaticci *et al*. (2001) demonstrated complete loss of tumorigenicity in 40% of animals that received tumours retrovirally transduced with a combination of both angiostatin and endostatin.

Antitumour effects have also been observed using specific interferons, endogenous glycoproteins that participate in triggering cellular antiviral mechanisms. Retroviral delivery of murine class I interferons (IFNs) -α and -β to endothelial-like cells *in vitro* significantly inhibited migration and invasion in chemotaxis and chemoinvasion Matrigel assays, respectively (Albini *et al*., 2000). In addition, a Matrigel morphogenesis assay showed that cells transduced with IFN-β were completely inhibited from growth and formation of capillary-like structures. However, IFN-α was unable to inhibit the endothelial cells from growing and forming capillary-like structures, suggesting that whereas IFN-β seems to exert an effect on endothelial cell differentiation, IFN-α does not. These IFNs were also tested in both immune-competent and athymic nude mice using the Matrigel sponge model. In this model, cells expressing IFN-β strongly inhibited angiogenesis whereas partial inhibition was observed with cells expressing IFN-α.

Antiangiogenic gene therapy has emerged as an alternative strategy to deliver antiangiogenic agents to cancer patients. The optimal method of gene delivery is still undetermined, as the immunogenicity and toxicity of many

viral vectors used for *in vivo* studies remains a concern regarding their clinical use. Use of viruses as gene therapy agents appears to be efficient in terms of cell entry and gene expression, whereas nonviral gene delivery using transporters such as liposomes offers lower risk of toxicity and immunogenicity, but may only offer suboptimal transgene expression. Molecular mechanisms of angiogenesis seem to vary among different tumour types, thus the efficacy of specific antiangiogenic agents is expected to vary in accordance with the tumour type. Systemic gene therapy may prove to be a method that can be used to alleviate the need for chronic drug treatment, which may be necessary with the tumour-directed approach. Further progress will require a more complete understanding of the mechanisms involved in the angiogenesis process and also improvements in gene vector technology.

CYTOKINE GENE THERAPY

Cytokine gene therapy of cancer is rooted in host immune defences, especially host T cell-mediated immunity involving cytotoxic T lymphocytes (CTLs) and T-helper cells (Th1s) (Lattime and Stanton, 1999). Tumours can be, and often are, recognized by T cells as abnormal. However, this recognition rarely leads to a fully activated immune response because the Th1 cells most often fail to secrete the cytokines necessary for sustaining proliferation and activation of the CTL cells. When a fully activated CTL response is initiated, the results can be dramatic and often correspond to cases of spontaneous remission or extended latency time to development of metastases after primary tumour removal. Efforts to enhance the CTL response by systemic administration of cytokines have been unsuccessful largely owing to serious side effects, rapid degradation and/or rapid elimination of the cytokine from the circulation, and in older patients with an immune system that is generally less responsive to cytokine stimulation (Hsueh *et al*., 1996). Cytokine gene therapy for cancer seeks to circumvent the systemic dosing problems by delivering the appropriate cytokine stimulation directly to the tumour site and thus support, elicit and enhance a full host CTL immune response in a more natural paracrine milieu. (See also chapters on *Models for Immunotherapy and Cancer Vaccines*; *Signalling by Cytokines*.)

Several gene therapy approaches are being pursued in order to deliver the appropriate cytokine to the tumour site and/or the most immunologically relevant site for induction and enhancement of antitumour immunity (Fernandez *et al*., 1998). Studies have shown that, under various circumstances, the list of relevant cytokines that promote an antitumour immune response include TNF-α, G-CSF, GM-CSF, IFN-γ, IL-1, IL-2, IL-4, IL-6, IL-7, IL-12 and IL-18 while TGF-β actually appears to suppress antitumour immune responses. Of these cytokines, IL-2 and IL-12 have perhaps shown the most promise owing to their specific mode of action via T cell-mediated immunity, but they also have significant difficulties. Cytokines that primarily stimulate immunity via T cells include GM-CSF, IL-6, IL-7 and IL-18. Eosinophils are a major effector cell for IL-4 and IL-2 while neutrophils are involved in G-CSF mode of action. In addition, G-CSF, IFN-γ, TNF-α, IL-2, IL-4 and IL-7 significantly enhance antitumour macrophage activity. Dendritic cell activity is also significantly enhanced, both at the tumour site and in surrounding tissue, by G-CSF, GM-CSF, IL-2, IL-18 and especially IL-4. Endothelial cells may also be involved as an effector cell for most of these cytokines since several are known to induce vascular adhesion molecule expression profiles in such a way as to promote inflammatory cell invasion at the site of cytokine expression. Obviously, with the range of mechanisms represented by the major effector cell types upon which these cytokines act, finding the optimal cytokine(s) for enhancing and stimulating antitumour immunity against any particular tumour will be critical to the underlying success or failure of these approaches.

Finding the right delivery method is also critical. The most common method of delivering the cytokine gene is to deliver cells that have been engineered *ex vivo*, via transfection or transduction, to produce the cytokine of interest. Live attenuated virus has also been used in order to infect a local tissue indiscriminately. Tumour cells, lymphocytes, endothelial cells and fibroblasts have all been tried as specific cellular targets. In the case of tumour cells, both engineered primary tumour cells from patients and engineered allogeneic, established tumour cell lines have been used as inoculums. There are many advantages and disadvantages to either tumour cell source, but generally, the use of autologous primary tumour cells guarantees the presentation of host tumour antigens as part of the antitumour response while established tumour cell lines provide a simpler, less labour-intensive method of delivering the cytokine but has the disadvantage of reducing the chances of a shared antigen profile. Lymphocytes are also an attractive target for gene therapy using cytokines because of their natural functions – proliferation and activation in response to immunogenic stimulation and accumulation at the site of immunogenicity. Unlike lymphocytes, endothelial cells are a potentially renewable source from which engineered endothelial cells could specifically expand into the growing vasculature of a solid tumour. In fact, endothelial cells have been shown to tolerate genetic engineering, migrate into and survive in an angiogenic site after intravenous injection (Ojeifo *et al*., 1995). Still, primary tumour cells, lymphocytes and endothelial cells are difficult to isolate, expand and manipulate *in vitro*. Fortunately, dermal fibroblasts are a readily available source of autologous cells that can be isolated and grown *in vitro* with relative ease, thus making it simpler to retrovirally transduce and select them. This means that fibroblasts are superior to other cells from an

ex vivo manipulation standpoint – whether they are superior, compared with the other possible target cell types, in their ability to deliver and stimulate antitumour immune responses when expressing the relevant cytokine remains to be seen. Still, there is clearly preclinical and clinical evidence that delivering immunomodulating cytokines in this manner enhances host antitumour activity resulting in significant suppression of tumour growth, although to date none have shown complete remission and/ or protection.

Since normal immune responses involve a highly coordinated effort involving numerous cell types and a veritable orchestra of cytokines, merely overexpressing a particular cytokine, or even two, is unlikely to provide the most effective antitumour response. With this in mind, dendritic cells (DCs) are being pursued as yet another target for cytokine gene therapy. Because of their role in directly presenting tumour antigens to and activating T cells, DCs are absolutely critical to a robust cell-mediated immune response. Many of the difficulties in obtaining and growing DCs have been overcome and numerous reports have outlined the successful transduction and maintenance of DC function after *ex vivo* manipulation (Szabolcs *et al*., 1997). The display of tumour antigen by DCs coupled with enhanced cytokine stimulation provided by gene therapy holds significant promise for increasing the already encouraging results obtained by cytokine gene therapy with other targets.

GENE-MODIFIED DENDRITIC CELLS

There have been a number of gene therapy approaches to fighting cancer by genetically modifying immune system cells to enhance their antitumour effect. One approach utilizes modification of CTLs for adoptive T cell therapy, and the other utilizes dendritic cells to enhance either presentation of tumour antigens or provide stronger costimulatory signals to activate antigen-specific CTLs more efficiently. Dendritic cell-based gene therapy has shown progress and several animal studies have provided encouraging results.

Dendritic cells are targets for cancer gene therapy because they are potent antigen presenters and possess high levels of various costimulatory molecules important for activation of $CD4^+$ and $CD8^+$ T lymphocytes. The cytotoxic T lymphocyte response is critically important for tumour rejection and eradication, and means to enhance this response through modification of dendritic cells are being vigorously pursued. Gene transfer of a tumour-specific antigen to dendritic cells *ex vivo* and infusion into the patient have the advantage of providing a renewable supply of antigen for presentation and preclude the need to identify specific peptides within the protein that induce strong CTL responses. Adenoviral vectors are commonly used for this purpose, although other viral and nonviral vectors have been used. (See also chapter on *Models for Immunotherapy and Cancer Vaccines*.)

In a study by Kaplan *et al*. (1999), a murine B16 melanoma tumour model was used to test the ability of modified dendritic cells to induce an antitumour immune response. B16 cells express murine homologues of human melanoma-associated antigens such as gp100, TRP-1, TRP-2 and MART-1. Therefore, murine dendritic cells isolated from the bone marrow were transduced with gp100 or TRP-2 using an adenoviral vector and infused into animals before or after tumour formation. The pre-immunization model showed substantial protection from tumour-induced death after infusion of 5×10^5 gp100 or TRP-2 expressing dendritic cells and challenge with a lethal dose of B16 tumour cells at 15 days post-immunization. Within 30 days, 100% of control mice died whereas in mice immunized with dendritic cells expressing gp100 or TRP-2, only 20% and 0% developed tumours with delayed kinetics, respectively. The effect was long lasting as 75% of mice receiving a second injection of B16 cells 50 days after the first B16 injection remained tumour free. Although the results with preimmunization were impressive, a more clinically relevant model involves treatment of pre-existing tumours. To this end, the authors injected a lethal dose of B16 cells into mice 4 days before treatment with transduced dendritic cells. Surprisingly, only 20% of the mice given gp100 expressing dendritic cells and 60% of mice given TRP-2 expressing dendritic cells remained tumour free. The outcome was improved by infusion of a mixture of the two dendritic cell populations.

Similar results were observed by Wan *et al*. (1999), who used the same murine tumour model. Immunization with dendritic cells transduced by adenovirus with gp100 could prevent tumours in 70% of mice challenged with B16 cells. In addition, treatment after establishment of pulmonary tumours resulted in an 80% reduction in lung metastases.

Less impressive results were obtained using non-immunogenic mouse fibrosarcoma cell (NFSA) tumours stably expressing MART-1, a melanocyte lineage-specific antigen expressed by many malignant melanomas (Ribas *et al*., 1999). When these mice were challenged with tumour cells after preimmunization twice with dendritic cells transduced with MART-1, only 21% were completely protected. Treatment of established tumours resulted in only a small delay in tumour growth. These results may be due to the poor immunogenicity of MART-1. A strong and sustained antitumour response may ultimately require transfer of more than one tumour antigen gene to dendritic cells to confer a more antigenically diverse immune response and to guard against poorly immunogenic antigens and development of tumour resistance due to down-regulation of specific tumour antigens.

$CD4^+$ T cells appear to be important for development of antitumour immunity. Depletion of $CD4^+$ T cells from animals before immunization with modified dendritic cells

significantly abrogates the antitumour response (Kaplan *et al*., 1999; Ribas *et al*., 1999; Wan *et al*., 1999). The central effecter role of $CD4^+$ T cells was demonstrated by Wan *et al*. (1999). They found that protective immunity conferred by transfer of gene modified dendritic cells could be maintained after $CD8^+$ T cell depletion. Furthermore, tumour rejection was independent of signals from IL-12 and CD40. These data may reflect a shift in our understanding of the antitumour response due to $CD8^+$ and $CD4^+$ cells and factors which influence generation of strong immunity to weakly immunogenic tumour antigens.

The second approach to cancer gene therapy utilizing dendritic cells involves transfer of the gene for CD40 ligand (CD40L). CD40L is usually expressed on activated $CD4^+$ T cells, whereas its receptor, CD40, is expressed predominantly on B cells, activated macrophages and dendritic cells. Stimulation of dendritic cells through CD40/CD40L is important for activation of dendritic cells, which then in turn up-regulate various T cell costimulatory molecules such as ICAM-1 and B7. The eventual result is stimulation of $CD8^+$ CTLs.

Enhancing the activation of dendritic cells is hypothesized to increase the CTL immune response against tumours and forms the basis for gene transfer of CD40L into dendritic cells. It is thought that CD40L expressed on dendritic cells will interact with CD40 in a *cis* fashion, to cause autoactivation and more potent activation of CTLs. This hypothesis was recently tested by Kikuchi *et al*. (2000). Using B16 and CT26 (undifferentiated colon adenocarcinoma) tumour models, dendritic cells transduced with CD40L were administered to mice bearing established tumours. For both models, significant and sustained tumour growth delay was observed, as was a significant increase in survival. Additionally, a two-tumour model in which tumours were established in each flank was tested for antitumour CTL activity. When CD40L transduced dendritic cells were injected into one of the tumours, both the left and right tumours were growth suppressed. Immunohistochemical analysis of the tumours showed high-level $CD8^+$ T cell infiltration and some $CD4^+$ T cells, whereas control mouse tumours had very low levels of T cell infiltration.

There are a number of advantages to the CD40L system compared with transduction of tumour antigens to dendritic cells. The most obvious is that in the CD40L system, a tumour antigen does not have to be identified. Furthermore, modified dendritic cells have the potential to present many tumour antigens owing to their interaction with tumour cells, thus allowing the immune system to respond to a variety of tumour antigens.

DRUG RESISTANCE GENE THERAPY

Transfer of drug resistance genes into haematopoietic stem cells is used as an approach to circumvent chemotherapy induced myelosuppression. The major goal of drug resistance cancer gene therapy is selective survival of haematopoietic progenitors while concurrently inhibiting tumour growth. This strategy results in the selection of drug-resistant cells and thus may allow more dose-intensive chemotherapy regimens to treat drug-resistant tumours. Bone marrow stem cells are the ideal target for this approach based on their multilineage differentiation potential and their capacity for self-renewal. The majority of the gene transfer studies have used recombinant murine leukaemia retrovirus-based vectors. However, this class of retrovirus requires a breakdown of the nuclear membrane for integration, limiting the transduction efficiency of quiescent stem cells. More recently, HIV-based lentiviral vectors have been shown to transduce nondividing cells stably, increasing the potential for gene transfer into stem cells. Despite the difficulties in transducing stem cells, studies have shown that transduction of only a few early progenitors with a drug resistance gene is adequate for selection and repopulation of the marrow.

Gene transfer of the multidrug resistance gene-1 (*MDR-1*), O^6-methylguanine–DNA methyltransferase (*MGMT*) and dihydrofolate reductase (*DHFR*) have each demonstrated the ability to increase cellular resistance to chemotherapeutic agents (see chapter on *Models for Drug Development and Drug Resistance*). The most widely studied drug resistance gene is *MDR-1* that encodes the 170-kDa ATP-dependent multipass-transmembrane efflux pump P-glycoprotein which can expel a number of lipophilic compounds from the cell. MDR-1 provides resistance to a variety of drugs including the anthracyclines, *Vinca* alkaloids, actinomycin D and paclitaxel. Increased drug resistance following *MDR-1* gene transfer into human and murine haematopoietic cells has been documented in many studies and is the basis for clinical trials that are now under way. Initial murine studies demonstrated that human *MDR-1* transduced haematopoietic cells had preferential survival *in vivo* after treatment with MDR drugs. Expression of human MDR-1 in long-term repopulating stem cells was shown by serial transplantation of transduced marrow with paclitaxel selection and resulted in MDR-1 expression 17 months after the initial transduction. This demonstrated that continuous expression can be maintained with *in vivo* selection. Simultaneous resistance of *MDR-1* transduced haematopoietic cells *in vivo* and sensitization of murine tumour cells was shown in Balb/C recipients following treatement with paclitaxel. The mice tolerated high doses of paclitaxel while demonstrating a significant decrease in tumour growth (Hanania and Deisseroth, 1997).

The O^6-alkylguanine–DNA alkyltransferase enzyme, encoded by the *MGMT* gene, increases cellular resistance to alkylating agents including the chlorethylating agent BCNU and the methylating agent temozolomide. Mutant forms of MGMT are resistant to the alkyltransferase modulator O^6-benzylguanine (BG) and provide a greater

resistance to the chemotherapeutic agents compared with wild-type forms. Two well-characterized mutants, G156A and P140K, have been shown to protect transduced marrow from BG and BCNU or temozolomide-induced myelosuppression and result in significantly improved survivial *in vivo* (Ragg *et al*., 2000; Sawai *et al*., 2001). Following drug treatment, an increase in MGMT transduced cells reached nearly 100% in bone marrow (Ragg *et al*., 2000) and peripheral blood (Ragg *et al*., 2000; Sawai *et al*., 2001) and secondary transplant experiments proved that selection occurred at the stem cell level (Ragg *et al*., 2000; Sawai *et al*., 2001). MGMT-associated enrichment was best demonstrated in a model in which G156A transduced marrow was selected up to 1000-fold, even when initially present at a repopulation disadvantage (Davis *et al*., 2000). In a tumour xenograft study, G156A transduced marrow was transplanted into *nu/nu* athymic mice bearing a tumour cell line expressing high levels of AGT. Here, mice were protected from a dose-intensive regimen of BG and BCNU while the tumour exhibited a significant growth delay (Koc *et al*., 1999). These results provide strong evidence that MGMT transduced marrow may increase the therapeutic index of BG and BCNU to allow aggressive cancer treatment. Clinical trials using G156A *MGMT* gene transfer have also been proposed. In contrast to MDR substrate drugs, BCNU is more toxic to early haematopoietic progenitors and therefore may facilitate stronger selection for MGMT transduced cells.

Table 2 Host-directed cancer gene therapy strategies

Antiangiogenic gene therapy

Angiostatin: Proteolytic fragment of plasminogen protein that serves as an endogenous inhibitor of angiogenesis

Endostatin: Proteolytic fragment derived from the C-terminus of collagen XVIII that serves as an endogenous inhibitor of angiogenesis

Localized cytokine gene therapy (enhancing cell-mediated immunity)

Cytokines that promote an antitumour immune response: TNF-α, G-CSF, GM-CSF, IFN-γ, IL-1, IL-2, IL-4, IL-6, IL-7, IL-12, IL-18

Effector cell targets of cytokine gene delivery: Tumour cells, lymphocytes, endothelial cells, fibroblasts

Gene-modified dendritic cells (enhancing cell-mediated immunity)

Tumour-specific antigens displayed on dendritic cells: gp100, TRP-1, TRP-2, MART-1

CD40L gene delivery to dendritic cells: CD40L expression on dendritic cells serves to activate cytotoxic T cells, allowing multiple antigen recognition without prior knowledge of the antigens involved

Drug resistance gene therapy (haematopoietic protection/ enhancing the chemotherapeutic index)

MDR-1: Efflux pump providing resistance to anthracyclines, *Vinca* alkaloids, actinomycin D and paclitaxel

MGMT: Encodes O^6-alkylguanine–DNA alkytransferase protein: resistance to cytotoxic alkyl lesions arising from BCNU and temozolomide treatment. Mutant forms of MGMT are resistant to the endogenous MGMT inhibitor, O^6-benzylguanine (BG)

DHFR: Dihydrofolate reductase is required for thymidine synthesis. Mutant forms of DHFR are resistant to inhibitors of the endogenous enzyme (methotrexate, trimetrexate)

Gene transfer of the *DHFR* gene is used to protect haematopoietic cells from antifolate chemotherapeutic agents such as methotrexate (MTX) and trimetrexate (TMTX), which bind DHFR and inhibit nucleotide synthesis. Certain mutant forms exhibit reduced affinity for antifolates, and therefore gene transfer of these variants is used to provide increased cellular protection from these agents. Mice transplanted with haematopoietic cells retrovirally expressing the human DHFR mutants L22R and L22Y were protected from MTX- and TMTX-induced cytopenia, respectively. Recently, gene transfer of a nucleoside transporter, hENT2, which is insensitive to the nucleoside analogue NBMPR, was shown to confer resistance in murine cells to antifolate drugs which target other key enzymes involved in purine/pyrimidine synthesis. In this model, tumour cells that express high levels of a sensitive nucleoside transporter are sensitized, whereas haematopoietic cells are protected, thereby increasing the antitumour activity of folate analogues (Patel *et al*., 2000).

These murine models indicate that drug-resistant gene transfer can lead to enrichment of drug-resistant bone marrow, which may allow more intensive chemotherapy regimens to treat tumours resistant to conventional chemotherapy. However, the success of this strategy will ultimately be determined in clinical trials which are ongoing. Of note, this powerful selection strategy can also be used to co-select for a second therapeutic gene in the treatment of genetic disorders of the haematopoietic system and studies to investigate this possibility are also in progress.

Finally, to summarize the above sections, host-directed gene therapy strategies are outlined in **Table 2**. (See also chapters on *Antisense and Ribozyme Therapy*; *Gene Therapy–Tumour Suppressor Replacement/Oncogene Suppression*.)

ACKNOWLEDGEMENT

This work was supported by Public Health Service Grants RO1CA84578, RO1ES06288, UO1CA75525 and P30CA43703.

REFERENCES

Albini, A., *et al*. (2000). Inhibition of angiogenesis and vascular tumor growth by interferon-producing cells: a gene therapy approach. *American Journal of Pathology*, **156**, 1381–1393.

Beecken, W. D., *et al*. (2001). Effect of antiangiogenic therapy on slowly growing, poorly vascularized tumors in mice. *Journal of the National Cancer Institute*, **93**, 382–387.

Bentires-Alj, M., *et al*. (2000). Cytosine deaminase suicide gene therapy for peritoneal carcinomatosis. *Cancer and Gene Therapy*, **7**, 20–26.

Bischoff, J. R., *et al*. (1996). An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. *Science*, **274**, 373–376.

Black, M. E., *et al*. (1996). Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proceedings of the National Academy of Sciences of the USA*, **93**, 3525–3529.

Chen, C. T., *et al*. (2000). Antiangiogenic gene therapy for cancer via systemic administration of adenoviral vectors expressing secretable endostatin. *Human Gene Therapy*, **11**, 1983–1996.

Chou, J., *et al*. (1990). Mapping of herpes simplex virus-1 neurovirulence to gamma 134.5, a gene nonessential for growth in culture. *Science*, **250**, 1262–1266.

Coffey, M. C., *et al*. (1998). Reovirus therapy of tumors with activated Ras pathway. *Science*, **282**, 1332–1334.

Davis, B. M., *et al*. (2000). Limiting numbers of G156A O(6)-methylguanine-DNA methyltransferase-transduced marrow progenitors repopulate nonmyeloablated mice after drug selection. *Blood*, **95**, 3078–3084.

Dewey, R. A., *et al*. (1999). Chronic brain inflammation and persistent herpes simplex virus 1 thymidine kinase expression in survivors of syngeneic glioma treated by adenovirus-mediated gene therapy: implications for clinical trials. *Nature Medicine*, **5**, 1256–1263.

Feldman, A. L. and Libutti, S. K. (2000). Progress in anti-angiogenic gene therapy of cancer. *Cancer*, **89**, 1181–1194.

Fernandez, N., *et al*. (1998). Active specific T-cell-based immunotherapy for cancer: nucleic acids, peptides, whole native proteins, recombinant viruses, with dendritic cell adjuvants or whole tumor cell-based vaccines. Principles and future prospects. *Cytokines, Cellular and Molecular Therapy*, **4**, 53–65.

Hanania, E. G. and Deisseroth, A. B. (1997). Simultaneous genetic chemoprotection of normal marrow cells and genetic chemosensitization of breast cancer cells in a mouse cancer gene therapy model. *Clinical Cancer Research*, **3**, 281–286.

Hsueh, C. M., *et al*. (1996). Involvement of cytokine gene expression in the age-dependent decline of NK cell response. *Cellular Immunology*, **173**, 221–229.

Hunter, W. D., *et al*. (1999). Attenuated, replication-competent herpes simplex virus type 1 mutant G207: safety evaluation of intracerebral injection in nonhuman primates. *Journal of Virology*, **73**, 6319–6326.

Kaplan, J. M., *et al*. (1999). Induction of antitumor immunity with dendritic cells transduced with adenovirus vector-encoding endogenous tumor-associated antigens. *Journal of Immunology*, **163**, 699–707.

Khuri, F. R., *et al*. (2000). A controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer. *Nature Medicine*, **6**, 879–885.

Kikuchi, T., *et al*. (2000). Dendritic cells modified to express CD40 ligand elicit therapeutic immunity against preexisting murine tumors. *Blood*, **96**, 91–99.

Koc, O. N., *et al*. (1999). DeltaMGMT-transduced bone marrow infusion increases tolerance to O6-benzylguanine and 1,3-bis(2-chloroethyl)-1-nitrosourea and allows intensive therapy of 1,3-bis(2-chloroethyl)-1-nitrosourea-resistant human colon cancer xenografts. *Human Gene Therapy*, **10**, 1021–1030.

Kock, H., *et al*. (1996). Adenovirus-mediated p53 gene transfer suppresses growth of human glioblastoma cells *in vitro* and *in vivo*. *International Journal of Cancer*, **67**, 808–815.

Lang, F. F., *et al*. (1998). Enhancement of radiosensitivity of wild-type p53 human glioma cells by adenovirus-mediated delivery of the p53 gene. *Journal of Neurosurgery*, **89**, 125–132.

Lattime, E. G. and Stanton, S. (eds). (1999). *Gene Therapy of Cancer* (Academic Press, San Diego). 359–371.

Martuza, R. L., *et al*. (1991). Experimental therapy of human glioma by means of a genetically engineered virus mutant. *Science*, **252**, 854–856.

Mesnil, M., *et al*. (1996). Bystander killing of cancer cells by herpes simplex virus thymidine kinase gene is mediated by connexins. *Proceedings of the National Academy of Sciences of the USA*, **93**, 1831–1835.

Mineta, T., *et al*. (1994). Treatment of malignant gliomas using ganciclovir-hypersensitive, ribonucleotide reductase-deficient herpes simplex viral mutant. *Cancer Research*, **54**, 3963–3966.

Nielsen, L. L., *et al*. (1998). Adenovirus-mediated p53 gene therapy and paclitaxel have synergistic efficacy in models of human head and neck, ovarian, prostate, and breast cancer. *Clinical Cancer Research*, **4**, 835–846.

Norman, K. L. and Lee, P. W. (2000). Reovirus as a novel oncolytic agent. *Journal of Clinical Investigation*, **105**, 1035–1038.

Ojeifo, J. O., *et al*. (1995). Angiogenesis-directed implantation of genetically modified endothelial cells in mice. *Cancer Research*, **55**, 2240–2244.

Patel, D. H., *et al*. (2000). Retroviral transfer of the hENT2 nucleoside transporter cDNA confers broad-spectrum antifolate resistance in murine bone marrow cells. *Blood,* **95**, 2356–2363.

Puhlmann, M., *et al*. (2000). Vaccinia as a vector for tumor-directed gene therapy: biodistribution of a thymidine kinase-deleted mutant. *Cancer and Gene Therapy*, **7**, 66–73.

Putzer, B. M., *et al*. (1998). Combination therapy with interleukin-2 and wild-type p53 expressed by adenoviral vectors potentiates tumor regression in a murine model of breast cancer. *Human Gene Therapy*, **9**, 707–718.

Qiao, J., *et al*. (2000). Enhanced ganciclovir killing and bystander effect of human tumor cells transduced with a retroviral vector carrying a herpes simplex virus thymidine kinase gene mutant. *Human Gene Therapy*, **11**, 1569–1576.

Ragg, S., *et al*. (2000). Direct reversal of DNA damage by mutant methyltransferase protein protects mice against

dose-intensified chemotherapy and leads to *in vivo* selection of hematopoietic stem cells. *Cancer Research*, **60**, 5187–5195.

Ribas, A., *et al*. (1999). Characterization of antitumor immunization to a defined melanoma antigen using genetically engineered murine dendritic cells. *Cancer and Gene Therapy*, **6**, 523–536.

Sawai, N., *et al*. (2001). Protection and *in vivo* selection of hematopoietic stem cells using temozolomide, O^6-benzylguanine, and an alkyltransferase-expressing retroviral vector. *Molecular Therapy*, **3**, 78–87.

Scappaticci, F. A., *et al*. (2001). Combination angiostatin and endostatin gene transfer induces synergistic antiangiogenic activity *in vitro* and antitumor efficacy in leukemia and solid tumors in mice. *Molecular Therapy*, **3**, 186–196.

Strong, J. E., *et al*. (1998). The molecular basis of viral oncolysis: usurpation of the Ras signaling pathway by reovirus. *EMBO Journal*, **17**, 3351–3362.

Szabolcs, P., *et al*. (1997). Retrovirally transduced human dendritic cells express a normal phenotype and potent T-cell stimulatory capacity. *Blood*, **90**, 2160–2167.

Vitaliti, A., *et al*. (2000). Inhibition of tumor angiogenesis by a single-chain antibody directed against vascular endothelial growth factor. *Cancer Research*, **60**, 4311–4314.

Wan, Y., *et al*. (1999). Enhanced immune response to the melanoma antigen gp100 using recombinant adenovirus-transduced dendritic cells. *Cellular Immunology*, **198**, 131–138.

Watanabe, T. and Sullenger, B. A. (2000). Induction of wild-type p53 activity in human cancer cells by ribozymes that repair mutant p53 transcripts. *Proceedings of the National Academy of Science of the USA*, **97**, 8490–8494.

FURTHER READING

Carmeliet, P. and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature*, **407**, 249–257.

Fernandez, N., *et al*. (1998). Active specific T-cell-based immunotherapy for cancer: nucleic acids, peptides, whole native proteins, recombinant viruses, with dendritic cell adjuvants or whole tumor cell-based vaccines. Principles and future prospects. *Cytokines, Cellular and Molecular Therapy*, **4**, 53–65.

McCormick, F. (2000). Interactions between adenovirus proteins and the p53 pathway: the development of ONYX-015. *Seminars in Cancer Biology*, **10**, 453–459.

Melero, I., *et al*. (2000). Feeding dendritic cells with tumor antigens: self-service buffet or a la carte? *Gene Therapy*, **7**, 1167–1170.

Norman, K. L. and Lee, P. W. (2000). Reovirus as a novel oncolytic agent. *Journal of Clinical Investigations*, **105**, 1035–1038.

Roth, J. A., *et al*. (1999). p53 tumor suppressor gene therapy for cancer. *Oncology*, **13**, 148–154.

Roth, R. B. and Samson, L. D. (2000). Gene transfer to suppress bone marrow alkylation sensitivity. *Mutation Research*, **462**, 107–120.

Chapter 70

Models for Epithelial Carcinomas

Suzanne A. Eccles
Institute of Cancer Research, Sutton, Surrey, UK

CONTENTS

- Introduction: Basic Biology of Epithelial Cancers
- Epithelial Cancer Models
- Skin and Squamous Cell Carcinomas
- Gastrointestinal Tract Tumours
- Lung and Bronchial Carcinoma
- Breast Carcinoma
- Prostate Cancer
- Conclusions and Perspective

INTRODUCTION: BASIC BIOLOGY OF EPITHELIAL CANCERS

All adult tissues are derived from three embryonic germ layers: ectoderm generates the likes of epidermis and nervous system, mesoderm connective and supportive tissues and endoderm the many epithelial organs. Although tissues of mesenchymal origin account for up to 80% of the mass of an individual, over 90% of adult human tumours are epithelial carcinomas. This may be explained by the fact that many such tissues (e.g. skin, respiratory, urogenital and alimentary tracts) are exposed to environmental carcinogenic hazards. Common examples include ultraviolet rays in sunlight (skin cancers), air pollution and cigarette smoke (lung cancers), asbestos (mesothelioma) and less well defined dietary risk factors which may contribute to the development of gastrointestinal (GI) tract cancers. Internal organs such as mammary gland, liver, kidney and bladder may also be affected by toxic agents, some of which (such as aflatoxin, which affects the liver, and aniline dyes, which target the urinary tract) may exhibit tissue-specific effects. Viruses are also involved in the aetiology of certain carcinomas (e.g. papilloma virus, cervical cancer; hepatitis virus, hepatocellular carcinoma; and Epstein–Barr virus, nasopharyngeal carcinoma). (See also chapter on *Cell and Tissue Organisation*.)

Normal Epithelial Organisation

The evolution of complex multicellular organisms required the development of highly coordinated and specific cell interactions. During embryogenesis, cell contacts with other cells and extracellular matrix (ECM) components play major roles in directing morphogenesis, migration and differentiation in exquisitely orchestrated developmental processes. These processes are also modulated by soluble factors (generally termed 'growth factors') in the environment. Epithelial cells are the first to develop in the embryo, with motile, nonadhesive mesenchymal cells differentiating from them. Throughout history, the notion that malignant cells represented some form of embryonic stem cell has been a recurring theme. It was debated whether these were occult fetal tissues or cells which *de novo* recapitulated the proliferative and invasive character of such cells. It is now known that remodelling of tissues following injury and oncogenic progression can involve transient epithelial–mesenchymal phenotypic conversions and re-expression of genes normally only active in the embryo.

In the adult, normal epithelial cells are polarized and organized broadly into squamous sheets or secretory structures via intercellular attachments including desmosomes, tight junctions and adherens junctions (**Figure 1**). The basal cell layer secretes proteins that constitute the basement membrane, with which close contact is maintained by further adhesive interactions. The cell shape, rigidity and structure are maintained by a complex system of epithelial-specific intermediate cytokeratin filaments. Aggregations of epithelia are separated and supported by a connective tissue stroma incorporating blood vessels and loosely associated, more mobile mesenchymal cells. Although the cell–cell and cell–matrix contacts maintain tissue structure and integrity, it is becoming clear that the adhesion molecules involved provide far more than a mere physical supporting role. Recent studies have provided evidence that a cell's differentiation, gene expression, motility and even survival depend upon correct interactions with its immediate

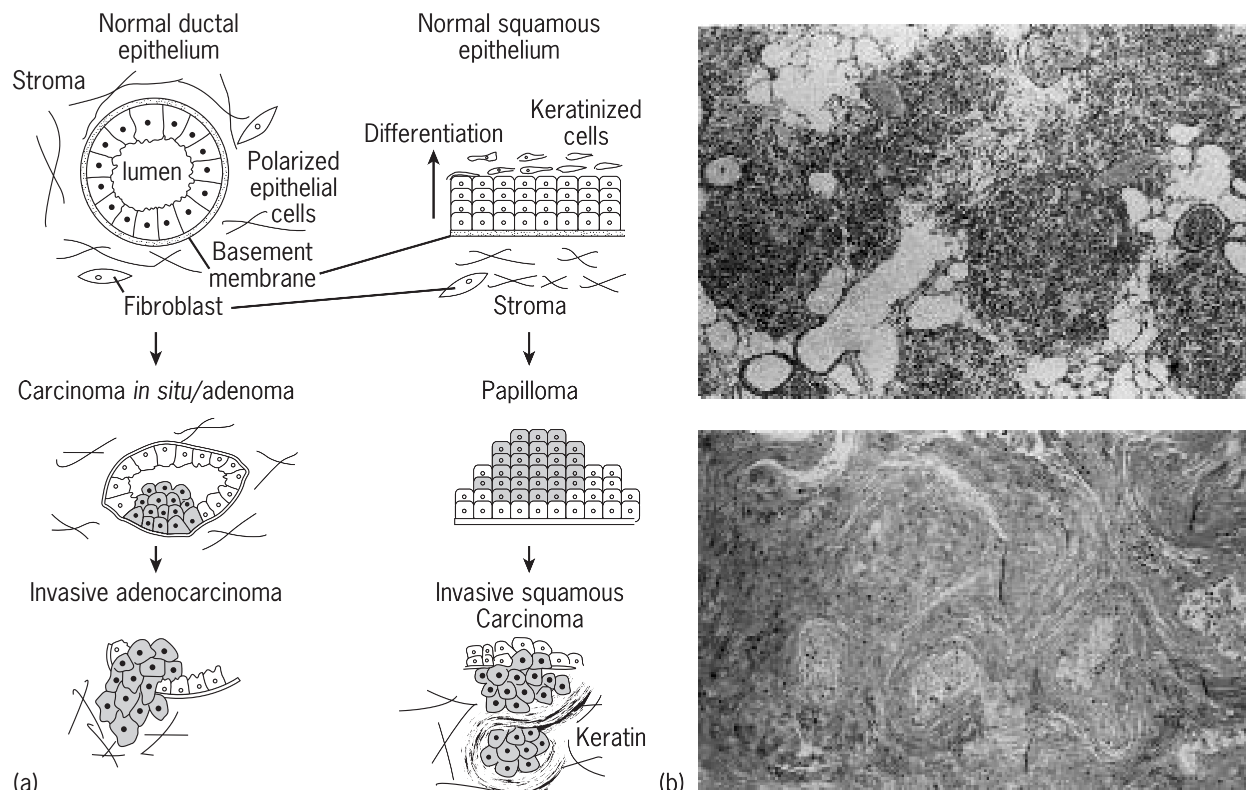


Figure 1 (a) Epithelial organisation; (b) examples of carcinomas: (top) lung adenocarcinoma and (bottom) oral squamous carcinoma.

environment. Many safeguards exist to ensure that cells which are incorrectly positioned do not survive (a process known as apoptosis or anoikis), and cancer cells are those which fail to respond to these normal social regulatory signals. (See also chapter on *Extracellular Matrix: the Networking Solution*.)

Regulation of Tissue Structure and Function: Homotypic Epithelial Adhesion is Down-regulated in Invasive Cells

One of the adhesion molecules important in maintaining the integrity of epithelial structures is the calcium-dependent glycoprotein E-cadherin. It is functionally inactivated in nearly all epithelial human cancers by a variety of mechanisms (**Figure 2**). Germ-line mutations in the E-cadherin gene are linked with predisposition to invasive gastric cancer, and in a transgenic mouse model of pancreatic carcinoma, loss of E-cadherin is causally linked with benign to malignant conversion. In human squamous carcinomas of the head and neck, E-cadherin expression is inversely correlated with the presence of lymph node metastases; similar findings have been reported in gastric and hepatocellular carcinomas. In some experimental studies, when the E-cadherin gene was introduced into carcinoma cells, which lacked expression, their invasive potential was lost and intercellular adhesion enhanced. In other cases, such as the human breast carcinoma cell line MCF-7, invasive variants continued to express high levels of E-cadherin, suggesting either that some function of the molecule was impaired, or alternatively that other 'positive' invasion signals may over-ride its influence in certain cell types (Christofori and Semb, 1999).

The intracellular domains of cadherins associate with catenins and the actin cytoskeleton. β-Catenin, when released from E-cadherin, can move to the nucleus and in association with transcription factors can induce expression of genes involved in the cell cycle. Although the mechanisms are far from fully elucidated, the circumstantial evidence in favour of an important regulatory role for E-cadherin in tumour progression is compelling. In developmental processes, it is clear that migratory cells must first detach from their neighbours, and this must involve a change in their adhesive status. It seems that the transient, orderly control of molecules such as the cadherins, if misregulated, could lead to the evolution of epithelial cells which have acquired the properties more usually associated with mesenchymal cells; lack of polarization and differentiation and a motile, migratory phenotype. (See also chapter on *Modelling Tumour–Tissue Interactions*.)

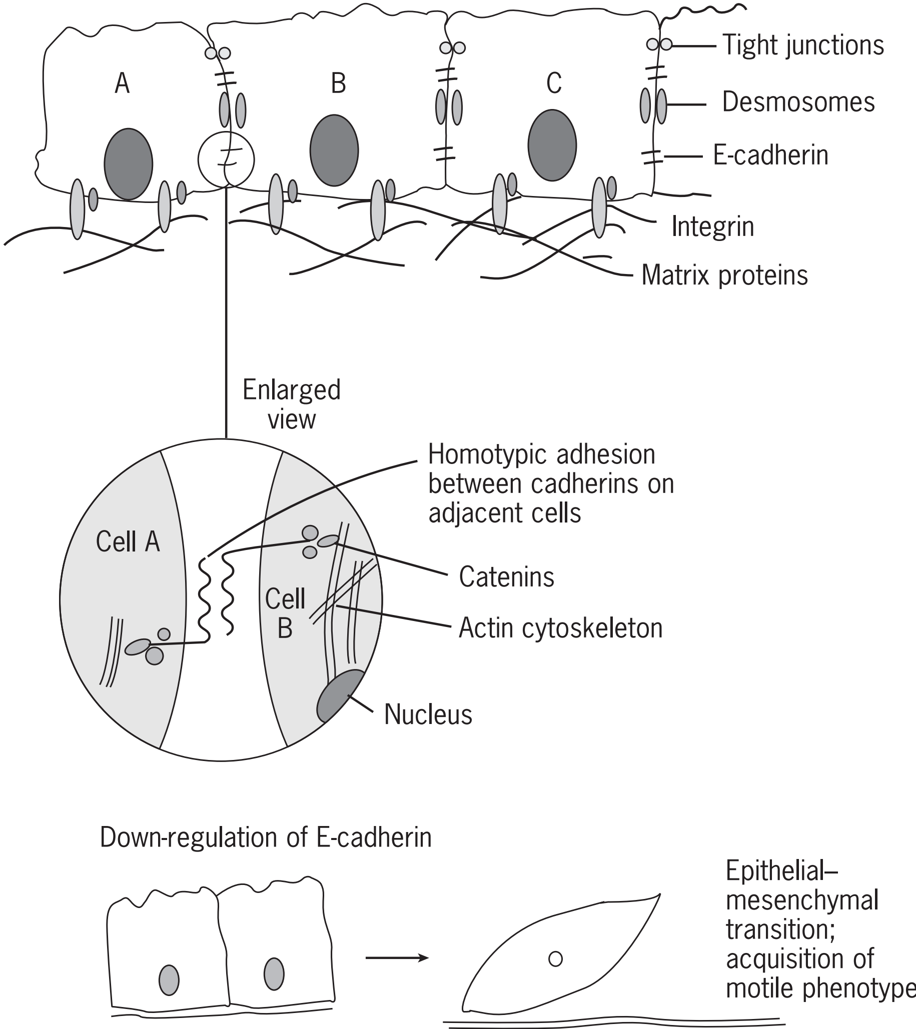


Figure 2 Mechanisms of inactivation of E-cadherin.

Epithelial–Mesenchymal Transformations in Embryogenesis and Metastasis

The transient phenotypic conversion of epithelial cells to a mesenchymal morphology is a common event during embryonic development. In the adult, epithelial cells normally only migrate during wound healing or tissue repair. In one experimental system utilizing a rat bladder carcinoma cell line, the epithelial–mesenchymal transition could be induced by fibroblast growth factor 1 (FGF-1) (a-FGF) or by culturing the cells on collagen type 1 fibres. The resulting cells showed a rapid down-regulation of cadherins, loss of their ability to form aggregates and acquisition of motility (**Figure 2**). Interestingly, an early response to FGF-1 was the production of matrix metalloproteinases, illustrating how such a phenotypic shift could provide the range of abilities required for invasion by either the embryonic or malignant cell (Valles *et al*., 1991).

Mesenchymal–Epithelial Interactions

During organ development, many morphogenic interactions between epithelia and mesenchyme occur and organogenesis has been studied *in vitro* in tissue explants. Mesenchymal factors have been found to be essential for growth and morphogenesis of epithelial cells in most if not all systems studied (Birchmeier and Birchmeier, 1993). The paracrine (indirect) signals may be transmitted by three different modes: cell–cell contact, cell–matrix interactions or diffusion of soluble factors. Direct contact between epithelial and mesenchymal cells has been shown to be important in the developing mammary gland and in the induction of new epithelia during kidney development. However, separation of the two cell types by filters has shown that mesenchymal induction of morphogenesis in other tissues (e.g. salivary gland) does not require direct contact. Components of the extracellular matrix such as collagens can support branching morphogenesis of

mammary epithelial cells, and soluble mediators such as 'scatter factor' and a variety of growth factors can also participate in morphogenesis (see below).

Fibroblasts, the classical mesenchymal cells, are nonpolarized, motile and generally found surrounded by extracellular matrix. They interact with components of the ECM via adhesion molecules such as integrins for which they express a variety of receptors. They secrete a variety of ligands that interact in a paracrine fashion with epithelial-specific receptors. In addition, it is known that several proteases can be produced by mesenchymal cells which bind to receptors on epithelial cells to mediate directed proteolysis, e.g. during tooth eruption, capillary sprouting or ductal branching. Coordinated expression of matrix-degrading proteases and their inhibitors is also involved in involution of the mammary gland. (See also chapters on *Mammary Tumour Induction in Animals as a Model for Human Breast Cancer*; *Modelling Tumour Tissue Interactions*.)

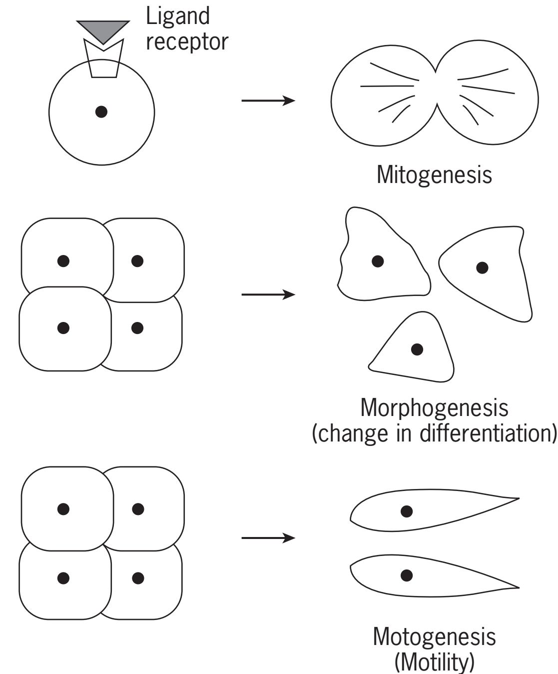


Figure 3 Pleiotropic effects of 'growth factors' (e.g. EGF, HGF).

Matrix and Growth Factors

Tissue development, renewal and repair occur with the assistance of a wide variety of regulatory growth factors that may provide both inhibitory and stimulatory signals. The responses induced depend upon interplay between the origin and state of differentiation of the target cell, the number and types of receptors it expresses and the concentration of growth factor(s) in its environment. In order to respond to these molecules, cells must express specific transmembrane receptors that can bind the growth factors. This process transmits signals via a cascade of intracellular activation events that culminate in a number of responses in the target cell, the most common of which is mitosis. However, the term 'growth factor' may be misleading since it is evident that these ligands induce pleiotropic responses and may alter states of differentiation, cytoskeletal organisation, motility and other cellular functions. The same growth factor may produce varying responses in either different cell lineages, or the same cell type at different stages of differentiation – again emphasizing the importance of 'context' in developmental decisions. For example, 'scatter factor' is secreted by fibroblasts and induces epithelial cells to dissociate (scatter) in culture; the cells also become dedifferentiated and more invasive (**Figure 3**). This molecule was found to be identical with hepatocyte growth factor (HGF) (a mitogen for liver cells) and to induce these diverse responses via a single receptor – the product of the c-*met* proto-oncogene – in all cell types. HGF has been implicated in acquisition of an invasive phenotype in many types of carcinomas.

A factor may also evoke different responses depending on the quantity available; fibroblasts respond to low concentrations of FGF-1 by proliferation and to high concentrations by movement towards it (chemotaxis). Once tissue architecture is defined, further growth and the replacement of those tissues (such as skin, intestine and haematopoietic system) which are self-renewing depends upon the proliferation of a limited number of multipotent 'stem' cells. These generate developmentally restricted progenitor cells whose progeny finally yield differentiated, fully functional, nonproliferative mature cells. This coordinated process must be closely regulated in order to maintain tissue integrity and function, and indeed it is the imbalance of proliferation and differentiation/cell death that is at the heart of malignant transformation. Many tumours have been shown to secrete growth factors and/or to over-express their receptors and indeed many of the oncogenes recently identified have been shown to encode these molecules or other components of the mitogenic signalling pathway. The potential of certain viruses to cause malignant transformation of cells is due to the integration of such oncogenes into their genomes. (See also chapters on *Human DNA Tumour Viruses*; *RNA Viruses*.)

Such observations led to the 'autocrine' hypothesis of tumour development, i.e. the idea that a tumour cell survives because it is no longer dependent on endocrine or paracrine growth stimulatory signals, but can produce its own. This leads to an autostimulatory loop and increased autonomy from normal growth controls. In the most extreme cases the receptor may be constitutively activated independently of ligand binding and hence constantly transmitting stimulatory signals to the cell. There is now

compelling evidence that mutation or overexpression of growth factors and/or their receptors may provide a growth advantage for tumour cells, indirectly stimulate vascularization on which the continued expansion of the tumour mass depends, and potentiate processes involved in invasion and dissemination.

Benign or Malignant? Carcinoma *In Situ*

As Virchow is supposed to have remarked, 'no man even under torture can say what is a tumour,' and even modern pathologists have to use an armamentarium of criteria to classify tumours and to discriminate benign from malignant lesions. In the past, diagnosis was ultimately determined only by the eventual outcome (survival or death), although then, as now, certain features of the location and appearance of the tumour provided clues as to its probable behaviour. We now like to think that tumour classification and staging have evolved to a sophisticated art, and there is a general tacit understanding of the difference between benign tumours (which are generally indolent, well-differentiated, circumscribed and localized masses) and malignant lesions with a tendency to develop more rapidly, invade locally and spread to distant sites.

For certain carcinomas there are defined premalignant or benign lesions which may represent stages in the evolution of a malignant tumour from its normal tissue of origin; other hyperplastic conditions do not appear to predispose to the later development of cancer, and for some tumours (e.g. pancreatic and renal carcinomas) there seems to be no premalignant state. Until recently, these lesions were described in morphological terms, but increasingly, molecular probes are being used to dissect the multiple genetic defects thought to be necessary to produce fully metastatic tumours. Since many tumours are known to have a long occult phase prior to detection, it is not surprising that the early stages of malignant conversion have been so elusive. The prime example of this is the case of so-called 'unknown primaries' where the first and only manifestation of cancer is a metastasis.

Although cancer is often described as the abnormal growth of cells, over-proliferation *per se* does not define malignancy. Hyperplasia represents the increased proliferation of essentially normal cells, e.g. after injury (physiological) or in psoriasis (pathological). Dysplasia may generate atypical or abnormal cells, but these generally do not proliferate indefinitely and are not invasive. Metaplasia is defined as a major switch in phenotype and these metaplastic cells may eventually overgrow and replace other cellular elements in the affected tissue; it is this latter category which may represent one of the earliest steps in oncogenesis, e.g. squamous cell metaplasia in the airways of the lung. At this point the lesion remains *in situ*, has generally not breached the surrounding basal membranes and may be encapsulated in a zone of connective tissue.

Tumour Progression, Heterogeneity and Clonal Dominance

It is generally accepted that most cancers arise from a single aberrant cell. The clonal nature of several malignant diseases has been demonstrated by examining genetic and biochemical markers such as the presence of chromosomal abnormalities, secretion of a single species of immunoglobulin (plasma cell tumours) and isoenzyme patterns. One possibility is that the progenitor cell is transformed (perhaps by the presence of an aberrant gene or exposure to an environmental carcinogen) and that this then proliferates to generate the clinically detectable tumour. An alternative hypothesis suggests that many cells undergo 'initiation,' but that only one or a few 'progress' to full malignancy. Certainly there are genetic traits where an individual is at a high risk of developing cancer in one or more organ systems, and in some instances this is due to mutation or the loss of genes important for the regulation of mitosis, repair of DNA damage or programmed cell death (apoptosis). In these individuals certain tissues may be primed and require only a single further event to generate a tumour. In such cases multiple primary tumours are not unusual.

Foulds first formulated the concept of tumour progression in 1975. At about this time, experimental models of carcinogenesis were demonstrating that certain compounds could cause cancer only if animals were exposed to sequential treatments with other compounds. Some chemicals worked if administered initially ('initiators'), followed by others that acted to 'promote' the oncogenic potential of the former, but were not themselves carcinogenic. Some complex mixtures (such as cigarette smoke) were described as 'complete carcinogens' because they contained both initiators and promoters. Nowell, in his clonal evolution theory, proposed that the carcinogenic event induced 'genetic lability' which permitted 'stepwise selection of variant sublines and underlies tumour progression.' This model, and its experimental background, formulated the concept of the multistep nature of carcinogenesis and the progression from normal cells through premalignant lesions to fully malignant cancers capable of metastasis. Tumours therefore can acquire increased heterogeneity of a variety of phenotypic and genotypic traits (some of which will provide a growth advantage to particular cell subpopulations) but can retain the unique markers that identify their monoclonal (single cell) origin. Recent work has begun to define the molecular genetic bases for some of these steps in the tumours where premalignant lesions are identifiable and accessible, notably colorectal carcinoma, squamous cell carcinomas and to a

lesser extent melanoma and breast carcinoma. (See also chapter on *Overview of Oncogenesis*.)

Metastasis

For some common tumours such as lung, breast and colon carcinoma and malignant melanoma there is a correlation between the size of the primary lesion and the probability of metastasis. For an individual tumour, size generally equates with time, and this observation probably reflects the fact that 'older' tumours have had longer in which to acquire multiple genetic alterations, now thought to underlie the well-recognized multistep nature of oncogenesis and progression. The cell population as it enlarges will generate increasing heterogeneity from which subpopulations with relatively greater autonomy from host environmental controls will emerge by adaptation or selection. It is clear that genes controlling tumour growth can be distinct from those controlling metastasis, since some tumours that are capable of indefinite growth fail to spread to distant organs. It is this aspect of cancer that is the most challenging both in terms of producing accurate models for the better understanding of the processes involved, and also in the development of effective systemic therapies. The importance of neoangiogenesis, not only in development of primary tumours, but also in facilitating metastasis and as a valuable therapeutic target is now fully appreciated (Fidler and Ellis, 1994). (See also chapters on *Angiogenesis Models* and *Tumour Metastasis Models*.)

EPITHELIAL CANCER MODELS

Given that there are over 200 different types of cancer, most of which are epithelial in origin, it is clearly not possible to discuss models of more than a fraction of them. The general principles involved in malignant transformation have been described above, and now it is possible to discuss in more detail specific cases where these processes can be mimicked *in vitro* and/or *in vivo*. We accept that much of cancer is genetically determined, but even when a specific gene (e.g. *APC* or *BRCA1*) is known to predispose to cancer, the host genetic background can contribute significantly to its effects. This means that the same gene in mouse and humans will never give exactly the same incidence, types and spectrum of behaviour of cancers. Nevertheless, manipulation of genes either by mutation, or overexpression or the 'knockout' of tumour-suppressor genes has proved very useful in both providing experimental cancers for study and also determining the function of these genes. (See also chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*.)

Similarly, animals can be exposed to carcinogenic chemicals, toxins, irradiation, viruses, etc., and cancers may ensue. In many cases the exposures are acute and high dose, and this may give different (often more highly immunogenic) tumours to those which arise from chronic, lower dose exposure as is usually the case in humans. More recently, cultured animal and human epithelial cells have been exposed to agents identified as having carcinogenic potential in order to simplify analysis, although this reductionist approach cannot match the complexity of interactions that can lead to metastatic cancer in the body.

It is important to define from the outset what the 'model' is designed to model (see chapter on *Advantages and Limitations of Models for Cancer and Malignant Cell Progression*). This may seem to be stating the obvious but there are very different imperatives if one is aiming to study, for instance, oncogenic transformation of a particular epithelial cell lineage *in vitro*, the process of chemical or viral carcinogenesis, chemoprevention or the biology or therapy of primary or secondary cancers. The next sections will cover models of several common cancers, and these will be cross-referenced to other cancer types where general principles can be defined.

SKIN AND SQUAMOUS CELL CARCINOMAS

Chemical Carcinogenesis–Squamous Cell Carcinomas: the Mouse Skin Model

Many of the classical concepts of multistage carcinogenesis have been derived from the mouse skin model in which the three basic stages of carcinogenesis (initiation, promotion and progression) have been demonstrated. The historical reasons for this approach are readily apparent since observations on scrotal skin cancer in chimney sweeps led Percival Pott to elaborate the first theories on environmental/chemical causes of cancer in the eighteenth century. Suspect substances could be 'painted' on to animal skin where the effects on the epidermal cells were readily observed. This system has led to the extraction and identification of substances from complex mixtures that are carcinogenic in humans, notably 1,2,5,6-dibenzanthracene and 3,4-benzo[*a*]pyrene from coal tar, polycyclic aromatic hydrocarbons (PAHs) from cigarette smoke and the diterpene esters from croton oil. However, this model has some limitations as a predictor of risk factors in humans. Marked differences in susceptibility exist not only between different species but also between strains of mice, and the structure of mouse skin is different from that of human skin in terms of morphology, hair cycle and permeability to exogenous compounds. Arsenic can cause skin cancers in humans but has not done so in any animal so far tested, and only about 70% of known carcinogens cause cancers when applied to the skin (reviewed in Eccles, 1987).

Nevertheless, this simple model has over many decades yielded vital information about which agents act at different stages of carcinogenesis, what their genetic targets are, the role of cofactors and epigenetic influences and finally has provided test systems for evaluating potential inhibitors of these events.

The 'initiation' stage is generally considered to require the direct interaction of the compound with DNA in stem cells leading to an irreversible genetic defect, often involving *ras* oncogenes. The H-*ras* locus has been shown to undergo many different changes that can lead to malignancy including mutagenic activation, amplification of the mutant allele and loss of the normal allele (Akhurst and Balmain, 1999). The way that skin tumour promoters seems to work is by expansion of the initiated stem cells offering the potential for buildup of further genetic changes, genomic instability and the conversion of a benign papilloma to a malignant carcinoma. In the absence of an initiation event, tumour promoters merely induce hyperplasia – overgrowth of nonmalignant cells (Slaga *et al*., 1996). H-ras induces a 'papilloma' phenotype, characterized by a high rate of cell proliferation and aberrant differentiation. This phenotype has been ascribed to activation of the epidermal growth factor receptor (EGFR) and effects on different protein kinase C (PKC) isoforms (these are intracellular signalling molecules). Premalignant progression requires additional genetic events which at this stage are inhibitable by transforming growth factor beta (TGFβ). Cells that have low levels of TGFβ are more readily transformed, exemplifying the fact that progression can be linked to gain of oncogene function and/or loss of suppressor gene function. Interestingly, at later stages of progression, TGFβ can act to stimulate tumour invasion, indicating that cellular context is important to gene function (Ackhurst and Balmain, 1999) (**Figure 4**).

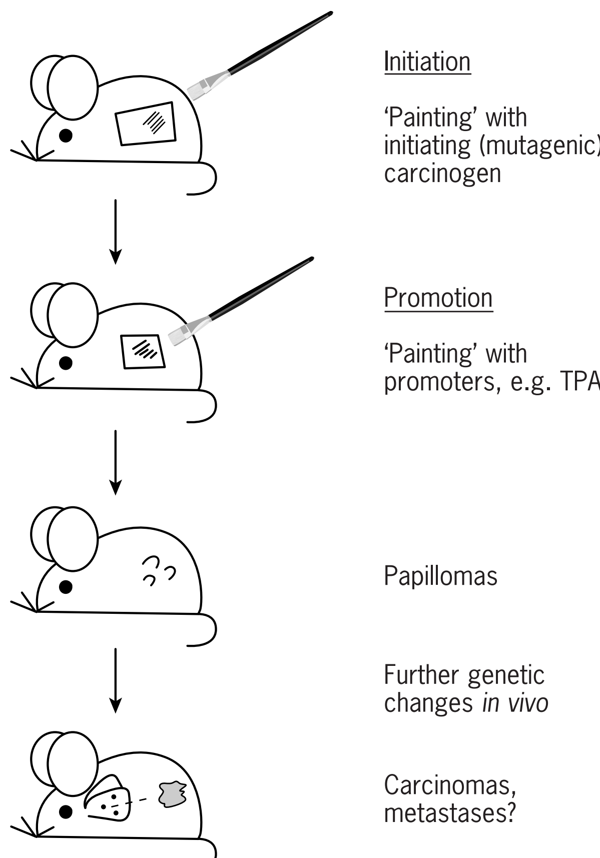


Figure 4 Skin carcinogenesis model.

Intervention/Therapy Studies

Inositol hexaphosphate (IP6) has been demonstrated to prevent carcinogenesis in a variety of experimental systems (Shamsuddin, 1999), although it seems that its metabolism and biodistribution may be different in mice and humans. Its mechanisms of action are not fully understood, but it appears to have several interesting properties: inhibition of cell proliferation, enhanced differentiation of malignant cells, down-regulation of mutant p53 and stimulation of DNA repair. In a classical two-stage initiation–promotion model using the initiator 7,12-dimethylbenzanthracene (DMBA) and the promoter TPA, it was shown that giving IP6 in the drinking water was able to inhibit papilloma formation significantly if given in the initiation (first 3 weeks) period, but not if administration was delayed until the promotion period (weeks 4–23).

In Vitro Models of Skin Tumour Progression

More recently, murine and human cell lines have been derived from keratinocytes (the skin epithelial cells) at various stages of progression. One of the most widely used is the human HaCaT series (Fusenig and Boukamp, 1998). The original cell line was obtained from normal human skin cells. Nonmalignant cells will grow in culture, but generally undergo senescence and die after 60–70 cell divisions. Immortalization (one of the first stages in malignant progression) may occur spontaneously in rare clones, or can be induced by various means. These include overexpression of the telomerase (*hTERT*) gene which acts to keep chromosomal 'ends' or telomeres long enough to allow repeated DNA replication and cell division, mutation of the *p53* 'guardian of the genome' gene involved in DNA repair or introduction of viral genes which also overcome senescence by interfering with cell cycle control. Similar cultures can be made from oral, bladder, rectal and oesophageal squamous epithelial cells.

By exposure of early cultures of HaCaT to various agents, a panel of cell lines has been obtained which show different genetic changes and malignant potential. Ultraviolet (UV) B (exposure to which is known to be a

risk factor for skin cancers, including melanoma which does not arise from epithelial cells) induced p53 mutations and loss of senescence genes on chromosome 3p. Interestingly, exposure of the HaCaT cells to inorganic arsenic (which as noted above causes skin and other cancers in humans) also reduced p53 levels and increased the levels of another gene, *mdm2*, which regulates p53 in normal cells (Hamadeh *et al*., 1999). This suggests that different risk factors for skin carcinogenesis may impact on key cell regulatory elements, notably the p53–Mdm2 loop which regulates cell cycle arrest in damaged cells. The loss of this control enables cells with DNA mutations to continue to proliferate. It is now also possible to reconstruct human skin epidermis *in vitro* by growing keratinocytes either on dead de-epidermized dermis (DED), dermal equivalent (collagen–fibroblast lattices) or synthetic substrates (Lenoir-Viale, 1996). Although primarily developed for skin graft research and cosmetic testing, these systems may also lend themselves to studies of malignant progression and therapeutic interventions.

In Vivo Xenograft Models of Squamous Cell Carcinoma Tumour Progression

So far, examples have been given of how normal human keratinocytes in tissue culture can be immortalized by various different agents, and some of the genes involved. At this point the cells are only able to generate benign papillomas, and not tumours when transplanted into immunodeficient mice. However, tumorigenic transformation can be induced by either introduction of mutant H-ras or exposure to heat. Growth of HaCaT cells at 40 °C induced chromosomal alterations including gains of regions of chromosome 11 which contains genes such as cyclin D1, involved in cell cycle control (Boukamp *et al*., 1999). This is interesting since it suggests that the effects of sunburn could be due to both UV exposure and high temperatures which produce different, but potentially complementary, genetic changes in skin epithelial cells. Malignant conversion was also associated with loss of chromosome 15 and a decrease in expression of the thrombospondin (*TSP*) gene. This molecule is important in controlling angiogenesis (the development of new blood vessels) on which tumour growth and metastasis depend. It is now thought that the ‘angiogenic switch’ is a key element in the conversion of benign to malignant tumours, since if a focus of cells is unable to induce a new blood supply, its growth and invasion are limited. The cells that had become malignant could be prevented from forming tumours by restoring chromosome 15 or the *TSP* gene to them. These procedures have now generated cell lines which together encompass all of the stages of human skin tumour progression from hyperplasia through to invasive cancers.

When cell lines of increasing malignancy (HaCaT, A5, II4RT) were compared, it was found that specific matrix metalloproteinases (MMP3, 9 and 10) increased coordinately (Bachmeier *et al*., 2000). These enzymes are very important in many aspects of tumour growth and invasion, releasing sequestered tumour growth factors and angiogenic factors, breaking down the ECM, thus potentiating tumour cell invasion, and MMP-9 has also been suggested to be involved in the angiogenic switch by releasing active vascular endothelial growth factor (VEGF) (Bergers *et al*., 2000). Xenografts can also be established directly from human cancers, or via cell lines derived from human cancers, although they rarely metastasize and recapitulate all of the malignant features of the original tumour. It is now appreciated that the malignant phenotype is best expressed if the tumours are grown in orthotopic (anatomically correct) sites rather than the convenient subcutaneous (s.c.) site in athymic mice. In some cases the latter site has been shown to downregulate key genes involved in invasion and angiogenesis. When a human basaloid oral SCC was transplanted s.c. or in the oral submucosa of athymic mice, the resulting lesions showed clear differences, with preferential growth of either the basaloid or squamous cells, which each exhibited different malignant features.

Microenvironmental/Host Influences in Skin Carcinogenesis

It is clear that the microenvironment and cell–cell interactions play a key role in tumorigenic conversion although the processes involved are not fully understood. Vaccariello *et al*. (1999) used an organotypic tissue model which mimics premalignancy to monitor genetically tagged malignant II4 cells cultured with either normal human keratinocytes or immortal HaCaT cells. Clonal expansion of these cells was prevented by direct contact with the former. Further progression of the tumorigenic HaCaT clones was found to require the *in vivo* environment which induced further genetic changes and selection of tumour cell subpopulations with enhanced autonomous growth. Various growth factors were induced which acted in an autocrine manner to stimulate both tumour cell proliferation and migration, which is a key element of an invasive phenotype (Fusenig and Boukamp, 1998). HaCaT cell clones capable of either benign or malignant (invasive) growth can be placed as surface transplants on to the subepidermal stroma of athymic (immunodeficient) mice to explore tumour–host interactions further. At 1 week after seeding, both types of transplants appear similar to healing wounds, but later the malignant cells revealed loss of cell polarity and defective basement membranes and cell–cell contacts.

In humans, UV irradiation is the main cause of skin tumours and the effects of UVB and chemical carcinogens

have been investigated in relation to the role of the extracellular matrix. Stenback *et al*. (1999) showed that these two agents induced different stromal reactions, notably in the synthesis and deposition of collagen types I and III. Further reproducible changes occurred during tumour development and progressive loss of differentiation. In a syngeneic model of squamous cell carcinoma similar to the HaCaT system, Dong *et al*. were able to explore differential gene expression during tumour progression in immunocompetent hosts. They used primary Balb/c mouse keratinocytes, a spontaneously transformed cell line (Pam 12) and metastatic cell lines derived from lymph nodes (Pam LY) and lung (Pam LU). A variety of genes were upregulated in the metastatic cell lines including the Growth Regulated Oncogene (GRO) alpha family and other genes linked to cell shape, motility and chemotaxis, signal transduction and angiogenesis (Dong *et al*., 1997). GRO alpha is a member of the C–X–C family of chemokines which promote chemotaxis of granulocytes and endothelial cells by binding to CXC receptor 2. Tumours overexpressing GRO alpha were fast growing, highly angiogenic and metastatic, but this advantage was diminished in hosts with defective CXC receptor 2, indicating that tumour malignancy depends on critical interactions with the microenvironment (Loukinova *et al*., 2000).

One of the genes which showed changes in expression (e.g. *Ly-6A/E*) is homologous to the human E48 SCC antigen which is implicated in cell–cell recognition, and hence may contribute to tumor progression via immune escape mechanisms. Upregulation of this gene would not be selected for *in vitro* or in immunodeficient hosts, indicating the importance of including syngeneic animal systems in tumour modelling. Both human and murine squamous cell carcinomas (SCC) have been shown to express cytokines involved in inflammation, e.g. interleukins 1α, 6 and 8 and granulocyte–macrophage colony-stimulating factor (GM-CSF). This is associated with a metastatic phenotype and only occurs *in vivo* and not during *in vitro* transformation, again emphasizing the role of the host environment in potentiating later stages of tumour progression.

Transgenic and ‘Knockout’ Mouse Models for Squamous Cell Carcinomas

These systems are covered in depth in other chapters (see *Transgenic Technology in the Study of Oncogenes and Tumour Suppressor Genes*; *Gene Knockouts in Cancer Research*), but a few observations are worthy of note. As stated previously, tumour growth and invasion are sustained partly by the binding of ‘growth factors’ to their receptors on the cell membrane. Key signalling systems in epithelial carcinomas include the c-erbB family (EGFR, c-erbB 2, 3 and 4) which are activated by various ligands including EGF, transforming growth factor alpha (TGFα) betacellulin and the heregulins (Eccles *et al*., 1995; Eccles, 2000) and also the insulin growth factor receptor. To determine the role of receptors and ligands in tumour development, transgenic mice have been made which express either amplified or mutant genes. The ability to introduce genes stably into the germ line of mice has converted the study of multistage chemical carcinogenesis to multistage molecular carcinogenesis. By combining genes with regulatory sequences that target expression to specific tissues, scientists can study neoplastic progression in animals with defined genetic traits. Cooperativity between different genes can be tested by crossing different transgenic lines, and effects of environmental agents can also be included. The epidermis is an attractive tissue for such studies, as it is accessible and can stand as a model for epithelial cancers in general.

Transgenic mice overexpressing insulin-like growth factor 1 under the control of the keratin 5 (epidermal basal cell specific) promoter develop squamous papillomas, some of which convert to carcinomas (DiGiovanni *et al*., 2000), thus illustrating at least two of the major phases of tumour development, although metastases were not reported. In another model, mutant p53 was targeted to the epidermis of transgenic mice using a keratin -1 promoter, and their skin was treated with the classical promoter TPA. In control mice papillomas developed and regressed following cessation of treatment, whereas in the transgenic mice the papillomas progressed to metastatic carcinomas. The accelerated tumorigenesis was ascribed to increased genomic instability and abnormal amplification of centrosomes. Microarray analysis may be able to determine which genetic changes are linked to this ‘progressor’ phenotype. In another double (TGFα + v-fos) transgenic model generated by the same group, TPA also accelerated carcinoma progression, but additional genetic events were required and a high incidence of c-Ha-ras mutations (but not p53 mutations) were found in the papillomas. Interestingly, not all double transgenics show enhanced tumour formation, since in some systems loss of p53 can paradoxically be inhibitory, although this can be overcome by other specific genetic synergies or additional genetic changes (Wang *et al*., 2000) (**Table 1**).

Human papillomaviruses (HPVs) are DNA viruses that have been implicated as immortalizing agents which induce hyperplasia in cutaneous and mucosal epithelia, and they have a role in the development of certain human squamous cell carcinomas, e.g. cervical carcinoma. They act by interfering with the function of the *p53* and retinoblastoma (*Rb*) tumour-suppressor gene products since two of the viral proteins E6 and E7 can bind to these molecules. Recently, transgenic mouse models have been produced in which the HPV E6 or E7 genes have been delivered to specific epithelial cell populations to explore their role in more detail, and their interactions with other genes implicated in malignant progression (Eckert *et al*., 2000).

Table 1 Some transgenic models for epithelial neoplasia and preneoplasia

Transgene	Promoter	Result
IGF-1	Keratin 5	Skin papillomas and carcinomas
Mutuant *p53*	Keratin 1	Increased incidence of metastatic carcinomas with TPA
TGFα + v-*fos*	Keratin 1	Spontaneous papillomas, increased sensitivity to TPA promotion to carcinomas; endogenous H-ras activated
SV40 Tag	WAP	DCIS and progression to breast carcinomas in mice
TGFα	MMTV	Preneoplasias and cancers in rats
v-Ha-*ras* + c-*myc*	MMTV	Synergistic induction of breast carcinomas
Wnt-1	MMTV	Breast carcinomas, some metastases
Polyoma virus mT	MMTV	Rapid development of breast carcinomas
c-*neu*	MMTV	Breast carcinomas
SV40 Tag	C(3) 1	Prostate cancers, no metastases
SV 40 Tag	Probasin	Prostate cancers + metastases
c-*myc*	C(3)1	PIN but no carcinomas

Mice deficient in certain genes have also been produced to explore potential susceptibility factors. NAD (P)H:quinone oxireductase 1 (NQOI) is a flavoprotein involved in the detoxification of quinones and hence protects against their mutagenic effects. NQOI −/− mice treated with a classical two-stage carcinogenesis protocol (skin painting with benzo[*a*]pyrene followed by repeated applications of TPA) developed a much higher incidence of skin tumours than wild type mice, indicating that NQOI can act as an endogenous factor in protection against certain carcinogens. Similar studies with other candidate genes may one day explain why certain individuals exposed to high levels of carcinogens (e.g. cigarette smoke) may be more likely than others to develop cancers.

GASTROINTESTINAL TRACT TUMOURS

Colorectal cancers appear to arise as a result of a combination of activation of oncogenes and inactivation of tumour-suppressor genes – a scenario that is likely to be common to other tumour types also. It has been calculated that mutations in at least four or five genes are required to produce a fully malignant cancer. Although genetic changes may occur predominantly at ‘preferred’ stages of tumour evolution, it seems to be the accumulation of acquired defects rather than their chronological appearance that determines tumour behaviour (Cho and Volgelstein, 1992). It has long been known that certain malignancies are associated with particular chromosomal aberrations. More recently, nonrandom ‘loss of heterozygosity’ (i.e. allelic deletions of specific chromosomal regions) has proved a powerful pointer to the presence of tumour suppressor genes. This type of analysis led to the identification of the *DCC* gene (Deleted in Colon Carcinoma), whose product turns out to be a cell adhesion molecule important in intestinal cell differentiation. The gene (*APC*) for familial adenomatous polyposis (FAP) when mutated in the germ line is associated with the development of multiple benign colorectal tumours and similar somatic mutations have been found in sporadic colorectal cancers.

During the search for the *APC* gene, another gene was found which is also frequently mutated somatically in colorectal carcinomas – the *MCC* gene. Also, a mutational ‘hotspot’ has been found in the *p53* gene (which plays a crucial role in cell cycle regulation) in some human liver cancers that could be a consequence of exposure to the hepato-carcinogenic agent aflatoxin B_1 (Paraskeva and Williams, 1992). These observations show how classical epidemiology and cytogenetic analyses are being translated into an understanding at the molecular level of how our ‘nature and nurture,’ i.e. hereditary and environmental influences, may conspire to render us more or less susceptible to particular cancers. In addition, the accumulation of genetic alterations at a particular point in time may determine if the neoplasm is benign or malignant.

Spontaneously Arising Animal GI Tumours and Chemically Induced Tumours

Colorectal carcinomas are rare in rodents, although incidences of 30–40% have been recorded in rats with bacterial infections or acute colitis (both possible risk factors in humans). Occasional lymph node metastases, but not liver metastases, were found. Most chemical carcinogens, when administered to rodents by oral or other routes, induce tumours in the gut but also other sites. The most commonly used agents are dimethylhydrazine, dibenzanthracene, methylcholanthrene, heterocyclic amines, azoxymethane, *N*-methyl-*N*-nitro-*N*-nitrosoguanidine (MNNG) and *N*-methyl-*N*-nitrosurea (NMU). In all cases the incidences and rate of development are unpredictable, rendering the models unsuitable for general use. However, cell lines have been derived from some tumours, and where the primary hosts were an inbred

strain, the tumours can be passaged into genetically identical hosts (e.g. colon 26 in Balb/c mice, MCA38 in C57BL/6J mice, PROb in BDIX rats and CC531 in WAG/Rij rats). In this case, the growth of the tumours (and in some cases metastasis) can be more readily controlled, and the effects of various interventions more easily measured.

The Min Mouse Model

Min (Multiple Intestinal Neoplasia) is a mutant allele of the mouse *Apc* gene with a nonsense mutation at codon 850. Mice carrying this mutation, like humans, develop multiple intestinal adenomas (benign colorectal tumours), although their frequency is significantly influenced by 'modifier' genes such as *Mom-1* carried by different inbred strains, and can be increased by treatment with chemical carcinogens such as ethylnitrosourea (Moser *et al*., 1995). The female mice also develop breast cancers. This model provides an interesting system in which to study the role of *Apc* and modifier genes in initiation and progression of two major cancer types. A second model Apc1638N is also available in which tumour multiplicity is also modified by *Mom-1*, but several other X-ray responsive loci are also involved. In this strain the female animals develop ovarian carcinomas (van der Houven van Oordt *et al*., 1999).

Intervention and Therapy Studies

In addition to studying genetic modifiers, these models are also extremely useful for exploring the potential role of agents which might prevent cancers developing in genetically prone individuals, and possibly eventually in the general population if effective, safe dietary supplements could be found. The identification of upregulated levels of prostaglandins in colorectal and other cancers has led to the suggestion that inhibitors such as aspirin may be beneficial. Daily doses of aspirin reduce the numbers and sizes of polyps developing in the small intestine, but not the colon of Min mice. The Min mice have also been crossed with mice transgenic for the human carcinoembryonic antigen (CEA) gene to generate tumours expressing a human antigen. Such animals are ideal for testing CEA-directed immunotherapy (Thompson *et al*., 1997) since the target antigen is expressed on appropriate normal tissues in addition to the tumour, whereas in xenograft models only the tumours are 'positive.'

Xenograft Models

Many human GI tumours have been established as xenografts, either directly from clinical material or from cell lines, and some of the most commonly used are SW480, LS174T, KM12, HT29, HCT116 and COLO 320. In general, the tumours grow best when they are implanted in the correct anatomical location (the caecal wall) rather than subcutaneously. Although different congenitally immunodeprived hosts have been used, it is not clear if doubly or triply deficient mice are generally more susceptible to tumour growth and metastasis than the generally used athymic (T-cell deficient) mouse (**Table 2**).

Table 2 Colon carcinoma xenografts + metastasis models

Tumour	Injection site	Host	Metastases
WiDR	Spleen	Athymic mice	Liver
GX-39	Spleen	Athymic mice	Liver
GX-39	Intravenous	Athymic mice	Lung
LS174T	Intraportal vein	Athymic rats	Liver
LS174T	Spleen	Athymic mice	Liver
HT-29	Spleen	Athymic mice	Liver
HT-29	S.c.	SCID mice	Lung
TK3, TK-4, TK-9	Caecal wall	Athymic mice	Liver
HM-7	Spleen	SCID mice	Liver
C-1H	Spleen	SCID moce	Liver
LS180	Intraportal	Athymic mice	Liver, nodes

Metastasis Models

The commonest sites of metastasis of GI tumours are the liver and mesenteric lymph nodes. Tumours grown in the caecal wall are generally able to colonize these sites, while failing to spread from s.c. sites. This is technically demanding, and various 'short cuts' have been devised, notably the introduction of cells directly into the portal circulation via a mesenteric vein, or into the spleen, whence cells also rapidly enter the portal circulation. The spleen can be left *in situ* or removed a few minutes after injecting cells to avoid overgrowth of tumour at the primary site. These procedures have proved invaluable both for selecting cells with a higher propensity for metastasis by repeated cycling through the liver, and for testing therapeutic interventions or the effects of introducing genes into the cells by transfection prior to transplant.

LUNG AND BRONCHIAL CARCINOMA

In most Western countries, lung cancer is the main cause of cancer deaths, and even in Japan it has now overtaken stomach cancer as the most frequent malignant disease. In the last 50 years the death rates from cancer have increased by 200% in males and 300% in females; 90% of this is accounted for by smoking. The predominant type of cancer is adenocarcinoma, and it generally spreads to the contralateral lung, mediastinum, pleural cavity and lymph nodes.

In Vitro Models

In addition to a plethora of rodent and human tumour cell lines, a three-dimensional *in vitro* organ culture model for bronchial carcinoma has been designed using human bronchial mucosa organ cultures and lung cancer cell lines. This system offers opportunities for various *in vitro* investigations on therapeutic and diagnostic modalities of lung cancer, and invasion can also be quantified (Al-Batran *et al*., 1999).

Chemically Induced Tumours

Benign and malignant laryngeal cancers have been induced in hamsters and bronchogenic cancers in other species; urethane and PAHs are potent initiators of lung tumours in mice. Nicotine-derived nitrosamino ketone (NNK), a tobacco-specific carcinogen, induces lung cancers in rats by whichever route it is administered. These tumours have been examined for genetic alterations, and these have been found to include K-*ras* activation and loss of the *p16* and *Rb* tumour-suppressor genes. Alterations of these genes are also common in human lung cancer. Genetic linkage analyses to identify human lung cancer susceptibility genes is difficult owing to population heterogeneity and variable exposure to carcinogens. The mouse lung tumour model has become a valuable alternative in this search and also as a vehicle for chemoprevention studies.

Chemoprevention and Intervention Studies

A/J mice administered benzo[*a*]pyrene or the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl) butan-1-one (NNK) develop pulmonary adenomas. This model has proved useful in exploring the potential of compounds to inhibit carcinogenesis in the postinitiation period – which is what would be required for any realistic cancer intervention studies in humans. Myo-inositol and dexamethasone have been shown to reduce the formation of lung neoplasms in this model, and when coadministered there was an additive inhibitory effect.

Transplantable Models: Lewis Lung Carcinoma

One of the oldest and most widely used syngeneic transplantable tumours is the Lewis lung carcinoma (3LL) which grows in C57BL mice. It is highly malignant and metastatic. Although its natural primary site is the lung, it has been generally used as a model of lung metastasis. In common with many tumours, it grows well in the lungs following intravenous injection, and may also spread to the liver; it is also capable of metastasis from subcutaneous sites. It has featured in hundreds of publications testing agents as diverse as cytotoxic drugs, proteinase inhibitors, vaccines, vitamins, gene therapy and immunotherapy. Recently, it has also been used as a model of primary lung cancer with metastasis to mediastinal lymph nodes.

Xenograft Models

A wide variety of human lung carcinomas have been grown as xenografts in athymic mice, although generally at the s.c. site. Recently, intrapleural inoculation of cells was shown to result in a high incidence of lung tumours and lymphatic metastasis, compared with low take rates when cells were injected intravenously. An athymic rat lung cancer xenograft model has also been described, although since these animals have more residual immunity than athymic mice, preirradiation was necessary to enhance tumour growth. Metastasis to lymph nodes has been observed in these models. Although interesting, tumour growth quantitation is difficult and the animals can suffer from cachexia. With the development of green fluorescent protein- or genetically tagged cells, or polymerase chain reaction (PCR)-based methods to detect low numbers of human cells in mouse tissues, these problems may be resolved (Chishima *et al*., 1997).

BREAST CARCINOMA

Although the earlier detection of breast cancer through screening programmes and improvements in surgery and adjuvant therapy have improved survival rates, there are still around 15 000 deaths from this disease in the UK each year and 43 000 in the USA. This is due to the development of metastatic disease that is refractory to therapy. It is still the leading cause of death from malignant disease in Western women. The majority present with apparently localized disease, but many relapse owing to presence of occult micrometastases.

Changes within breast lobules, notably atypical lobular hyperplasia and lobular carcinoma *in situ* (LCIS) are generally recognized as risk markers for subsequent development of carcinoma. Ductal carcinoma *in situ* (DCIS) appears to predispose to cancer at the same site, whereas there is no evidence that invasive carcinomas arise directly from LCIS. However, it is still not clear whether all *in situ* carcinomas are inherently capable of evolving into malignant metastatic cancers, but in view of the heterogeneity observed within these lesions, it is unlikely. With an increasing number of cases of DCIS being identified by mammography, it is important for further analysis at the molecular level to try to define those most likely to progress.

Recently, a number of genetic alterations have been identified that are beginning to shed more light on the

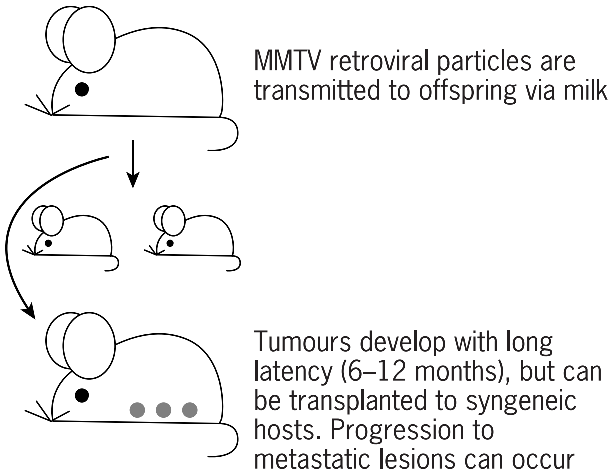
aetiology and progression of breast cancer. The identification of two genes (*BRCA1* and *-2*) associated with familial breast cancers may seem of little relevance to most cases, which are not genetically linked (see chapter on *Inherited Predispositions to Cancer*). However, by analogy with *APC* (see chapter on *Lower Gastrointestinal tract*) it is likely that their isolation and cloning will yield insights into the control of proliferation and differentiation of epithelial cells in the mammary gland. Secondly, overexpression of the proto-oncogene c-*erb*B-2 has been shown to be an independent prognostic indicator in breast cancer. Its presence in primary tumours has been linked with a higher probability of brain and visceral metastases, shorter survival and a poor response to therapy.

Breast cancer models will be discussed in other sections (see chapter on *Mammary Tumour Induction in Animals as a Model for Human Breast Cancer*), so in this chapter only a brief outline will be given. The reader is referred to two excellent special issues of journals: The *Journal of Mammary Gland Biology and Neoplasia* (1996) featured ‘Experimental Models of Development, Function and Neoplasia’ and a recent issue of *Oncogene* (2000) has been devoted entirely to ‘Mouse Models for Breast Cancer.’

Syngeneic Rat and Mouse Tumour Models

About 60–70 years ago, some of the first mouse tumour models were developed from inbred strains found to have a naturally high incidence of breast cancers. This was subsequently shown to be due to the transmission of the mouse mammary tumour virus (MMTV). The tumourigenicity is due in part to the activation of proto-oncogenes such as *Wnt-1* by proviral insertion into the host DNA. Although metastasis in the primary host is rare, when transplanted into syngeneic recipients, metastasis to lungs, nodes and viscera can be promoted by selection, producing useful experimental models, and the cells can also be readily grown *in vitro* (Barnett and Eccles, 1984). Human breast tumours show varying degrees of oestrogen responsiveness, and *in toto* the various MMTV mouse models span the spectrum from highly endocrine dependent (growing only during pregnancy and regressing following birth) to hormone-independent (**Figure 5**).

Rat mammary tumours may occur sporadically at low incidence and can be induced by the use of chemical carcinogens such as DMBA and NMU, and also by high doses of oestrogen (**Figure 6**). The primary hosts generally developed multiple lesions of varying malignancy and hormone dependence, and these proved useful for intervention studies with tamoxifen and other oestrogen antagonists. Metastases were again rare, and owing to the effects of the carcinogen, lung primary tumours



Oncogenesis is due to insertional activation, by the virus, of cellular oncogenes such as *Wnt-1*, *Wnt 3*, *Int2/FGF3*, *hst1/FGF4*, *Int-3*

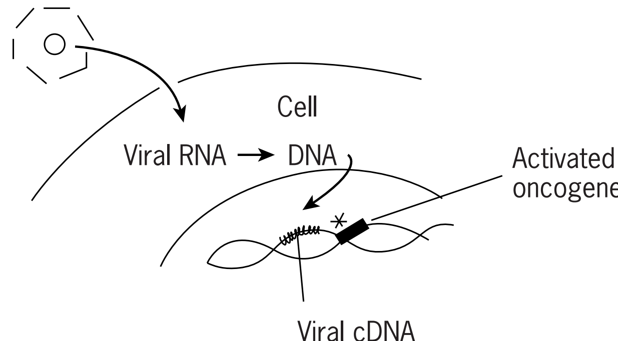


Figure 5 Mouse mammary tumour virus oncogenesis.

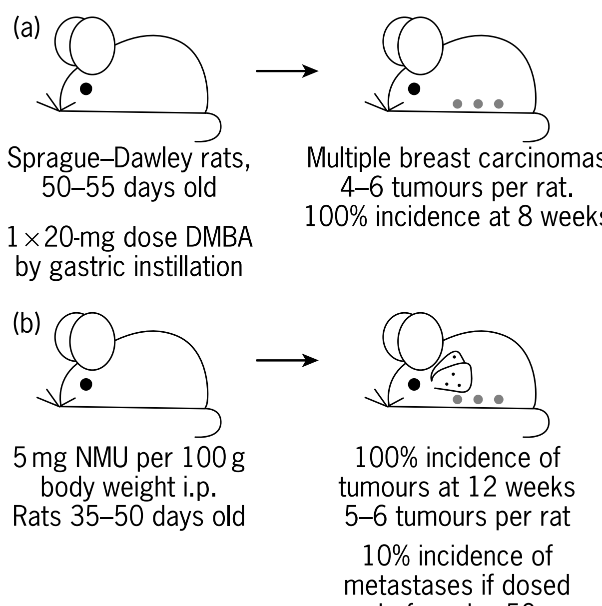


Figure 6 Rat mammary carcinogenesis. (a) ‘Huggins’ model; (b) NMU model.

were sometimes mistaken for secondary tumours. Again, these tumours could be transplanted and developed into useful models of hormone-dependent or independent growth and metastasis (Briand, 1983; Nicolson, 1988; Noble *et al*., 1975; Eccles *et al*., 1994). However, chemically induced tumours tend to have high levels of *ras* mutations, which are rare in human breast cancers, and since the human aetiology is complex (with a multiplicity of risk factors but no primary causative agent identified), it has been hard to model the natural history accurately. Now that genetic changes implicated in human breast cancer are emerging, it is possible to model these in rodents to study their effects.

Human Cell Lines and Xenografts

There are a relatively small number of commonly used human breast carcinoma cell lines, but these have been well characterized. The most popular is MCF-7, which is also reliably oestrogen sensitive, and generally only grows as a xenograft if the athymic mice (which have naturally low oestrogen levels) are hormonally supplemented. Other common cell lines are the MD Anderson (MDA) series, some of which (e.g. MDA MB 435) will metastasize from mammary fat pads in athymic mice. Breast carcinomas are not particularly easy to grow from clinical material (the success rate is around 30%), but this has been improved by the use of tissue fragments, co-inoculation with host fibroblasts and/or Matrigel and transfection of angiogenic cytokine genes into cell lines (Price *et al*., 1990). In addition, three-dimensional mammary primary culture models, as in other systems (e.g. skin and prostate), have proved invaluable for analysing important interactions controlling gene expression and differentiation (Ip and Darcy, 1996). Recently, breast cell lines which will metastasize to bone have been developed, which should prove very useful for studying this difficult clinical problem (Sung *et al*., 1997) (**Table 3**).

Transgenic Models and 'Knockouts'

Some excellent transgenic models of breast carcinoma have been developed which seem to mimic many aspects of human tumour biology including, in some cases, metastasis. The genes of interest (c-*erb*B-2, TGFα, c-*myc*) can be readily and specifically targetted to the mammary gland using breast specific promoters or by fusing genes to the long terminal repeat of the mouse mammary tumour virus. Both transgenic rat and mouse models have been established, and also a system in which cells from the mammary epithelium are removed, transfected with genes of interest using retroviruses and returned to the fat pad of the donor animal, avoiding the necessity for germ-line transduction (Edwards *et al*., 1996) (**Figure 7**). (See also chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*.) Some transgenic animals have been designed specifically to produce lesions, which mimic early-stage disease (DCIS) with the intention of studying determinants of cancer progression to invasive disease. Recently, targetted mutagenesis has been used to delete specifically the gene *Brca1* from the mouse mammary gland. *BRCA1* mutations are thought to be responsible for a significant proportion of familial breast cancers. The studies were complex because the homozygous 'knockout' mouse fails to survive, and the heterozygote does not develop tumours. However, if the second allele in these mice is mutated in the glands, they proceed to develop abnormalities and later mammary cancers, which, like their human counterparts, show increased genetic instability (Xu *et al*., 1999). This is still not a perfect model since in women loss of only one allele is

Table 3 Breast carcinoma xenografts

Tumour	Origin	Hormone sensitive[a]	ER/PR[a]	Metastases
MCF-7	Pleural effusion	Y	Y	Transfected sublines
MDA MD 231	Pleural effusion	N	N	Selected sublines, bone
MDA MB 361	Brain metastasis	+/−		
MDA MB 435	Pleural effusion	N	–	Selected sublines
MDA MB 468	Pleural effusion			
CAL51	Pleural effusion	N	N	Y (i.p.)
IIB-BR-G	Breast cancer	N	N	Y
ZR-75-1	Ascitic effusion	+/−	+/−	
SK-BR-3	Pleural effusion	N	N	
BT-20	Primary breast cancer	N	N	
T47D	Pleural effusion	Y	Y	
GI 101	Breast recurrence	N	N	Y
BT474	Primary breast cancer	Y	Y	Y
UISO-BCA-NMT-18	Breast cancer	Y		Y

[a]Y = yes; N = no; ER/PR, oestrogen receptor and progesterone receptor.

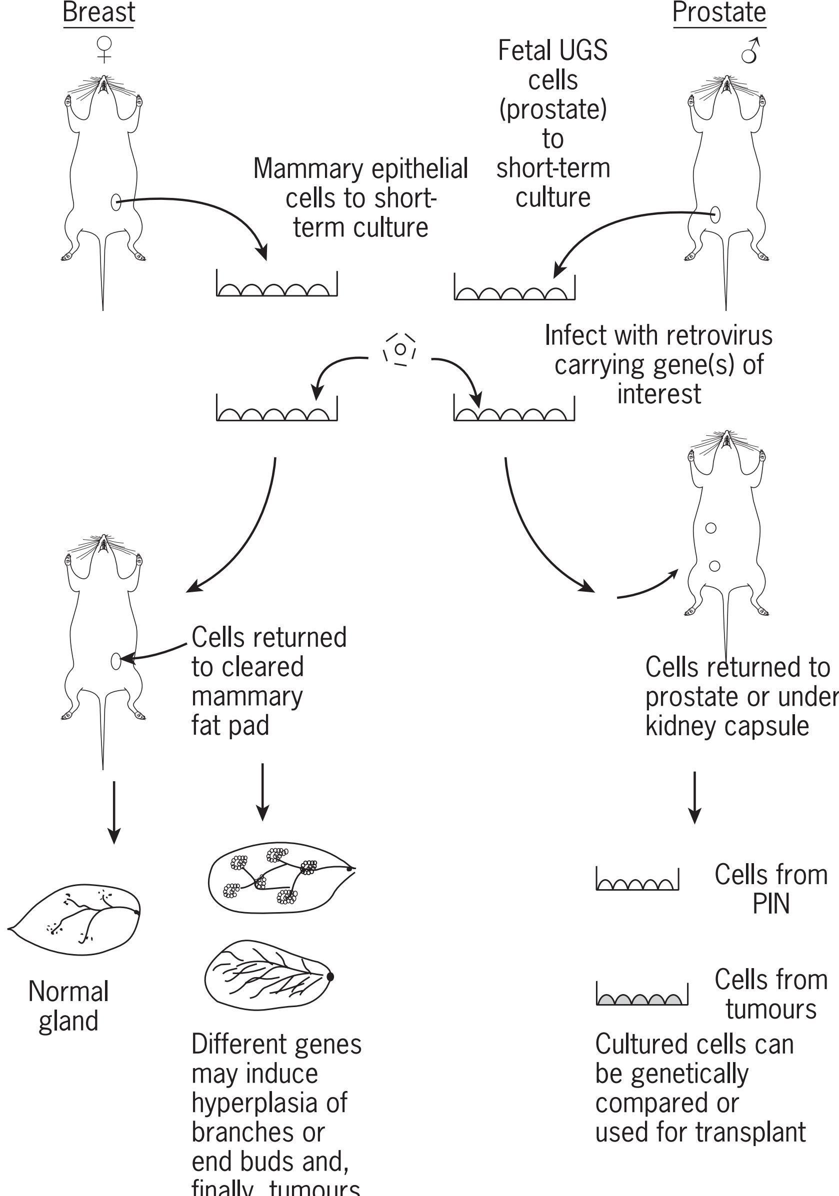


Figure 7 Examples of tissue reconstitution models for epithelial cancers.

sufficient to confer a very high risk of developing breast cancer.

PROSTATE CANCER

Prostate cancer has the highest incidence of any non-cutaneous malignancy in the West, and is the second highest cause of cancer deaths in males in Europe and the USA (Foster *et al.*, 1999). The incidence is greater than 30% in men over 50 years and 80% in men over 80 years. The incidence is also steadily rising, and at diagnosis around 75% of patients have locally extensive or metastatic disease. In 1978, Isaacs *et al.* listed the ideal criteria for an animal model of prostate cancer, which could equally well be applied to most epithelial cancers. The list was comprehensive and included features such as morphological and biochemical profiles similar to human disease, slow growth, appropropriate sites of metastasis (lymph nodes and bone), hormone and drug responsiveness similar to humans, progression towards hormone independence, etc. For practical purposes, the cells should also be transplantable in syngeneic hosts, be able to grow *in vitro*, and should express markers such as prostate-specific antigen (PSA).

Rodent Models

The best known series of rat prostate tumours (the 'Dunning' model) was derived from a spontaneous tumour (R3327) which arose in a Copenhagen male rat in 1961.

It has subsequently spawned multiple sublines of varying hormone dependence, metastatic capacity and marker profiles which together mimic most aspects of the human disease spectrum. It is probably the most widely used model available, and has been used to test antiandrogens, immunotherapy and vaccines, and many other novel therapies, and most recently has been used to discover important tumour- and metastasis-suppressor genes including *KAI1* and *MKK4/SEK1* (Tennant *et al*., 2000). It has also been used to develop a bone metastasis model. A second rat model system originated in a closed colony of Lobund–Wistar rats maintained by Dr Morris Pollard. This is the only rat strain that is inherently predisposed to spontaneous metastasizing hormone-sensitive prostate carcinomas. These animals have been widely used for intervention studies since they show progression from benign lesions (prostatic intraepithelial neoplasia (PIN) to fully malignant, hormone-independent tumours (Bostwick *et al*., 2000).

Xenograft Models

To date, about 25 xenograft models of human prostate cancer have been described, although the most commonly used are the PC3 (M), DU145 and LNCaP since these (unlike some other xenografts) also grow *in vitro*; of these only the latter is androgen sensitive (**Table 4**). Human prostatic carcinoma cells are generally very slow growing, and this has hampered their establishment as models, with only about 5% of implants yielding progressively growing, transplantable tumours (van Weerden and Rominj, 2000). Again, with the appreciation of the importance of orthotopic sites of growth, models which, when grown in the prostate, will disseminate to lymph nodes and other sites including bone have now been developed (Pettaway *et al*., 1996; Zhau *et al*., 2000).

Transgenic and Reconstitution Models

Several transgenic models have been developed using oncogenes such as those derived from the SV40 virus and polyoma virus, *ras* and *myc*. Prostate-specific gene expression has been achieved using regulatory elements of the rat probasin gene, the PSA antigen and the androgen-responsive rat *C3* gene (reviewed in Green *et al*., 1998). One of the most useful models has been the TRAMP (transgenic adenocarcinoma mouse prostate) in which rat probasin promoter elements drive expression of SV40

Table 4 Prostate carcinoma xenograft models

Cell	Origin	Hormone sensitive[a]	Grows *in vitro*[a]	PSA[a]	Metastases[a]
LNCaP	Node	Y	Y	Y	Y selected sublines
ARCaP	Ascites	Y	Y	N	Y
LuCaP 23.1	Node	Y	Y	Y	
LuCaP 23.8	Node	Y	Y	Y	
LuCaP 23.12	Liver	Y	Y	Y	
DU145	Brain	N	Y	N	
PC3 (M)	Bone	N	Y	N	Y selected sublines
PC82	Prostate	Y	N	Y	
PC-133	Bone	N	N	N	
PC-135	Prostate	N	N	N	
PC-295	Node	Y	N	Y	
PC-310	Prostate	Y	N	Y	
PC-324	Prostate	N	N	N	
PC-329	Prostate	Y	N	Y	
PC-339	Prostate	N	N	N	
PC-346	Prostate	Y	N	Y	
PC-374	Skin	Y	N	Y	
MDA PCa-31	Liver		N	Y	
MDA PCa-40	Liver	N	N	N	
MDA PCa-43	Adrenal		N	Y	
MDA PCa-44	Skin	N	N	N	
LAPC-3	Prostate	N	N	Y	
LAPC-4	Node	Y	N	Y	
LAPC-9	Bone	Y	N	Y	
CWR21	Prostate		N		
CWR22	Prostate	Y	N	Y	
CWR 31	Prostate		N		
CWR91	Prostate		N	Y	

[a]Y, yes; N, no.

large T antigen. The transgenic mice develop PIN by 8–12 weeks of age that progress to adenocarcinomas with distant metastases by 24–30 weeks of age. Cell lines have been derived from tumours at various stages of progression, and since the mice are inbred, the tumours can be transplanted to new hosts.

A mouse prostate reconstruction (MPR) model has also been developed over the last decade. Fetal urinogenital sinus (UGS) tissue is dissected and can be separated into epithelial and mesenchymal cells. These cells can then be infected in short-term cultures with retroviruses carrying genes of interest prior to reimplantation into syngeneic mice. The MPR model was one of the first transgenic organ systems to study the effects of specific genes or gene combinations within the context of the prostate *in vivo* (Thompson *et al*., 2000). The cells do not have to be cultured *in vitro*, eliminating the selection of cell types with the capacity to grow under these artificial conditions (**Figure 7**). The effects of many genes have been studied, including TGFβ, *ras* and *myc*. The cells are effectively genetically 'tagged' by the unique retroviral insertion sites, allowing clonal evolution to be followed as the tumours progress.

Cells can also be derived from genetically mutated mouse strains to measure their susceptibility to oncogenesis. Ras + myc transduction leads to the development of localized tumours in normal mice, but when UGS was derived from p53 'knockout' mice, the resulting tumours gave a very high incidence of bone and soft tissue metastases. Early-passage cell lines derived from primary tumours and metastases in the same animal can be analysed for differential gene expression, which might be related to progression. This approach has identified several important candidate genes, including caveolin 1, which has been found to be overexpressed in both mouse and human prostate carcinoma metastases and to potentiate survival of tumour cells in the absence of androgens.

CONCLUSIONS AND PERSPECTIVE

Over the last several decades, *in vitro* and *in vivo* models of increasing sophistication have been designed to assist in the elucidation of the nature of cancers. Epithelial cancers, because of their major social and economic importance, have been particularly well represented. While it is clear that no single model will suffice to encompass all aspects of the human cancer counterpart, there are now a great variety of systems which together provide most of the necessary tools for studies of basic cancer biology, the underlying genetic and molecular mechanisms, tumour–host interactions and therapeutic interventions. For most of the major cancer types, we have available both rodent and human tumour cell lines, and in some cases panels of cells ranging from normal through immortalized to fully malignant. Increasingly complex *in vitro* models are being devised whereby tumour cells can be made to interact with or invade normal tissues, bridging the *in vitro–in vivo* divide. In terms of *in vivo* models, there are a wide range of syngeneic and xenogeneic tumours available, and the increasing use of orthotopic (anatomically correct) sites of implantation, and of 'tagged' cells for sensitive detection of disseminated cells has enhanced our ability to study and treat metastatic disease.

Finally, the most dramatic breakthroughs have come from the development of genetically engineered mice in which oncogenes targeted to specific tissues, or knockout of suppressor genes has led to transgenic strains with high incidences of cancers. These can be used for primary intervention/prevention studies or, if on a suitable inbred genetic background, can yield new transplantable tumour models for studies of progression. Furthermore, the crossing of transgenic strains enables interactions between oncogenes, suppressor genes and modifier genes to be explored, since it is clear that full malignant conversion requires cooperation of multiple genetic and epigenetic factors. Once the models which best mimic human cancers have been defined (and this must await a fuller understanding of the latter), we will have ideal systems in which to test new therapies specifically targeting the underlying molecular abnormalities, and ultimately preventative treatments which will short-circuit the oncogenic process in those at risk due to hereditary factors or lifestyle.

REFERENCES

Ackhurst, J. J. and Balmain, A. (1999). Genetic events and the role of TGFβ in epithelial tumor progression. *Journal of Pathology*, **187**, 82–90.

Al-Batran, S. E., *et al*. (1999). Three dimensional *in vitro* co-cultivation of lung carcinoma cells with human bronchial organ culture as a model for bronchial carcinoma. *American Journal of Respiratory Cell and Molecular Biology*, **21**, 200–208.

Bachmeier, B. E., *et al*. (2000). Human keratinocyte cell lines differ in the expression of the collagenolytic matrix metalloproteinases-1, -8, and -13 and of TIMP-1. *Biological Chemistry*, **381**, 497–507.

Barnett, S. C. and Eccles, S. A. (1984). Studies of mammary carcinoma metastasis in a mouse model system I. Derivation and characterisation of cells with different metastatic properties during tumour progression *in vivo*. *Clinical and Experimental Metastasis*, **2**, 15–36.

Bergers, G., *et al*. (2000). Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature Cell Biology*, **2**, 737–744.

Birchmeier, C. and Birchmeier, W. (1993). Molecular aspects of mesenchymal–epithelial interactions. *Annual Review of Cell Biology*, **9**, 511–540.

Bostwick, D. G., *et al*. (2000). Prostatic intraepithelial neoplasia: animal models 2000. *The Prostate*, **43**, 286–294.

Boukamp, P., *et al*. (1999). Tumorigenic conversion of immortal human skin keratinocytes (HaCaT) by elevated temperature. *Oncogene*, **18**, 5638–5645.

Briand, P. (1983). Hormone-dependent mammary tumors in mice and rats as a model for human breast cancer (review). *Anticancer Research*, **3**, 273–282.

Chishima, T., *et al*. (1997). Metastatic patterns of lung cancer visualized live and in process by green fluorescent protein expression. *Clinical and Experimental Metastasis*, **15**, 547–552.

Cho, K. R. and Vogelstein, B. (1992). Genetic alterations in the adenoma–carcinoma sequence. *Cancer*, **70**, 1727–1731.

Christofori, G. and Semb, H. (1999). The role of the cell adhesion molecule E-cadherin as a tumor suppressor gene. *Trends in Biochemical Science*, **24**, 73–76.

DiGiovanni, J., *et al*. (2000). Constitutive expression of insulin-like growth factor-1 in epidermal basal cells of transgenic mice leads to spontaneous tumor promotion. *Cancer Research*, **60**, 1561–1570.

Dong, G., *et al*. (1997). Genes differentially expressed with malignant transformation and metastatic tumor progression of murine squamous cell carcinoma. *Journal of Cell Biochemistry, Supplement*, **28–29**, 90–100.

Eccles, S. A. (1987). Animal models of cancer. In: Waring, M. J. and Ponder, B. (eds), *Biology of Carcinogenesis.* 109–133 (MTP Press, Lancaster).

Eccles, S. A., *et al*. (1994). Preclinical models for the evaluation of targeted therapies of metastatic disease. *Cell Biophysics*, **24/25**, 279–291.

Eccles, S. A., *et al*. (1995). Significance of the c-erbB family of receptor tyrosine kinase in metastatic cancer and their potential as targets for immunotherapy. *Invasion and Metastasis*, **14**, 337–348.

Eccles, S. (2000). Cell biology of lymphatic metastasis: the potential role of c-*erb*B oncogene signalling. *Recent Results in Cancer Research*, **157**, 41–54.

Eckert, R. L., *et al*. (2000). Transgenic animal models of human papillomavirus-dependent disease. *International Journal of Oncology*, **16**, 853–870.

Edwards, P. A., *et al*. (1996). Genetic manipulation of mammary epithelium by transplantation. *Journal of Mammary Gland Biology and Neoplasia*, **1**, 75–89.

Fidler, I. J. and Ellis, L. M. (1994). The implications of angiogenesis for the biology and therapy of cancer metastasis. *Cell*, **79**, 185–188.

Foster, C. S., *et al*. (1999). The cellular and molecular basis of prostate cancer. *British Journal of Urology*, **83**, 171–194.

Fusenig, N. E. and Boukamp, P. (1998). Multiple stages and genetic alterations in immortalization, malignant transformation and tumor progression of human skin keratinocytes. *Molecular Carcinogenesis*, **23**, 144–158.

Green, J. E., *et al*. (1998). Workgroup 3: transgenic and reconstitution models of prostate cancer. *Prostate*, **36**, 59–63.

Hamadeh, H. K., *et al*. (1999). Arsenic disrupts cellular levels of p53 and mdm2: a potential mechanism of carcinogenesis. *Biochemical and Biophysical Research Communications*, **263**, 446–449.

Ip, M. M. and Darcy, K. M. (1996). Three-dimensional mammary primary culture model systems. *Journal of Mammary Gland Biology and Neoplasia*, **1**, 91–110.

Lenoir-Viale, M. C. (1996). Reconstruction of human skin epidermis *in vitro*. In: Shaw, A. J. (ed.), *Epithelial Cell Culture: A Practical Approach.* 179–200 (Oxford University Press, Oxford).

Loukinova, E., *et al*. (2000). Growth-regulated oncogene-alpha expression by murine squamous cell carcinoma promotes tumour growth, metastasis, leukocyte infiltration and angiogenesis by a host CXC receptor-2 dependent mechanism. *Oncogene*, **19**, 3477–3486.

Moser, A. R., *et al*. (1995). ApcMin: a mouse model for intestinal and mammary tumorigenesis. *European Journal of Cancer*, **31A**, 1061–1064.

Nicolson, G. L. (1988). Differential organ tissue adhesion, invasion and growth properties of metastatic rat mammary adenocarcinoma cells. *Breast Cancer Research and Treatment*, **12**, 167–176.

Noble, R. L., *et al*. (1975). Spontaneous and estrogen-produced tumors in Nb rats and their behaviour after transplantation. *Cancer Research*, **35**, 766–780.

Paraskeva, C. and Williams, A. C. (1992). Cell and molecular biology of gastrointestinal tract cancer. *Current Opinions in Oncology*, **4**, 707–713.

Pettaway, C. A., *et al*. (1996). Selection of highly metastatic variants of different human prostatic carcinomas using orthotopic implantation in nude mice. *Clinical Cancer Research*, **2**, 1627–1636.

Price, J., *et al*. (1990). Tumorigenicity and metastasis of human breast carcinoma cells in nude mice. *Cancer Research*, **50**, 717–721.

Shamsuddin, A. M. (1999). Metabolism and cellular functions of IP6: a review. *Anticancer Research*, **19**, 3733–3736.

Slaga, T. J., *et al*. (1996). The mouse skin carcinogenesis model. *Journal of Investigative Dermatology Symposium Proceedings*, **1**, 151–156.

Stenback, F., *et al*. (1999). The extracellular matrix in skin tumour development: a morphological study. *Journal of Cutaneous Pathology*, **27**, 327–328.

Sung, V., *et al*. (1997). Human breast cancer cell metastasis to long bone and soft organs of mice: a quantitative assay. *Clinical and Experimental Metastasis*, **15**, 173–182.

Tennant, T. R., *et al*. (2000). The Dunning model. *The Prostate*, **43**, 295–302.

Thompson, J. A., *et al*. (1997). Expression of transgenic carcinoembryonic antigen (CEA) in tumor-prone mice: an animal model for CEA-directed tumor immunotherapy. *International Journal of Cancer*, **72**, 197–202.

Thompson, T. C., *et al*. (2000). Mouse prostate reconstitution model system: a series of *in vivo* and *in vitro* models for benign and malignant prostatic disease. *The Prostate*, **43**, 248–254.

Vaccariello, M., *et al*. (1999). Cell interactions control the fate of malignant keratinocytes in an organotypic model of

early neoplasia. *Journal of Investigative Dermatology*, **113**, 384–391.

Valles, A. M., *et al*. (1991). Adhesion systems in embryonic epithelial to mesenchymal transformations and in cancer invasion and metastasis. In: Goldberg, I. D. (ed.), *Cell Motility Factors.* 17–34 (Birkhauser, Basel).

Van der Houven van Oordt, C. W., *et al*. (1999). The genetic background modifies the spontaneous and X-ray induced tumor spectrum in the Apc1638N mouse model. *Genes, Chromosomes and Cancer*, **24**, 191–198.

Van Weerden, W. M. and Rominj, J. C. (2000). Use of nude mouse xenograft models in prostate cancer research. *The Prostate*, **43**, 263–271.

Wang, X. J., *et al*. (2000). Co-operation between Ha-ras and fos or TGFα overcomes a paradox: tumor-inhibitory effect of p53 loss in transgenic mouse epidermis. *Molecular Carcinogenesis*, **29**, 67–75.

Xu, X., *et al*. (1999). Conditional knockout of BRCA1 in mammary epithelial cells results in blunted ductal morphogenesis and tumor formation. *Nature Genetics*, **22**, 37–43.

Zhau, H. E., *et al*. (2000). Establishment of human prostate carcinoma skeletal metastasis models. *Cancer*, **88**, 2995–3001.

FURTHER READING

Compagni, A. and Christofori, G. (2000). Recent advances in research on multistage carcinogenesis. *British Journal of Cancer*, **83**, 1–5.

Fodde, R., *et al*. (1999). Mechanisms of APC-driven tumorigenesis: lessons from mouse models. *Cytogenetics and Cell Genetics*, **86**, 105–11.

Greenhalgh, D. A., *et al*. (1996). Multistage epidermal carcinogenesis in transgenic mice. *Journal of Investigative Dermatology Symposium Proceedings*, **1**, 162–176.

Henninghausen, L. (guest ed.) (2000). Mouse models for breast cancer. *Oncogene*, **19**, No. 8 Review Issue.

Kobaek-Larsen, M., *et al*. (2000). Review of colorectal cancer and its metastases in rodent models: comparative aspects with those in humans. *Comparative Medicine*, **50**, 16–26.

Medina, C. and Daniel, C. (guest eds) (1996). Experimental models of development, function and neoplasia. *Journal of Mammary Gland Biology and Neoplasia*, **1**, 1–136.

Navone, N. M., *et al*. (1999). Model systems of prostate cancer: uses and limitations. *Cancer and Metastasis Reviews*, **17**, 361–371.

Shaw, A. J., (ed.) (1996). *Epithelial Cell Culture: A Practical Approach* (Oxford University Press, Oxford).

Yuspa, S. H. (1998). The pathogenesis of squamous cell cancer: lessons learned from studies of skin carcinogenesis. *Journal of Dermatological Science*, **17**, 1–7.

Website

http:mammary.nih.gov/Annapolis-guidelines (transgenic mouse models).

Chapter 71

Haematological Malignancies in Cancer Research

Jörg Haier and Garth L. Nicolson

Institute for Molecular Medicine, Huntington Beach, CA, USA

CONTENTS

- Introduction
- Haematopoiesis
- Role of some Tumour Viruses in Malignant Transformation
- Oncogenes and Tumour-suppressor Genes
- Models for Haematological Pathogenesis
- Principles of Haematological Tumour Models and their Application in Haematology
- Some Problems of Animal Models
- Clinical Relevance of Experimental Results

INTRODUCTION

Carcinomas account for approximately 90% of human cancers, whereas malignancies that arise from blood-forming cells and from cells of the immune system, leukaemias and lymphomas, account for about 8% of malignancies in humans. Acute leukaemia is the most common malignancy in children, affecting approximately 2000 children each year in the United States. The lymphomas are the seventh most common cause of death from cancer in industrialized countries, with more than 40 000 new cases annually in North America.

The leukaemias and lymphomas provide particularly good examples of the relationship between defective differentiation and malignancy. Since all of the different types of blood cells are derived from a common stem cell in the bone marrow, the descendants of these cells form the various lymphoid cell lineages by commitment to specific differentiation pathways. For example, some cells differentiate to form erythrocytes whereas others will differentiate to lymphocytes, platelets, macrophages, granulocytes, etc. Cells of each of these types undergo several rounds of division as they differentiate, but once they become fully differentiated, cell division ceases. In contrast, leukaemic or lymphoma cells fail to undergo terminal differentiation. Instead, they become arrested at early stages of maturation, and they retain their capacity for proliferation and continue to reproduce (**Figure 1**).

Compared with other cancers, haematological cancers are well disseminated near their inception, and they are capable of being transported throughout the body in the lymph and blood. Although they are disseminated, the leukaemias and lymphomas still show certain preferences of organ colonization and growth (**Figure 2**). Some of these malignancies, such as chronic lymphocytic B cell leukaemias and small-cell lymphocytic lymphomas, are found extensively in the liver, whereas others (acute lymphocytic T cell leukaemias, acute lymphocytic B cell leukaemias and prolymphocytic B cell leukaemias) are found predominantly in the bone marrow, and a few (polylymphocytic B cell leukaemias and immunoblastic sarcomas) are found in both compartments.

HAEMATOPOIESIS

There are basically three types of cellular elements present in the blood: red blood cells (erythrocytes), white blood cells (leucocytes), and platelets (thrombocytes). Since platelets do not contain a nucleus, sometimes they are not considered true cells. Each of these cells has its own function, differs morphologically from the others and has a lifespan characteristic for that particular cell type. In normal healthy people the destruction and production of lymphoid cells is balanced and, therefore, the number of cells present in the blood at any particular time is relatively constant. In the same way that ontogenesis occurs for the whole organism, the blood cells must also undergo certain stages of maturation to carry out their intended functions. In a healthy person, only the mature adult cells are found in the peripheral blood, whereas in many diseases, immature and abnormal forms of the cells may be present.

During human development there are changes in the sites of production, type and distribution of haematopoietic

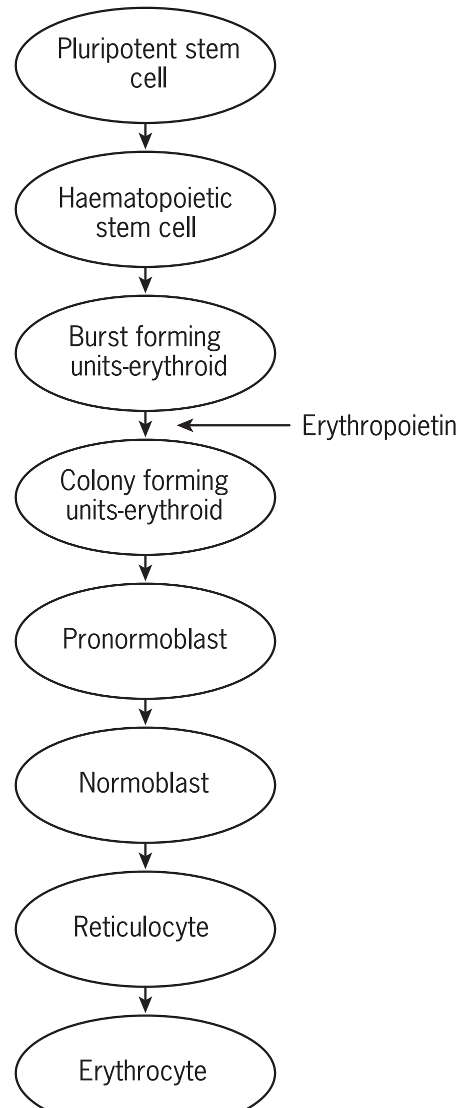


Figure 1 Stimulation of erythropoiesis by erythropoietin. Erythropoietin stimulates growth and differentiation of erythroid progenitor cells in the regulation of red cell production. Simplified scheme of erythroid precursors.

cells. In the fetus, haematopoiesis takes place at various intervals in the liver, spleen, thymus, bone marrow and lymph nodes. Within 2 weeks of embryonic life, primitive red blood cells are produced in the yolk sac. By the second month, granulocytes and megakaryocytes begin to appear and the liver and spleen become primary sites of cell development. Lymphocyte production starts at approximately the fourth month and monocytes are produced by the fifth month. At the end of the third trimenon the bone marrow becomes the primary organ of haematopoiesis. At birth, and continuing into adulthood, major blood cell production is confined to the bone marrow and is known as medullary haematopoiesis.

In children, the haematopoietic bone marrow (red marrow) is located in the flat bones of the skull, clavicle, sternum, ribs, vertebrae and pelvis and also in the long bones of the arms and legs. After puberty and during the adult life, the red marrow is normally confined to the flat bones (skull, clavicle, sternum, ribs, vertebrae, pelvis and the proximal ends of the long bones femur and humerus). The remaining marrow space is occupied by fat cells (yellow marrow), which can be replaced by haematopoietic cells under certain situations of intensive stimulation.

Haematopoiesis is tightly controlled by numerous cytokines/growth factors. In certain disease states, such as haemolytic anaemia, the bone marrow is unable to produce sufficient numbers of haematopoietic cells, and the secondary organs of haematopoiesis, such as liver and spleen, become reactivated as sites of extramedullary haematopoiesis. However, in haematological malignancies the primary and secondary organs of haematopoiesis may be infiltrated with malignant cells, resulting in a loss of the haematopoietic capacity.

Lineage-specific responses, such as increase of the number of erythrocytes during hypoxia or neutrophil production during bacterial infection, are governed by three generic categories of cellular competencies: (1) the capacity of haematopoietic stem cells to give rise to progenitors committed to a particular haematopoietic lineage; (2) the capacity of progenitors of a lineage to give rise to a differentiated progeny; and (3) the ability of the organism to perceive environmental stimuli by generating both extracellular and intracellular signals that either permit or repress these pathways of differentiation.

The production of blood cells in haematopoietic tissues involves the establishment, in early embryogenesis, of pluripotential haematopoietic stem cells. These cells are capable of a high degree of self-replication and have not yet committed themselves to the production of any specific lineage. Haematopoietic stem cells remain in residence in haematopoietic niches for the entire life of the host. They can also be removed from the original carrier and establish haematopoiesis for another lifetime in recipient organisms that have been lethally irradiated. Therefore, stem cells are (1) pluripotential and (2) can reconstitute the haematopoietic system because of their capacity for both self-renewal and production of committed progeny. Although committed progenitor cells cannot be distinguished morphologically, they can be quantified *in vitro*, because in semisolid medium (e.g. soft agar or methylcellulose) they form colonies containing morphologically recognizable, differentiated progeny cells. Committed progenitor cells are immediate precursors of blast forms, the earliest morphologically recognizable precursor cells of the stem cell lineage. The cell progeny of the blast forms, such as erythroblasts, can be identified morphologically and undergo predictable phenotypic changes as they progressively acquire attributes of the functional, terminally differentiated cell.

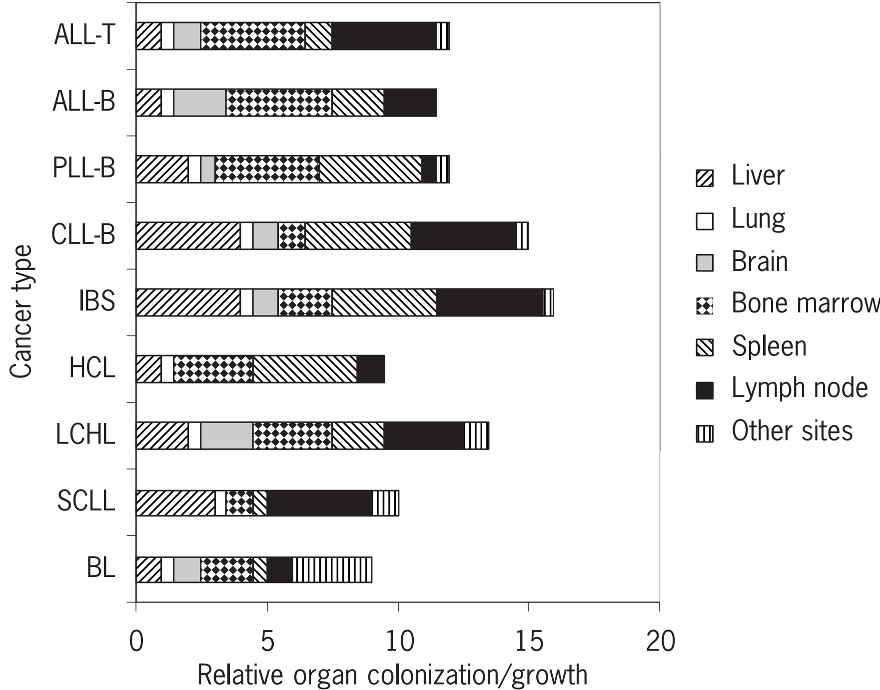


Figure 2 Organ preference of colonization/growth of selected human leukaemias and lymphomas. Organ colonization is scored as follows: 0, none; 0.5, organ sometimes involved; 1, organ involved; 2, organ moderately involved; 3, organ heavy involved; 4, organ replacement by disease. ALL-T, acute lymphocytic T cell leukaemia; ALL-B, acute lymphocytic B cell leukaemia; PLL-B, prolymphocytic B cell leukaemia; IBS, immunoblastic sarcoma; HCL, hairy cell leukaemia; LCHL, large cell histiocytic lymphoma; SCLL, small cell lymphocytic lymphoma; BL, Burkitt lymphoma. (From Nicolson, 1988, *Biochimica et Biophysica Acta*, **948**, 175–224.)

In the normal adult, all of these steps of replication and maturation take place exclusively in the bone marrow, an environment that provides growth factors, growth inhibitory factors, differentiation factors and adhesion molecules that aid in the retention, survival and growth of stem cells and progenitors in that particular niche. In the past 30 years, many factors involved in regulating growth and differentiation of haematopoietic cells in the bone marrow have been identified, usually in mouse models prior to their discovery in human systems. It should be recognized that murine models are highly relevant to human haematopoietic physiology. For example, virtually all *in vitro* assays for human haematopoietic progenitor cells were derived from similar methods established using murine cells. In fact, the first clear evidence that hematopoietic stem cells exist was derived from the seminal studies of Till and McCulloch (1961).

Current models of haematopoiesis propose that uncommitted pluripotent stem cells in the bone marrow maintain the ability to proliferate, reproduce themselves and differentiate. These stem cells are also able to repopulate the bone marrow after lethal damage and reconstitute normal haematopoiesis in the recipient organism. The first step of differentiation divides into lymphoid stem cells and haematopoietic stem cells, which are still considered multipotential cells. The haematopoietic stem cell is the precursor for red cells, granulocytes, monocytes and megakaryocytes (platelets). Progression of the multipotential cells to unipotential stem cells occurs through stimulation by various environmental factors within the different compartments of the organism. For example, the lymphocyte precursor or lymphoid stem cell leaves the bone marrow and undergoes differentiation into B lymphocytes in the lymph nodes or T lymphocytes in the thymus. In many cases the differentiation signals responsible have been identified. In the case of the primitive red blood cells that have been termed burst forming unit-erythroid cells that differentiate into colony-forming erythroid units, this occurs under the influence of the differentiation/growth factor erythropoietin (**Figure 1**). Similarly, colony-forming cells that are driven by specific differentiation/growth factors can be found for eosinophils, megakaryocytes, monocytes or granulocytes in response to Eos colony-stimulating factor (CSF), macrophage colony-stimulating factor (M-CSF), thrombopoietin, granulocyte colony-stimulating factor (G-CSF) or macrophage/granulocyte colony-stimulating factor (GM-CSF), respectively.

During the differentiation and maturation of haematopoietic stem cells, the cells undergo considerable changes in cell structure, size and shape. Progression from one stage to another does not appear to have discrete boundaries, but the expression of different cell surface molecules allows distinction between different precursor stages. In addition, metabolic pathways, cytoplasmic contents and nuclear chromatin structure differ between the different stages and are dependent on the committed lineage.

The lineage-specific differentiation of haematopoietic cells is determined to a large degree by specific haematopoietic growth factors and their cellular receptors. Once these haematopoietic growth factors bind to their receptors, specificity is mediated by the specific activation of their subsequent intracellular signalling pathways. This permits only a limited number of cells at certain differentiation stages to respond to the growth factors in the cellular environment. This ability to detect differentiation/growth factor signals can be introduced into non-haematopoietic cells by gene transfer of specific receptors (Roussel *et al*., 1987). Most growth factors interfere with more than one step of progenitor cell replication and cellular differentiation. In normal cells, replication and differentiation occur simultaneously. For example, the growth/differentiation factor G-CSF can stimulate growth of neutrophil progenitors and precursor cells, but this factor can also stimulate differentiation of these cells into mature neutrophils. Hence investigating and interpreting the actions of haematopoietic growth/differentiation factors can be difficult. Although there are additional limitations, such as unusual expression and availability of signalling proteins, our current understanding of the mechanisms of haematopoietic growth/differentiation factors has been achieved through the use of various models of transformed or leukaemic cell lines.

The control of haematopoiesis requires not only stimulation of cell proliferation and directed differentiation, but also tight regulation of its inhibition. Mitogenesis can be inhibited by various inhibitory factors, such as transforming growth factor-β (TGF-β). This antiproliferative factor interferes with tumour-suppressor gene products, such as Rb protein inhibiting the cell cycle progression from G_1 to S phase during the proliferation of progenitor cells (Laiho *et al*., 1990; Zentella *et al*., 1991). In addition, TGF-β can suppress the expression of growth factor receptors, such as the receptors for G-CSF, GM-CSF and interleukin-3 (IL-3) (Musso *et al*., 1990; D'Angeac *et al*., 1991). Other inhibitory factors of haematopoietic growth/differentiation include IL-4, IL-10, prostaglandins, interferon alpha and gamma (IFNα and IFNγ) and ferritin. (See also chapters on *Signalling by Cytokines* and *Signalling by TGF beta*.)

The ability of haematopoietic stem cells to undergo haematopoietic growth/differentiation in specific environmental compartments has been known for some time, but the mechanisms involved in this localization or 'homing' required a better understanding of adhesion molecules and their organ-specific distributions, mechanisms and activities. For example, it was known for some time that donor stem cells can be intravenously infused, but they repopulate only haematopoietic compartments in the bone marrow where they interact with stromal cells and then undergo further replication, proliferation and differentiation.

The organ-specific adhesive interactions involved in haematopoietic growth and differentiation are mediated by various types of adhesion molecules at the cell surface of haematopoietic cells, such as integrins and CD44, and their ligands on host organ stromal cells and their extracellular matrix that contain proteoglycans, heparan sulfate, fibronectin and vitronectin, among other molecules (Long, 1992). Adhesive interactions of circulating cells are discussed elsewhere in this section and the reader is referred to the relevant chapters (*Tumour Metastasis Models and Models for Tumour Cell–Endothelial Cell Interactions*). The expression of adhesion molecules at the haematopoietic cell surface is closely related to cell activation by various soluble growth factors or interleukins. Many interleukins or CSFs can modulate the expression of adhesion molecules and therefore interfere with the adhesive properties of stem cells at various stages of differentiation. This in turn, is an important point of hematopoietic regulation and lineage specific differentiation. In addition to the organ-specific expression of adhesion ligands, some CSFs or growth factors, such as GM-CSF and TGF-β1, can bind to extracellular matrix components resulting in localized regulatory activity of these factors (Mooradian *et al*., 1989).

ROLE OF SOME TUMOUR VIRUSES IN MALIGNANT TRANSFORMATION

Tumour viruses have been extremely important as models for cellular and molecular investigations of malignant cell transformation, and in some specific cases they appear to be involved in the formation and progression of human haematological malignancies (**Table 1**). The relatively small size of their genomes has allowed the discovery of various viral genes or oncogenes that have cellular homologues that play important roles in cell signalling and regulation of cell proliferation, survival and growth. Tumour viruses of six different families are capable of causing malignant transformation of mammalian cells, and one of the five DNA tumour viruses and the single known RNA tumour virus are important in haematological malignancies either as co-carcinogens or complete carcinogens. For example, the Epstein–Barr virus (EBV), which belongs to the Herpes virus group, is involved in the induction of Burkitt lymphomas or B cell lymphomas in AIDS patients, and the human T-cell lymphotropic virus (HTLV-1) is related to the appearance of adult T-cell

Table 1 Viruses associated with human malignancies

Virus group	Virus	Malignancy
DNA viruses		
Herpesviruses	Ebstein–Barr virus	African Burkitt lymphoma Nasopharyngeal carcinoma
Adenoviruses	Rarely in humans	
Hepadnaviruses	Hepatitis B virus	Hepatocellular carcinoma
Papovaviruses, Papillomaviruses	Human papillomaviruses (HPV 5, 8, 16, 18, 31, 33)	Squamous cell carcinoma Cervical carcinoma
Polyoma/SV40	Polyoma	Not in humans
RNA viruses		
Retroviruses	Abelson murine leukaemia virus (ALV)	Not in humans
	Moloney murine leukaemia virus (Mo-MLV)	Not in humans
	Human T cell lymphotropic virus type I (HTLV-1)	Adult T cell leukaemia
	Human T cell lymphotropic virus type II (HTLV-2)	Hairy T cell leukaemia
	Rous sarcoma virus (RSV)	Not in humans
	Avian erythroblastosis virus (AEV)	Not in humans

The retroviral group contains various examples of tumour viruses that do not directly cause malignancies in humans, but this group has contributed significantly to the discovery of various cellular oncogenes.

leukaemias. EBV is able to transform human B lymphocytes in culture, but the mechanism of this transformation remains incompletely understood. Multiple gene changes appear to be required to induce transformation, and the presence of a transformation-related viral gene may be only the first step in a multistep genetic process resulting in malignancy. Other viruses are also related to malignant transformation of haematological cells. For example, HTLV-1 is thought to be one of the causative agents of adult T cell lymphomas, whereas another related retrovirus (HTLV-2) is associated with the relatively rare formation of hairy T cell leukaemias. As a third virus from this group, HIV-1 does not appear to directly cause cancers, but it is likely to be responsible for the higher incidence of malignancies in AIDS patients, such as lymphomas or Kaposi sarcomas, possibly as a consequence of immunosuppression and loss of immune surveillance. Papillomaviruses, polyoma/SV40 viruses and adenoviruses appear to be important in certain malignancies, especially those involving certain epithelial cells, and can induce cell transformation by a common pathway that includes the alteration of cell cycle regulation by interfering with the key nuclear suppression proteins Rb and p53. (See also chapters on *Human DNA Tumour Viruses* and *RNA Viruses*.)

ONCOGENES AND TUMOUR-SUPPRESSOR GENES

Oncogenes and tumour-suppressor genes are large groups of genes that can together cause or contribute to the abnormal behaviour of malignant cells (**Table 2**). Suppressor genes were originally identified as genes whose expression was related to the suppression of tumour formation or progression. Elevated expression or uncontrolled activity of oncogene products or their normal cellular gene counterparts or proto-oncogene products can result in excess proliferation of tumour cells or inhibition of their differentiation to terminal phenotypes and cell death, whereas loss of function or expression during malignant transformation is characteristic for tumour-suppressor gene products. Oncogenes and suppressor genes related to malignant transformation have been identified in certain haematological malignancies, and some examples will be described in the following sections.

Oncogenes

The first well characterized example of oncogene activation by chromosome translocation was the c-*myc* oncogene involved in human Burkitt lymphomas and mouse plasmocytomas, which are malignancies of antibody-producing B lymphocytes. Both of these malignancies are characterized by chromosome translocations involving the genes that encode immunoglobins. For example, virtually all Burkitt lymphomas have translocations of a fragment of chromosome 8 to one of immunoglobulin gene loci on chromosome 2, 14 or 22. The fact that the immunoglobulin genes are actively expressed in these tumours suggested that the translocations activate a proto-oncogene from chromosome 8 by inserting it into the immunoglobulin loci. This possibility was investigated by analysis of tumour DNAs with probes for known oncogenes, leading to the finding that the c-*myc* proto-oncogene was at the chromosome 8 translocation breakpoint in Burkitt

Table 2 Oncogenes and tumour-suppressor genes involved in haematological malignancies and their proposed cellular functions

Type	Name	Malignancies in humans	Cellular function of gene product
Oncogenes	*abl*	Chronic myelogenous leukaemia Acute lymphocytic leukaemia	Nonreceptor protein tyrosine kinase
	Bcl-2	Follicular B cell lymphoma	Regulation of apoptosis
	DEK-CAN	Acute myelogenous leukaemia	Transcription factor
	E2A/pbx1	Acute lymphocytic leukaemia	Transcription factor
	ErbA	Avian erythroblastosis	Thyroid hormone receptor
	ErbB	Avian erythroblastosis	Receptor protein tyrosine kinase (growth factors)
	ets	Avian erythroblastosis	Transcription factor
	evi	Acute myelogenous leukaemia	Transcription factor
	fms	Stem cell leukaemia	Receptor protein tyrosine kinase
	Gp55 SFFV	Erythroleukaemia	Receptor ligand
	Hox gene family	Acute T cell lymphoma Myeloid leukaemias	Regulation of apoptosis Transcriptions factors
	IL-3	Pre B cell leukaemia	Receptor ligand
	lyl	Acute T cell lymphoma	Transcription factor
	mpl	Myeloproliferative leukaemia	Receptor protein tyrosine kinase
	Myb	Avian myelocytoblastosis	Transcription factor
	myc	Burkitt lymphoma	Transcription factor
	PML/RARα	Acute promyelocytic leukaemia	Retinoic acid receptor
	rasN	Acute myelogenous leukaemia Acute lymphocytic leukaemia	GTP-binding protein
	rel	Non-Hodgkin lymphomas	Transcription factor
	scl	Acute T cell lymphoma	Transcription factor
	tal	Acute T cell lymphoma	Transcription factor
	ttg	Acute T cell lymphoma	Transcription factor
Tumour-suppressor genes	*Rb*	Chronic megaloblastic leukaemia	Cdk4/cyclin D-dependent regulation of G_1 cell cycle phase
	INK4	Leukaemias	Cdk4/cyclin D regulation of G_1 cell cycle phase
		Lymphomas	
	p53	Leukaemias Lymphomas	Cdk cell cycle inhibitor p21 Apoptosis induction by DNA damage
	tan-1	Acute lymphocytic leukaemia	Receptor ligand

lymphomas. The translocations inserted c-*myc* into an immunoglobulin locus, where it was expressed in an unregulated manner. Such uncontrolled expression of the c-*myc* gene, which encodes a transcription factor normally induced in response to growth factor activation, is sufficient to drive cell proliferation and contribute to tumour development.

Translocations of other proto-oncogenes frequently result in rearrangement of coding sequences, leading to the formation of abnormal gene products. A prototype of this process is the translocation of the *abl* proto-oncogene from chromosome 9 to chromosome 22 in chronic myelogenous leukaemia forming the morphologically characteristic Philadelphia chromosome. This translocation leads to fusion of *abl* with its translocation partner, the *bcr* gene. The result of this gene fusion is production of a Bcr/Abl fusion protein in which the normal amino terminus of the Abl proto-oncogene protein has been replaced by Bcr amino acid sequences. The fusion of Bcr sequences results in aberrant activity and altered subcellular localization of the Abl protein tyrosine kinase, an important event leading to cell transformation.

Although most oncogenes stimulate cell proliferation, the oncogenic activity of some transcription factors results from inhibition of cell differentiation. For example, thyroid hormone and retinoic acid induce cell differentiation in a variety of cell types. These hormones diffuse through the plasma membrane and bind to intracellular receptors that act as transcriptional regulatory molecules. Mutated forms of the retinoic acid receptor (PML/RARalpha) can act as oncogene proteins in chicken erythroleukaemia and human acute promyelocytic leukaemia cells, respectively. The mutated oncogene receptors appear to interfere with the action of their normal proto-oncogene homologues,

thereby blocking cell differentiation and maintaining the leukaemic cells in an active proliferating state. As a consequence of this knowledge, high doses of retinoic acid can overcome the effect of the PML/RARalpha oncogene protein in acute promyelocytic leukaemia cells and induce their differentiation. Therefore, patients with this type of leukaemia can be treated effectively by administration of high doses of retinoic acid, which can induce differentiation and block continued cell proliferation. Such treatment with retinoic acid results in remission of the leukaemia in those patients, although this favourable response is temporary and patients eventually relapse with retinoic acid-resistant leukaemia. The use of this treatment is thus of substantial benefit in the initial treatment of acute promyelocytic leukaemia, providing the first example of a clinically useful drug targeted against an oncogene protein (**Figure 3**).

Another example of the role of differentiation in development of haematological malignancies involves the E2A gene products, E12 and E47, which are critical for proper early B cell development and commitment to the B cell lineage. Loss of E2A activity also results in a partial block at the earliest stage of T cell lineage development. The E2A gene products play a role early in thymocyte development that is similar to their function in B cell lineage determination. Furthermore, the lack of E2A gene products can result in the development of T cell malignancies, and it has been proposed that E2A inactivation is a common feature of a wide variety of human T cell proliferative disorders, including those involving the E2A heterodimeric protein partners tal-1 and lyl-1. The tal oncogene proteins are phosphorylated exclusively on serine and act as a substrate for the mitogen-activated signaling protein kinase ERK-1, resulting in alterations in DNA-binding affinity and modifications of their transcriptional activities (Yunis and Tanzer, 1993; Prasad and Brandt, 1997). Another example is the *HOX11* proto-oncogene that is normally expressed in embryogenesis where it directs the synthesis of the lymphoid cells in the spleen. In adult tissues, HOX11 expression is silenced by an unknown mechanism. Aberrant expression of HOX11 occurs in T cell acute lymphoblastic leukaemia, where it is thought to be involved in T cell immortalization. The *HOX11* gene may be repressed in normal T cells and erythroid cells by the action of negative elements which may be deleted or mutated in leukemia cells (Brake *et al*., 1998). Other members of the *HOX* gene family may be also involved in haematopoietic cell differentiation where they act as transcriptions factors (Crooks *et al*., 1999; Daga *et al*., 1999).

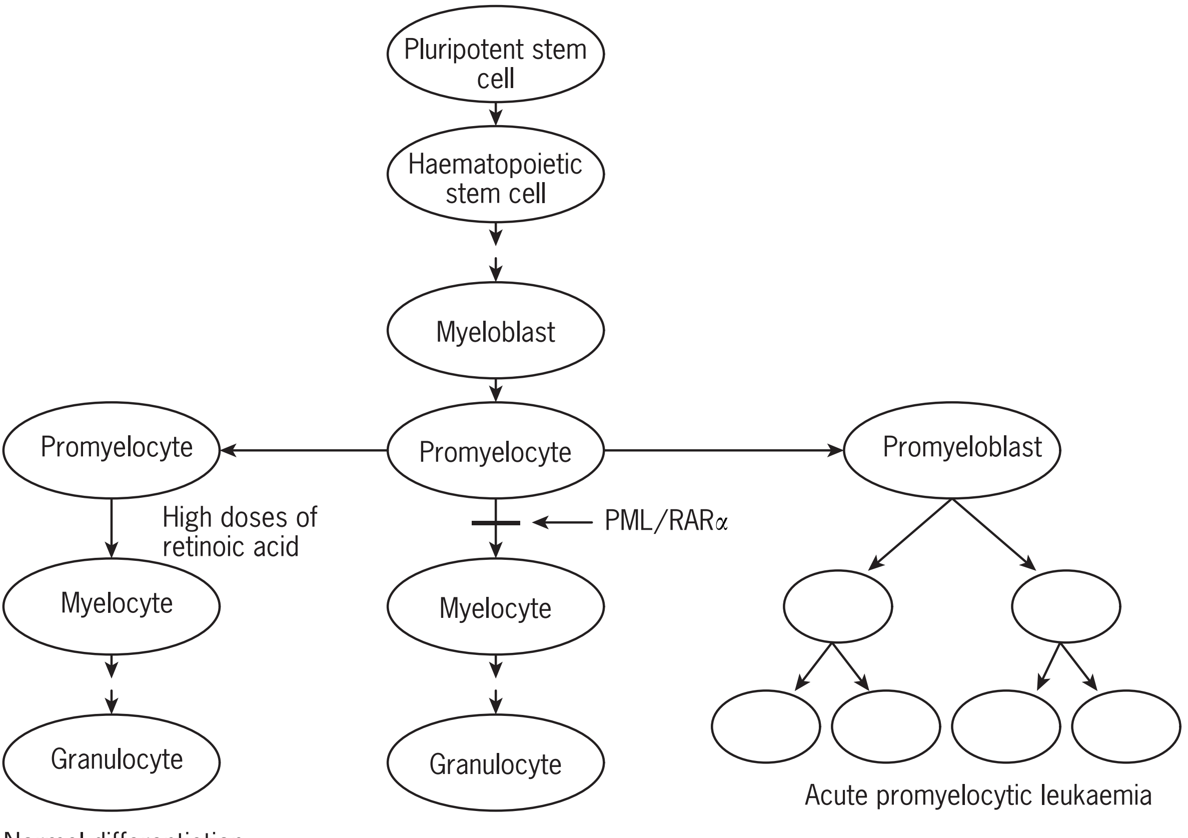


Figure 3 Pathogenesis of acute promyelocytic leukaemia and its treatment with high doses of retinoic acid.

Tumour-suppressor Genes

A different mechanism of abnormal tumour cell growth and survival is the loss of genes or gene products that regulate cell proliferation and programmed cell death (apoptosis). For example, the *bcl-2* oncogene product is involved in the regulation of apoptosis, but its generation by chromosome translocation with subsequent abnormal amplification results in protection against programmed cell death and formation of lymphomas. Abnormal Bcl-2 expression inhibits apoptosis and maintains cell survival leading to reduced environmental dependence on tumour cell growth and survival, such as that found in lymphomas that are growth stimulated by their organ factors found in their environment.

Inactivation or lost expression of tumour-suppressor genes results in the loss of negative regulators of cell proliferation or programmed cell death. The classical example for tumour-suppressor genes is the retinoblastoma *Rb* gene, which is involved in cell cycle regulation. Instead of acting only in haemotological malignancies, most of the tumour-suppressor genes are involved in many types of human tumours. For example, INK4, which regulates cell cycle progression from G_1 to S phase, is often lost in leukaemias and lymphomas, but also in melanomas, brain tumours and carcinomas. This gene encodes the Cdk inhibitor p16 that normally regulates Rb inactivation by the Cdk4/cyclin D kinase necessary for the release of G_1 cell cycle progression. Therefore, inactivation of INK4 results in increased activity of Cdk4/cyclin D complexes, hyperphosphorylation of Rb and loss of an important cell cycle restriction point. (See also chapter on *Regulation of the Cell Cycle*.)

DNA damage normally causes induction of the p53 gene product that activates transcription of the Cdk inhibitor p21. This inhibitor blocks cell cycle progression and DNA replication. Reduced expression or loss of function of p53 therefore prevents DNA damage-induced cell cycle arrest with a subsequent accumulation of mutations and genetic instability in these cells. In addition, p53 regulates DNA damage-induced apoptosis and inactivation of this tumour-suppressor gene product can result in resistance of transformed cells against radiation, chemotherapy, growth factor deprivation or hypoxia, treatments that normally induce apoptosis.

MODELS FOR HAEMATOLOGICAL PATHOGENESIS

Haematological malignancies provide several particularly good examples of the relationship between cell differentiation, uncontrolled proliferation and malignant transformation. As mentioned above, infections with the HTLV-1 virus can induce the formation of acute T cell leukaemia (ATL). Since the molecular mechanisms of the HTLV-1-induced ATL are such good examples of oncogenic transformation, the mechanisms of viral tumorigenesis and its investigation from clinical background to molecular pathways will be described in the following section.

Epidemiological, serological, experimental and molecular biological evidence supports the viral pathogenesis of ATL. This evidence is that: (1) the incidence of HTLV-1 infections and the occurrence of ATL are similar with regard to their geographical distribution; (2) infections with HTLV-1 can be found in most patients with ATL (Kozuru *et al*., 1996; Tamiya *et al*., 1996); (3) proviral HTLV-1 DNA is integrated into the genome of leukaemic cells; (4) clonal proliferation of infected leukaemic cells has been demonstrated (Yoshida *et al*., 1984); and (5) infection of T cells with HTLV-1 can immortalize these cells (Miyoshi *et al*., 1981; Yoshida *et al*., 1982). The integration of HTLV-1 proviruses into the genome of leukaemic cells appears to be random, which suggests a role of HTLV-1 in malignant transformation independent of the integration site of the provirus. In contrast, viral oncogenesis in carcinoma cells is normally associated with a close relationship of viral integration and activation of oncogenes or suppression of tumour-suppressor genes. The viral gene expression also results in a transcription of the *trans*-activation protein Tax that can initiate viral gene expression and replication. In addition, this protein can interfere with transcription of cellular genes, such as those for IL-2, IL-2 receptor α, GM-CSF, c-*fos* or c-*jun*. Many proteins with an increased Tax-induced expression are involved in the regulation of cellular proliferation. Furthermore, expression of Tax proteins can lead to immortalization of T cells in an IL-2-dependent matter (Grassman *et al*., 1989). However, additional genetic or epigenetic alterations and clonal selection appear to be required for leukaemogenic transformation and the clinical occurrence of ATL.

PRINCIPLES OF HAEMATOLOGICAL TUMOUR MODELS AND THEIR APPLICATION IN HAEMATOLOGY

Availability of Blood Cells

Blood represents a ‘renewable resource’ for obtaining human cells. Therefore, normal and malignant cells may often be obtained using relatively easy procedures used for the isolation and maintenance of these cells from peripheral blood. Large numbers of human lymphocytes (1–2 million per millilitre of blood) can be isolated by simple and inexpensive techniques. For example, lymphocyte preparations can be obtained by density-gradient centrifugation, further isolation of monocytes/macrophages utilizes their ability to adhere to plastic. In addition,

monoclonal antibodies bound to magnetic beads and recognizing specific surface molecules of haematological cells can be used to select enriched or depleted T or B cell subpopulations. The latter cells also allow the generation of continuously growing cell lines via retroviral transformation, using viruses such as Epstein–Barr virus.

Murine bone marrow cells are usually isolated from femoral bones using standardized preparation techniques. The ease of isolation and high yield of cells without loss of viability provide excellent cell cultures for experimental studies. Similarly, human bone marrow aspirates or preparations from bone after their surgical removal can be used, but the potential risk for viral infections has to be considered when using human tissues.

In addition to the isolation of primary cells from patients or animals with malignant haematological diseases, other techniques are available to obtain haematological cells for experimental investigations. For example, short-term culture of human lymphocytes can be established from peripheral blood (Hungerford *et al*., 1959). With a few modifications, this technique has become a standard procedure for human genetic analyses. However, these cells can be cultured for only a few days, and the resulting cell cultures are heterogeneous in their cellular characteristics. Therefore, primary cultures are often not suitable for experimental procedures. Alternatively, diploid stem cell and permanent cell lines can be established or they are available commercially (from the American Type Culture Collection (ATCC) and the European Collection of Animal Cell Cultures (ECACC). Such diploid nontransformed cells are more homogeneous and demonstrate stable proliferation for up to 50 passages in culture. Permanent cell lines can be obtained by viral transformation of untransformed cells, resulting in immortalized cells. In addition, specific genetic modifications have been introduced that result in mutational changes in oncogenes/tumour-suppressor genes, knockouts of these genes or specific retroviral transformation by gene insertion. The disadvantage of these cells is that they often lose characteristics of normal cellular behaviour, such as cell growth and differentiation properties, availability of metabolic pathways, and expression of surface molecules, among other changes. (See also chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*; *Models for Tumour Growth and Differentiation*.)

Cell Culture Techniques

Modern haematological research requires the use of cell cultures that exist as more or less complex cellular systems. Various techniques have been established for the establishment and maintenance of these cell cultures (see chapter on *Basic Tissue Culture in Cancer Research*). Although principles of mammalian cell culture can be transferred to haematological cells, some specific problems have to be considered because of the specific nature and behaviour of these cells. First, as described above, the formation of leukaemias and lymphomas is closely related to disturbances in cell differentiation and maturation, which are strongly dependent on microenvironmental factors, such as the presence of cytokines or growth factors. For example, addition of serum (horse or fetal bovine) to cell cultures can result in variation of cellular behaviour since serum contains numerous growth factors with significant variations among different batches. Use of serum-free culture conditions in colony-forming assays can support colony formation and overcome these difficulties. In addition, specific conditions usually provided by the bone marrow matrix may influence proliferative properties and differentiation of blood progenitor cells. Therefore, culture conditions have to be adapted to the requirements of the specific cells. Second, blood provides specific conditions of circulating fluid where haematological cells occur as single cells in a complex solution. Consequently, culture of haematological cells frequently requires nonadherent culture techniques and cell suspensions.

Although this chapter cannot do justice to some of the specific technical problems of cell culture, the following will discuss some of the specific problems in haematological cell culture. Cell culture in suspension is required for cells that do not show cell adhesion to plastic surfaces or matrix components under normal culture conditions. This technique can be used for tumour cells that do not require or have reduced requirements for cell adhesion for their survival and growth. Culture as single cells in suspension has some advantages compared to the culture of adherent cells. The suspension cell cultures can be easily subcultured by simple dilution without the requirement of degradative enzymes, such as trypsin. As long as culture conditions (fresh medium) provide sufficient nutrients and toxic metabolic waste products are removed so that they do not accumulate, these cultures more or less grow continuously without a lag time after their dilution into fresh medium. In addition, homogeneous suspensions are available in large culture volumes and high cell densities.

Short-term culture of peripheral lymphocytes is a standard technique in human genetics and provides an almost unlimited source of diagnostic material for mutational analysis, investigation of chromosomal abnormalities and identification of hereditary alterations in genome structure. However, since the cells originate from cancer patients, special considerations of biological safety are required to prevent infection with viral agents, such as HBV, HCV or HIV-1. Isolation of peripheral lymphocytes is based on density gradient centrifugation (usually in a density gradient of Ficoll) of heparinized whole blood. Subsequently, addition of mitogens stimulates cell proliferation to obtain large numbers of lymphocytes for chromosomal analysis by staining techniques, *in situ* hybridization or isolation and direct sequencing of DNA.

Frequently, haematological cells need specific conditions for cell proliferation or maintenance of their differentiation properties. This can be obtained only by addition of other cells that normally provide the specific microenvironmental signals required by the haematological cells. These feeder cell layers frequently consist of embryonic fibroblasts, macrophages or thymocytes that have been inactivated for cell proliferation but not survival by radiation or mutagen treatment, leaving feeder cells that contain the necessary metabolic activity. These feeder cells secrete matrix proteins, growth factors and/or other stimulating factors, and they provide specific surfaces for cell–cell contacts and cell adhesion. In addition, bone marrow matrix has been used for haematological cell cultures.

In contrast to cellular differentiation found in the complex environment of *in vivo* models of cellular differentiation, in haematological malignancies analysis of single cells or cell lines permits the identification of specific cellular interactions and allows for their manipulation or modulation. The advantage of these model systems for cellular differentiation is at the same time one of their disadvantages. Neglecting the complex interactions of haematological cells with their microenvironment, such as bone marrow matrix and the presence of specific growth factors and cytokines, the relevance of *in vitro* results to human pathobiology may be limited. The haematological development from totipotent cells to pluripotent stem cells and finally to terminal differentiated blood cells is characterized by the loss of proliferative activity.

Various *in vitro* model systems have been developed for studying different steps of haematological cell differentiation, where the cellular differentiation can be induced by physiological or artificial substances and cell–cell or cell–matrix interactions. For example, growth factors, such as TGF-β, EGF or HGF, hormones, such as corticoids, glucagon or thyroxin, vitamins, such as all-*trans*-retinoic acid, or various inorganic ions can be used for induction of cellular differentiation (Osborne *et al*., 1982; Nakamura *et al*., 1989; Massague, 1990; Bonicinelli *et al*., 1991). In addition, induction of cell differentiation can be achieved by homologous or heterologous cell–cell interactions. Homologous cell–cell interactions require a high cell density to permit efficient cell–cell communication. However, for the investigation of haematological cell differentiation, heterologous stimulation by fibroblasts and thymocytes, among other cell types, is more frequently used. Current *in vitro* models in haematology have been based on the development of major experimental techniques. Earlier Burgess and Metcalf (1980), Metcalf (1989) and Cross and Dexter (1991) developed various cloning methods for haematopoietic cells based on semisolid agar culture containing colony-stimulating growth factors. The usage of feeder layer-cell culture was introduced by Dexter *et al*. (1977). Lymphoblastoid (such as HL-60 cells) and erythroid cell lines (such as Friend–erythroleukaemia cells (MEL)) are widely used examples of established progenitor cells for studying haematopoiesis (Rossi and Friend, 1967). Depending on the stimulating agents, their cellular differentiation into various terminal differentiated blood cells can be induced and studied (**Table 3**).

Table 3 Cellular differentiation of haematopoietic progenitor cell lines depending on the differentiation inducing agent

Cell line	Inducing agent	Terminal differentiated cell type
Friend-erythroleukaemia (MEL)	DMSO	Erythrocytes
HL-60	DMSO	Granulocytes
	Vitamin D_3	Monocytes
	Phorbol ester	Monocytes
	GM-CSF	Eosinophils

(From Eisen *et al*., 1977, Collins *et al*., 1977.)

Animal Models for Haematological Malignancies

Animal models based on knockout or transgenic technology are now widely used in basic and applied haematological research (see chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*). The evaluation of animal models for haematological malignancies has to take into account various aspects of host and donor organisms as well as the target process(es) in the malignant tumour cells that is under investigation with a particular model. For example, for investigating the metastatic properties of leukaemia cells, the following aspects of the metastatic cascade have to be considered: (1) the natural site of leukaemogenesis, which in most cases is the bone marrow; (2) the exit site of the leukaemia cells from the bone marrow; (3) the effects of natural immune response of the host in the bone marrow and during the metastatic spread in the circulation; (4) the homing of leukaemic cells to other organs, including the bone marrow; and (5) the growth and elimination of leukaemic cells in the target organ(s).

The establishment of animal models for haematological malignancies currently relies on transgenic technology and retroviral infection of target cells. As technological advances become available they offer an opportunity to inactivate or modulate various endogenous genes in mice (or rats), leading to new possibilities in haematological research on the various steps of transformation and progression. In general, perturbations of gene expression in transgenic systems often lead to measurable alterations of physiological functions within the animal model. The technical details of this are described elsewhere in this section (see chapter on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*), but for an understanding of haematological models we will

discuss a few important features that have to be considered in the evaluation of these models. Little is known about the stepwise deregulation of specific genes leading to haematological malignancy, but aberrant gene expression in transgenic mice has been correlated with the formation of various haematological malignancies. For the investigation of these processes, specific target genes are usually dysregulated using transgenic modifications. However, frequently transgenic dysregulation of intracellular signalling pathways can result in activation of alternative pathways or interfere with signalling that is not directly involved in the regulation of the target gene (Sheppard *et al*., 1998). This can lead to cellular behaviour normally not found in evolving human malignancies. Therefore, the careful characterization of transgenic or infected cell clones has to include evaluation of morphological criteria, histopathological appearance, cytoskeletal architecture, expression of cell surface molecules, clonality and immunological properties of the resulting cells. For example, the expression of p210Bcr/Abl in haematopoietic progenitor cells in transgenic mice can contribute to two clinically distinct haematopoietic malignancies, CML and ALL, indicating that this transgenic system provides a novel transgenic model for human Philadelphia chromosome-positive leukaemias. In addition, proliferation of blast cells can occur in the lung or other secondary organs, which might represent an extramedullar blast crisis comparable to the natural clinical occurrence in humans (Honda *et al*., 1998).

Chromosomal translocations are found as primary or secondary mutations in patients with leukaemias. As a result of these translocations, fusion can occur between various genes that encode transcription factors or other potential oncogenes/tumour-suppressor genes. The role of these fusion genes in leukaemogenesis is often unknown. For functional studies of the fusion genes *in vivo*, they can also be expressed in mouse bone marrow cells via retroviral transduction (Cuenco *et al*., 2000). The clinical features of these animals are important, and such morphological investigations can increase our knowledge about the role of affected genes in haematological malignancies.

Multipotential progenitor and stem cells occur at a low frequency in haematopoietic tissues. Therefore, it is often difficult to obtain sufficient numbers of cells for experimental studies that delve into the molecular events involved in the regulation of lineage-affiliated genes within these multipotent cells. As an example to solve this problem, the myeloproliferative leukaemia virus (MPLV) has been used to generate a phenotypically diverse array of haematopoietic progenitors from adult mouse bone marrow and embryonic blood. These cells can be expanded to perform further investigations previously impossible using analogous primary cells (Thompson *et al*., 1999).

In some cases it is preferable to use animal models of human haematological malignancies for some studies. For example, rabbit models of lymphoma with specific chromosomal abnormalities have proved to be very useful for clarifying the role of EBV in human EBV-associated lymphoma and provide a means for studying prophylactic and therapeutic regimens. Another example is the Si-IIA-EBV or Cyno-EBV that can be inoculated intravenously, orally, subcutaneously or intraperitoneally and can subsequently induce T cell lymphomas in rabbits in about 2–5 months (Hayashi and Akagi, 2000).

SOME PROBLEMS OF ANIMAL MODELS

Although animal models of human haematological malignancies have proved to be useful, especially for *in vitro* studies, there are also some limitations that must be considered. Long-established leukaemia and lymphoma cell lines from mice or other animals and their sublines are heterogeneous in their malignancies, resulting in variations in their behaviours. For example, cells of high malignant potential are often less susceptible to lysis by the complement-dependent natural immune response than less malignant cells. The more malignant cells also have higher affinities for adhesion within the bone marrow and are less dependent on the microenvironment for their survival, differentiation and proliferation. This phenotypic diversity or heterogeneity is not specific to haematological malignancies, but because of the single cell characteristic of tumour cell–host organ interactions this phenomenon requires specific consideration when using haematological malignancies.

Host factors influencing metastasis formation have been attracted research attention for several decades. These studies have shown that there are important systemic and local environmental factors, such as immune resistance mechanisms that impact on tumour properties. For example, the antitumour T cell response, natural killer cells and macrophages can influence the metastatic process. It has been known for some time that malignant haematological cells are more susceptible to host immune mechanisms than cells from carcinomas. Therefore, many models of haematological malignancies use immunodeficient animals as the host. Furthermore, coagulation factors and platelets interact with circulating cells, particularly in the microcirculation. As discussed above, growth factors, cytokines and/or hormones are involved in the regulation of differentiation, proliferation and survival of leukaemia and lymphoma cells. Many of these factors are also very important regulators of host defence, inflammation and wound repair mechanisms. Thus these environmental factors can result in the selection of malignant cells that escape the above environmental controls. The complexity of the relationship between tumour cells and animal hosts is further discussed elsewhere in this section. (See also chapter on *Human Tumours in Animal Hosts*.)

Major problems for the investigation of tumour immunology in haematological malignancies are the reliability

and reproducibility of antitumour effects in spontaneously arising and nonimmunogenic haematological malignancies (see chapter on *Models for Immunotherapy and Cancer Vaccines*). For example, B cell malignancy-derived Ig may be considered a model tumour antigen for vaccine development. However, as a nonimmunogenic self-antigen, it must also be first rendered immunogenic by chemical or genetic fusion to carriers that enable the induction of protective antitumour immunity in murine tumour models. For example, multiple injections of intact irradiated murine B cell leukaemia/lymphoma cells (BCL1) has been used to trigger a dose-dependent antitumour immune response in syngeneic mice. However, Newcastle disease virus (NDV) infection or transfer of cytokine genes into these cells for investigation of their role immune response did not affect tumorigenicity or immunogenicity of BCL1 cells (Morecki *et al*. 1998). Administration of these vaccines as fusion proteins or the use of naked DNA vaccines may allow efficient targeting of antigen-presenting cells *in vivo* (Biragyn and Kwak, 1999).

In summary, the very complex interactions of haematological cells with different physiological processes, such as immune response, host defence and inflammation, among others, require consideration of many factors influencing the experimental systems. Therefore, the use of experimental systems requires careful investigation of these factors for reproducible and reliable results. Moreover, these complex interactions are frequently neglected in experimental design of *in vitro* investigations. Finally, the translation of these observations to *in vivo* systems often yields unexpected results.

CLINICAL RELEVANCE OF EXPERIMENTAL RESULTS

Molecular analysis of malignant transformation in haematological malignancies has led to significant contributions in clinical treatment strategies. For example, as a consequence of the molecular understanding of leukaemogenesis in patients with acute promyelocytic leukaemia (APL) the development of differentiation-inducing therapy by all-*trans*-retinoic acid (RA) is now a standard clinical procedure. Nearly all patients achieve complete remission by the treatment of all-*trans*-RA; however, clinical remissions are usually of brief duration, and these patients often subsequently develop RA-resistant disease. The mechanisms of RA resistance in APL cells are poorly understood and most clinical approaches have not been successful in overcoming RA resistance. Therefore, further understanding of the complex regulatory mechanisms of progenitor cell maturation and differentiation will be necessary for successful long-lasting treatment. Using this example, studying RA-resistant APL model systems *in vitro* and *in vivo* may be useful for investigating the molecular aspects of blocking leukaemic cell differentiation and the molecular mechanisms of RA resistance (Kizaki *et al*., 1999).

REFERENCES

Biragyn, A. and Kwak, L. W. (1999). B-cell malignancies as a model for cancer vaccines: from prototype protein to next generation genetic chemokine fusions. *Immunological Reviews*, **170**, 115–126.

Bonicinelli, E., *et al*. (1991). Hox gene activation by retinoic acid. *Trends in Genetics*, **7**, 329–334.

Brake, R. L., *et al*. (1998). Multiple negative elements contribute to repression of the HOX11 proto-oncogene. *Oncogene*, **17**, 1787–1795.

Brown, B. A. (1993). *Hematology. Principles and Procedures* (Lea & Febiger, Philadelphia).

Burgess, A. W. and Metcalf, D. (1980). The nature and action of granulocyte-macrophage colony-stimulating factors. *Blood*, **56**, 947–958.

Collins, S. J., *et al*. (1977). Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture. *Nature*, **270**, 347–349.

Crooks, G. M., *et al*. (1999). Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. *Blood*, **94**, 519–528.

Cross, M. and Dexter, T. M. (1991). Growth factors in development, transformation and tumorigenesis. *Cell*, **64**, 271–280.

Cuenco, G. M., *et al*. (2000). Human AML1/MDS1/EVI1 fusion protein induces an acute myelogenous leukemia (AML) in mice: a model for human AML. *Proceedings of the National Academy of Sciences of the USA*, **97**, 1760–1765.

Daga, A., *et al*. (2000). The retroviral transduction of HOXC4 into human CD34(+) cells induces an *in vitro* expansion of clonogenic and early progenitors. *Experimental Hematology*, **28**, 569–574.

De'Angeac, A. D., *et al*. (1991). Transforming growth factor type beta 1 (TGF-beta1) down-regulates interleukin-2 production and up-regulates interleukin-2 receptor expression in a thymoma cell line. *Journal of Cellular Physiology*, **147**, 460.

Dexter, T. M., *et al*. (1977). Conditions controlling the proliferation of haematopoietic stem cells *in vitro*. *Journal of Cellular Physiology*, **91**, 335–345.

Eisen, H., *et al*. (1977). Induction of spectrim in Friend erythroleukemic cells. *Proceedings of the National Academy of Sciences of the USA*, **74**, 3898–4002.

Grassman, R., *et al*. (1989). Transformation to continuous growth of primary human T lymphocytes by human T cell leukemia virus type I X-region genes transduced by a Herpesvirus saimiri vector. *Proceedings of the National Academy of Sciences of the USA*, **86**, 3351.

Hayashi, K. and Akagi, T. (2000). An animal model for Epstein–Barr virus (EBV)-associated lymphomagenesis in the human: malignant lymphoma induction of rabbits by EBV-related

herpesvirus from cynomologus. *International Journal of Pathology*, **50**, 85–97.

Honda, H., *et al*. (1998). Development of acute lymphoblastic leukemia and myeloproliferative disorder in transgenic mice expressing p210bcr/abl: a novel transgenic model for human Ph1-positive leukemias. *Blood*, **91**, 2067–2075.

Hungerford, D. A., *et al*. (1959). The chromosome constitution of a human phenotypic intersex. *American Journal of Human Genetics*, **11**, 215–236.

Kizaki, M., *et al*. (1999). A novel retinoic acid-resistant acute promyelocytic leukemia model *in vitro* and *in vivo* (review). *International Journal of Molecular Medicine*, **4**, 359–364.

Kozuru, M., *et al*. (1996). High occurrence of primary malignant neoplasms in patients with adult T-cell leukemia/lymphoma, their siblings, and their mothers. *Cancer*, **78**, 1119–1124.

Laiho, J., *et al*. (1990). Growth inhibition by TGF-beta linked to suppression of retinoblastoma protein phosphorylation. *Cell*, **62**, 175.

Long, M. W. (1992). Blood cell cytoadhesion molecules. *Experimental Hematology*, **20**, 288.

Massage, J. (1990) The transforming growth factor-β family. *Annual Review of Cellular Biology*, **6**, 597–641.

Metcalf, D. (1989). The molecular control of cell division, differentiation, commitment and maturation of haematopoietic cells. *Nature*, **339**, 27–30.

Miyoshi, I., *et al*. (1981). Type C virus particles in a cord T cell line derived by cocultivating normal human cord leukocytes and human leukemic T cells. *Nature*, **294**, 770.

Mooradian, D. L., *et al*. (1989). Transforming growth factor-β1 binds to immobilized fibronectin. *Journal of Cellular Biochemistry*, **41**, 189.

Morecki, S., *et al*. (1998). Tumorigenicity and immunogenicity in a murine model of B-cell leukemia/lymphoma (BCL1). *Leukemia Research*, **22**, 831–835.

Musso, T., *et al*. (1990). Transforming growth factor β down-regulates interleukin-1 (IL-1)-induced IL-6 production by human monocytes. *Blood*, **76**, 2466.

Nakamura, T., *et al*. (1989). Molecular cloning and expression of human hepatocyte growth factor. *Nature*, **342**, 440–443.

Osborne, H. B., *et al*. (1982). Effect of dexamethasone on HMBA-induced Friend cell erythrodifferentiation. *Cancer Research*, **42**, 513–518.

Prasad, K. S. and Brandt, S. J. (1997). Target-dependent effect of phosphorylation on the DNA binding activity of the TAL1/SCL oncoprotein. *Journal of Biological Chemistry*, **272**, 11457–11462.

Rossi, G. B. and Friend, C. (1967). Erythrocytic maturation of (Friend) virus-induced leukemic cells in spleen clones. *Proceedings of the National Academy of Sciences of the USA*, **58**, 1373–1380.

Roussel, M. F., *et al*. (1987). Transforming potential of the c-fm proto-oncogene (CFS-1 receptor). *Nature*, **325**, 549.

Sheppard, R. D., *et al*. (1998). Transgenic N-myc mouse model for indolent B cell lymphoma: tumor characterization and analysis of genetic alterations in spontaneous and retrovirally accelerated tumors. *Oncogene*, **17**, 2073–2085.

Tamiya, S., *et al*. (1996). Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia. *Blood*, **88**, 3065–3073.

Thomson, A. M., *et al*. (1999). Haemopoietic progenitor cell lines generated by the myeloproliferative leukaemia virus: a model system to analyse murine and human lineage-affiliated genes. *British Journal of Haematology*, **107**, 33–48.

Till, J. E. and McCulloch, E. A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, **14**, 213.

Yoshida, M., *et al*. (1982). Isolation and characterization of retrovirus from cell lines of human adult T cell leukemia and its implication in the disease. *Proceedings of the National Academy of Sciences of the USA*, **79**, 2031.

Yoshida, M., *et al*. (1984). Monoclonal integration of HTLV in all primary tumors of adult T-cell leukemia suggests causative role of HTLV in the disease. *Proceedings of the National Academy of Sciences of the USA*, **81**, 2534.

Yunis, J. J. and Tanzer, J. (1993). Molecular mechanisms of hematologic malignancies. *Critical Reviews in Oncology*, **4**, 161–190.

Zentella, A., *et al*. (1991). Early gene responses to transforming growth factor-beta in cells lacking growth-suppressive RB function. *Molecular and Cellular Biology*, **11**, 4952.

FURTHER READING

Balkwill, F. R. (1995). *Cytokines: A Practical Approach* (IRL Press, Oxford).

Chiba, S. (1998). Homeobox genes in normal hematopoiesis and leukemogenesis. *International Journal of Hematology*, **68**, 343–353.

Cooper, G. M. (1996). *The Cell. A Molecular Approach* (ASM Press, Washington, DC).

Dainiak, N. (1991). Surface membrane-associated regulation of cell assembly, differentiation, and growth. *Blood*, **78**, 264.

Devita, V. T., *et al*. (1999). *Cancer. Principles and Practice of Oncology*, 5th edn (Lippincott, Philadelphia).

Dexter, T. M., *et al*. (1984). Long-term marrow culture: an overview of technique and experience. In: Wright, D. G. and Greenberger, J. S. (eds), *Long-term Bone Marrow Culture*. Kroc Foundation Series 18. 57–96 (Liss., New York).

Freshney, R. I. (1993). *Culture of Animal Cells. A Manual of Basic Techniques*, 3rd edn (Wiley-Liss, New York).

Nicolson, G. L. (1988). Cancer metastasis: tumor cell and host organ properties important in metastasis to specific sites. *Biochimica et Biophysica Acta*, **948**, 175–224.

Papayannopoulou, T. and Abkowitz, J. (1991). Biology of erythropoiesis, erythroid differentiation, and maturation. In: Hoffman, R., *et al*. (eds), *Hematology: Basic Principles and Practice*. 252–263 (Churchill Livingstone, New York).

Stamatoyannopoulos, G., *et al*. (1994). *The Molecular Basis of Blood Diseases*, 2nd edn (Saunders, Philadelphia).

Chapter 72

Models for Melanomas and Sarcomas

Laura L. Worth, Eugenie S. Kleinerman and Menashe Bar-Eli
University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

CONTENTS

• Models for Melanoma
• Models for Sarcomas

MODELS FOR MELANOMA

Malignant melanoma is a common human cancer with high mortality rates. The incidence of this disease is currently increasing faster than that of any other malignancy. The primary cause of melanoma is thought to be exposure to ultraviolet (UV) radiation (Fountain *et al*., 1990). Other risk factors include fair skin, a hereditary predisposition toward the disease, increased age and race. In the white population, for example, the disease is increasing rapidly and was projected to affect approximately one in 75 persons by the start of the twenty-first century (Ahmed, 1997). The steady overall increase has been attributed to various causes ranging from the reduction of the stratospheric ozone to the changing lifestyles and attitudes that affect sun exposure. Although it is easily diagnosed and wholly curable if detected early, the high mortality rate (6700 deaths each year in the United States) is due to the propensity for melanoma to metastasize if left untreated (Dooley, 1994).

Most human malignant melanomas develop as a consequence of a lengthy multistep process. This process involves formation of naevi from normal melanocytes, a radial growth phase (RGP) and a subsequent vertical growth phase (VGP; metastatic phenotype), which results in decreased dependence on growth factors, reduction of contact inhibition and diminished anchorage dependence (Lu and Kerbel, 1994). The prevalent working model for melanoma metastasis development defines it as a series of interrelated, sequential steps involving the survival and growth of unique subpopulations of cells with metastatic properties that preexisted within the parental neoplasm (Fidler, 1990).

The molecular steps leading to human malignant melanoma and subsequent metastasis are not well known.

Several animal models have been established to study the progression of melanoma. **Table 1** summarizes the major melanoma models of both human and nonhuman origin. Each model will be discussed here for its relevance to study the human disease.

Table 1 Experimental models in melanoma research

Type	Site	Description
Mouse models		
DMBA + UV-induced melanoma	*In vivo*	Mouse skin treated with carcinogen and UV irradiation
Allografts	*In vivo*	Murine melanoma cell lines (B16, K-1735) transplanted into syngeneic mice
Transgenic: SV-40	*In vivo*	Tyrosinase proporter-driven SV-40 induced-melanoma
Transgenic: c-Ret	*In vivo*	Metallothionein-driven c-Ret induced-melanoma
Transgenic: HGF	*In vivo*	Metallothionein-driven hepatocyte growth factor-induced melanoma
Transgenic: $p16^{INK4a}$ + H-ras	*In vivo*	Tyrosinase-driven Ha-ras in $p16^{INK4a}$ knockout mice
Human/mouse xenografts	*In vivo*	Human tumours injected s.c. or i.v. into nude or SCID mice
Human models		
Artificial skin reconstructs	*In vitro*	All components of human skin cells are cocultured to simulate human skin
Xenografts of human skin	*In vivo*	Human skin grafted on to SCID/RAG mice and melanomas are induced
Other animal models		
Monodelphis domestica (opossum)	*In vivo*	Melanomas are induced by UV irradiation
Sinclair swine	*In vivo*	Pigs develop spontaneous melanomas
Syrian hamster	*In vivo*	DMBA-induced or spontaneous melanomas
Guinea pig	*In vivo*	DMBA-induced melanomas
Xiphophorus fish	*In vivo*	UV-induced or spontaneous melanomas

Mouse Models

Currently there are no spontaneous murine melanoma models, and the existing models consist of induced melanomas. The induced murine melanoma models can be divided into two groups: (1) induction of melanomas by means of UV radiation and carcinogens and (2) induction of melanoma by specific genetic manipulations (transgenic mice). These models are syngeneic and have the advantage of an intact immune system. This allows the investigators to study not only the molecular events within the tumours, but also to look into the immune responses against these tumours. Since the murine skin differs from the human skin in its architecture, these models, however, may not represent the actual molecular events involved in the progression of human melanoma.

In group (1), investigators used the carcinogen 7,12-dimethylbenz[*a*]anthracene (DMBA) as an initiator followed by UV radiation, croton oil or TPA as promoters to induce melanomas in mice. Nearly all of the mice developed skin cancer, and around 20% of the animals developed primary cutaneous melanoma (Donawho *et al.*, 1992). Several melanoma cell lines have been generated (B16 and K-1735) and established in culture. These cell lines were injected as allografts to generate several sublines with different metastatic capabilities (Epstein *et al.*, 1967; Romerdahl *et al.*, 1989; Donawho *et al.*, 1992). These syngeneic models are extensively in use to study the molecular changes associated with the acquisition of the metastatic phenotype.

Recently, transgenic models to generate melanomas has been the subject of renewed research. Indeed, several transgenic mouse models are available. All of these mice were generated by utilizing the expression of viral or human oncogenes under inducible or tissue specific promoters (**Table 1**). For example, Mintz and her colleagues used the tyrosinase promoter-driven SV-40 antigen to induce melanoma in mice (Bradl *et al.*, 1991). Other investigators used the c-*Ret* or hepatocyte growth factor/scatter factor genes driven by the metallothionein promoter to generate melanomas (Iwamoto *et al.*, 1991; Takayama *et al.*, 1997). Recently, transgenic mice bearing melanomas were established by crossing knockout mice for $p16^{INK4a}$ with mice expressing activated H-ras in mouse melanocytes using the tyrosinase promoter (Chin *et al.*, 1997). In general, these transgenic animal models suffer the disadvantage that the genes they utilized to induce melanoma were not necessarily have been shown to be involved in the progression of human melanoma *in situ*. In addition, these models provide limited potential to study the aetiology and the genetic events involved in the progression of human melanoma because unlike in the human skin, mouse melanocytes are localized deep in the hair shafts in the dermis.

Human/Mouse Xenografts

This is the most widely used model to study the regulation of human melanoma tumour growth and metastasis. There are close to 50 human melanoma cell lines available to the scientific community at large. These cell lines represent several stages in the progression of human melanoma. Most of the cell lines were established from metastatic lesions (lymph nodes, brain or skin metastases) and as such they fall under the category of VGP. Several other cell lines were established from skin tumours at the RGP, while very few cell lines representing primary cutaneous melanoma are available owing to the difficulties of growing them in culture. This wide range of cell lines (sometimes from the same patient at different stages of the disease) enables us to compare their gene expression *in vitro*, to manipulate their gene expression pattern *ex vivo* and subsequently to analyse the effect of this manipulation *in vivo* by implanting them as xenografts in an orthotopic model (subcutaneous injections) (Huang *et al.*, 1996, 1998; Xie *et al.*, 1997). The human cell lines are injected into the immunocompromised nude or SCID mice. In general, there is a good correlation between the tumourigenicity of the tumours in patients from which they originated and their ability to grow subcutaneously in nude or SCID mice. The incidents of spontaneous metastasis from the primary tumours to other organs in these animal models are very low. An alternative approach has been utilized to study the metastatic potential of these cell lines by using the experimental lung metastasis assay. In the assay, melanoma cells (usually 1×10^6 or 2×10^6) are injected into the tail vein of nude mice and a few weeks later the number of lung colonies counted under a dissecting microscope. Our experience with this assay shows that in most cases a direct correlation exists between the behaviour of tumour cells in the patient and their metastatic potential in nude mice. Unlike the environment setting in the patient, however, the environment in these models is lacking certain components of the immune response. In addition, growth and angiogenic factors released by the implanted human tumour cells may not work efficiently on the stroma cells within their immediate host environment. Despite these deficiencies, these models are widely used to study the effect or the role of target genes on tumour growth and metastasis of human melanoma.

Human Models

Artificial Skin Reconstructs

In this artificial model (Meier *et al.*, 2001), all cell types of the human skin are cultured together to mimic the skin environment. It involves coculture of terminally differentiated epidermal cells, a basement membrane layer and a dermal compartment consisting of fibroblasts embedded in collagen. All layers are arranged in a physiological context

that functions like the human skin. The model has been used as an alternative means to animal testing for skin irritation (cosmetic products), wound healing and UV-induced damage (Ponec and Kempenaar, 1995; Simonetti *et al*., 1995; Kuroyanazi *et al*., 1996). Recently, however, Herlyn and his colleagues have utilized this model to study the biology of melanocytes and the progression of human melanoma. When normal melanocytes were applied to these reconstructs, they retained their dendritic morphology and did not divide for a period of 35 days. Melanoma cells, on the other hand, exhibited the same biological properties of human melanoma cell *in vivo*. For example, RGP melanoma cells when incorporated onto the reconstructs, grew as single cells or small nests within the epidermis/dermis junction and did not invade into the dermis. VGP melanoma cells formed clumps and clusters within the epidermis/dermis junction with some of the cells crossing into the dermis. Metastatic melanoma cells exhibited accelerated proliferation with the majority of the cells invading deep into the dermis (Satyamoorthy *et al*., 1999). These data point to the usefulness of this model to study the behaviour of melanoma cells in the nonanimal setting. As in the xenograft/nude/SCID mice model, the skin reconstructs are ideal to study altered gene expression with respect to melanoma.

Mouse/Human Chimaera: Xenografts of Human Skin on to SCID Mice

Using the same principles to induce melanomas in mice, i.e. DMBA coupled with UV radiation, Herlyn's group has tried to induce human melanoma in human skin grafted on to SCID/RAG mice. The combination treatments of DMBA followed by UVB irradiation produced high incidents of squamous cell carcinomas and a human nodular malignant melanoma in one mouse. Grafting skins from normal donors or from patients with melanoma could be a valuable tool to study the molecular mechanisms associated with the progression of human melanoma (Soballe *et al*., 1996; Attilasoy *et al*., 1998).

Other Animal Models

Other animal models to generate melanomas have been the subject of considerable research. Among the models worth mentioning are those in the opossum (Ley *et al*., 1989), Sinclair swine (Millikan *et al*., 1974), Syrian hamster (Chernozemsai and Raichev, 1966), guinea pig (Clark *et al*., 1976) and the *Xiphophorus* hybrid fish (Setlow *et al*., 1989) (**Table 1**). In the South American opossum, melanomas were induced by UV irradiation and were found to metastasize spontaneously to other organs. The Sinclair swine model is interesting in the sense that the pigs developed spontaneous melanomas at birth, which also regressed spontaneously in adult life. The mechanisms by which these tumours regress are still unknown and may provide valuable information for the progression of the human disease as the skin of the pig resembles the structure of the human skin. The drawbacks in all these animal models are the requirement to clone the species-specific human-homologous genes in each case. This challenge causes difficulties in comparing human melanoma with those developed in these animal models.

MODELS FOR SARCOMAS

Osteosarcoma is a primary bone malignancy that usually occurs during the adolescent and young adult years when the long bones are undergoing rapid growth. Current therapy consists of preoperative chemotherapy, surgical intervention (either amputation or limb-sparing surgery depending upon the resectability of the tumour) followed by postoperative chemotherapy. Despite this aggressive therapy, 30–40% of patients will relapse, usually within 1 year after diagnosis. Although some patients will develop local recurrence of the tumour or metastases at other bony sites, the majority of patients will relapse in the lung. Patients who relapse have a poor prognosis. Salvage chemotherapeutic regimens have had limited success. Consequently, new chemotherapeutic agents and novel therapeutic approaches are needed.

As a result, investigators are using osteosarcoma cell lines and developing animal models in which to identify new therapies for testing in clinical trials. In the following sections the benefits and disadvantages of various cell lines and xenographic and syngeneic animal models of osteosarcoma will be discussed. Finally, the canine model of osteosarcoma, the only spontaneously occurring animal model will be described.

Cell Lines

A large number of osteosarcoma cell lines are available through the American Type Culture Collection (ATCC) (see www.ATCC.org). An advantage of using cell lines to assess the effect of chemotherapeutic and immunomodulatory agents is that a large number of agents can be screened relatively rapidly. Three human cell lines, SAOS-2, MG-63 and TE-85, have been well characterized with regard to their sensitivity to chemotherapeutic and immunomodulatory agents. For example, the antitumour effects of the chemotherapeutic agents etoposide, adriamycin, amsacrine, topotecan, vinblastine and cisplatin and the immunomodulating cytokines TNF-α, interleukin-1 and interferon-γ were assayed for their cytotoxic effect on these three cell lines (Jia *et al*., 1995). Interestingly, SAOS-2 cells were most sensitive to the effects of TNF-α and interferon-γ, while MG-63 cells were most sensitive to interleukin-1, and TE-85 cells were resistant to TNF-α and IL-1, but sensitive to interferon-γ (Jia and Kleinerman,

1991). This heterogeneous response is similar to what happens in a tumour when only a portion of cells making up the tumour are sensitive to a particular agent. The population of cells that are sensitive die off while the resistant population continues to grow and ultimately results in the death of the patient. So if a single cell line is used to assess a new agent, it is not clear if other cell lines will also be sensitive or if a human tumour will be sensitive.

In the analysis of chemotherapeutic agents in combination with cytokines, interleukin-1α (IL-1α) enhanced the activity of etoposide in all three cell lines (Jia *et al.*, 1999a). Based on this result, a clinical trial was undertaken in patients with relapsed osteosarcoma (Worth *et al.*, 1997). Although the study had to be stopped prematurely because IL-1α production was halted, analysis of the patients treated on this study indicated that the combination of IL-1α and etoposide was effective in some patients. The number of osteosarcoma cell lines available for *in vitro* and *in vivo* studies is very limited and there is a need to establish more cell lines.

Animal Models

Human–Murine Xenografts

Xenografts have several advantages over cell lines. The most commonly used model introduces human cells in a murine host. A major drawback of xenograft models is that human cells will not grow in immunocompetent animals, the immune system of an animal will recognize the human cells as foreign and reject the tumour. Consequently, as in the melanoma models, immunocompromised mice must be used, and the effect of the immune system on therapy cannot be adequately assessed.

MNNG/HOS is human cell osteosarcoma line that produces tumours when implanted subcutaneously in nude mice (Chen *et al.*, 1997) (**Table 2**). This model is useful in that the tumours are well circumscribed and the effect of therapy can be easily quantified. However, this cell line does not grow in bone and does not metastasize. The effects of therapy observed on tumours growing subcutaneously may not be the same on tumours growing in the bone or in the lung.

A model devised by Meyers *et al.* (1990) is unique in that it uses surgical specimens, not cell lines, and immune-deprived mice, not nude mice. In this model, CBA/CaJ inbred mice are subjected to infant thymectomy and total body irradiation with subsequent marrow reconstitution. Tumour specimens obtained at the time of surgery are directly implanted into flank muscle. Tumour cells grow at the site of inoculation; however, the tumour does not metastasize, and this limits the usefulness of this model.

Another model (Berlin *et al.*, 1993) requires the v-Ki-*ras* oncogene transformation of the human osteosarcoma TE-85 cell line. This new cell line was designated KRIB. When KRIB cells are intravenously or intraosseously injected into nude mice, they reproducibly form pulmonary metastases.

We had concerns about some of the manipulations of the host and tumour cells that might decrease the relevance of the models in terms of evaluating the efficacy of various therapies, so we developed our own experimental osteosarcoma model. SAOS-2 cells are available through the ATCC. They will not grow when injected intravenously, intratumourally or subcutaneously in nude mice. A population of SAOS-2 cells that were able to grow on 0.9% agarose were selected and injected intravenously into a nude mouse. A rare lung metastasis that developed more than 6 months later was isolated, expanded in tissue culture and reinjected (Radinsky *et al.*, 1994). This process was repeated a total of six times to produce SAOS-LM6 (Jia *et al.*, 1999b). The advantage of this model is that there is a natural selection of cells without manipulation of the host or tumour cells. The formation of micrometastases in 4–6 weeks and macroscopic disease by 8 weeks is reproducible. This model has been demonstrated to be useful in evaluating standard chemotherapeutic agents as well as new immunotherapies (Jia *et al.*, 1999b; Worth *et al.*, 2000).

Syngeneic Models

In syngeneic models (**Table 3**), tumour cells are injected into the same species/haplotype animals. The advantage of a syngeneic animal model is that immunocompetent animals can be used. This allows the interaction of the immune system on the new therapy to be assessed. Consequently, responses seen in these animal models may more closely mimic the responses in humans. Three murine syngeneic models for metastatic osteosarcoma have been described.

Table 2 Human–murine xenograft models for osteosarcoma

Cell line	Heterotropic/orthotopic	Metastatic potential	Manipulation of host	Reference
MNNG/HOS	Heterotropic (s.c.)	Rare	None	Chen *et al.* (1997)
Surgical specimens	Orthotopic? (flank muscle)	None	TBI/infant thymectomy	Meyers *et al.* (1990)
KRIB	Orthotopic (tibia)	Moderate (0–22 range)	v-Ki-*ras* oncogene transformation of cell line	Berlin *et al.* (1993)
SAOS-LM6	Heterotropic (i.v.)	High (> 200)	None	Jia *et al.* (1999b)

Table 3 Syngeneic murine models for osteosarcoma

	Heterotropic/orthotopic injection	Metastatic potential	Reference
Radiation-induced	Heterotropic (muscle)	Can approach 96%	Pollak *et al.* (1992)
Spontaneous	Heterotropic (subcutaneous)	Nearly 100%	Asai *et al.* (1998)
Spontaneous	Orthotopic (intraosseous)	High	Khana *et al.* (2001)

The murine osteosarcoma in the first model originated in a C3H-Sed mouse that had been irradiated (Pollak *et al.*, 1992). A fragment of this tumour was then implanted in the lateral gastrocnemius muscle. The tumour produces abundant neoblastic bone that invades the proximal skeletal muscle and metastasizes to the lung. The incidence of metastases is dependent upon when the original tumour is amputated. If the original tumour is amputated when it is 12 mm, 65% of the animals will develop pulmonary metastases. If the tumoured leg is amputated when the mass approaches 16 mm, 96% of the animals will develop pulmonary metastases (average number of metastases 16 ± 6). The advantage of this model is that it uses immunocompetent mice. At the time of diagnosis most patients have microscopic metastatic disease. It is unclear if this model behaves in the same manner with the early development of microscopic pulmonary metastases.

A second syngeneic mouse model uses a murine osteosarcoma that was first described by Dunn and Andevont (1963). This tumour arose spontaneously in the tail vertebra of a mouse. When injected into the femoral muscle, tumours will grow, but will not metastasize. Asai *et al.* (1998) took this cell line and cycled it through a mouse eight times to produce a cell line that is highly metastatic to the lung. This model is advantageous in that it reproducibly produces a large number of metastases. The disadvantage is that this is not an orthotopic model. The cells are implanted subcutaneously and not in bone. The authors conclude that this model is easier to use since subcutaneous tumours are easier to inject and measure as compared with intraosseous injections; however, it is not clear if therapies that are effective on subcutaneous tumours will have the same effect on intraosseous masses. Organ-specific growth factors may play a key role in the growth of tumour cells, thus tumour cells that grow in the lung may not grow in the skin and vice versa.

A recent report describes a new syngeneic, orthotopic model (Khana *et al.*, 2001). K7 and K12 are murine osteosarcoma cell lines. K7 cells were injected intraosseously into the tibia of a Balb/c mouse. A spontaneous metastasis arose in the lung, was isolated and surgically implanted into a paraosteal muscle flap. A highly metastatic pulmonary cell line was established from a pulmonary metastasis and was designated K7M2. This model has several advantages. It is syngeneic, orthotopic and results in spontaneous pulmonary metastases. Cell lines with a high metastatic and low metastatic potential can be compared. If differences can be identified between these cell lines, it is possible that new therapies could be designed to decrease the high metastatic potential of some tumour cells.

Spontaneous Canine Osteosarcoma

Similar to humans, large-breed dogs develop osteosarcoma that metastasizes to the lung. At the time of diagnosis nearly 90% of dogs have pulmonary micrometastases. This is based on the observation that dogs who have their primary tumoured limb amputated and receive no further therapy have a median survival of 3–4 months and die of pulmonary metastases. Dogs with osteosarcoma were instrumental in the preclinical of evaluation of an agent, L-MTP-PE, that activates macrophages to kill tumour cells (MacEwen *et al.*, 1989). Treatment of dogs with L-MTP-PE increased the survival time of these dogs from a median of 77 to 222 days. Based on these results, L-MTP-PE was introduced into clinical trials. A phase III multi-institutional randomized trial involving L-MTP-PE in patients with newly diagnosed osteosarcoma was completed recently and for the first time in more than 15 years there is an increase in the 5-year disease-free survival rate. The dog osteosarcoma model is a vary useful model for preclinical studies. The only disadvantage to this model is that although more than 8000 dogs are diagnosed with osteosarcoma per year, not all pet owners are willing to seek expensive therapy for dogs or to enter clinical trials.

Clinical Relevance of Experimental Tumour Behaviour

Each model system used to identify new therapeutic approach to osteosarcoma, be it the *in vitro* cell culture, a xenographic or syngeneic system, or the osteosarcoma that occurs spontaneously in the dog, has its unique set of advantages and disadvantages. By keeping the limitations in mind, we will be able to identify new therapies that could have a significant impact on survival rates of patients with osteosarcoma.

REFERENCES

Ahmed, I. (1997). Malignant melanoma: prognostic indicators. *Mayo Clinic Proceedings*, **72**, 356–361.

Asai, T., *et al.* (1998). Establishment and characterization of a murine osteosarcoma cell line (LM8) with high metastatic

potential to the lung. *International Journal of Cancer*, **76**, 418–422.

Atillasoy, E. S., *et al*. (1998). UVB induces atypical melanocytic lesions and melanoma in human skin. *American Journal of Pathology*, **152**, 1179–1186.

Berlin, O., *et al*. (1993). Development of a novel spontaneous metastasis model of human osteosarcoma transplanted orthotopically into bone of athymic mice. *Cancer Research*, **53**, 4890–4895.

Bradl, M., *et al*. (1991). Malignant melanoma in transgenic mice. *Proceedings of the National Academy of Sciences of the USA*, **88**, 164–168.

Chen, D. S., *et al*. (1997). Retroviral vector-mediated transfer of an antisense cyclin G1 construct inhibits osteosarcoma tumour growth in nude mice. *Human Gene Therapy*, **8**, 1667–1674.

Chernozemski, I. and Raichev, R. (1966). Two transplantable lines from melanomas induced in syrian hamsters with 9,10-dimethyl-1,2-benz[*a*]anthracene (DMBA). *Neoplasma*, **13**, 577–582.

Chin, L., *et al*. (1997). Cooperative effects of INK4a and ras in melanoma susceptibility *in vivo*. *Genes and Development*, **11**, 2822–2834.

Clark, W. H., Jr, *et al*. (1976). The developmental biology of induced malignant melanoma in guinea pigs and a comparison with other neoplastic systems. *Cancer Research*, **36**, 4079–4091.

Donawho, C., *et al*. (1992). Immunobiology of primary murine melanomas. *Journal of Immunotherapy*, **12**, 187–193.

Dooley, T. (1994). Recent advances in cutaneous melanoma oncogenesis research. *Oncology Research*, **6**, 1–9.

Dunn, T. B. and Andervont, H. B. (1963). Histology of some neoplasms and non-neoplastic lesions found in wild mice maintained under laboratory conditions. *Journal of the National Cancer Institute*, **31**, 872–901.

Epstein, J. H., *et al*. (1967). Production of melanomas from DMBA-induced 'blue nevi' in hairless mice with ultra-violet light. *Journal of the National Cancer Institute*, **38**, 19–30.

Fidler, I. J. (1990). Critical factors in the biology of human cancer metastasis: twenty-eighth GHA Clowes memorial award lecture. *Cancer Research*, **50**, 6130–6138.

Fountain, J. W., *et al*. (1990). Genetics of melanoma. *Cancer Surveys*, **9**, 645–671.

Huang, S., *et al*. (1996). Enforced c-KIT expression renders highly metastatic human melanoma cells susceptible to stem cell factor-induced apoptosis and inhibits their tumourigenic and metastatic potential. *Oncogene*, **13**, 2339–2347.

Huang, S., *et al*. (1998). Loss of AP-2 results in downregulation of c-KIT and enhancement of melanoma tumourigenicity and metastasis. *EMBO Journal*, **17**, 4358–4369.

Iwamoto, T., *et al*. (1991). Aberrant melanogenesis and melanocytic tumour development in transgenic mice that carry a metallothionein/ret fusion gene. *EMBO Journal*, **10**, 3167–3175.

Jia, S. F. and Kleinerman, E. S. (1991). Antitumour activity of TNF-alpha, IL-1, and IFN-gamma against three human osteosarcoma cell lines. *Lymphokine and Cytokine Research*, **10**, 281–284.

Jia, S. F., *et al*. (1995). Antitumour effects of IL-1 alpha and VP-16, ADR, mAMSA, VBL, TPT, cDDP against osteosarcoma cells lines. *Proceedings of the National Meeting of the American Association for Cancer Research*, **36**, A28757.

Jia, S. F., *et al*. (1999a). Interferon-α enhances the sensitivity of human osteosarcoma cells to etoposide. *Journal of Interferon and Cytokine Research*, **19**, 617–624.

Jia, S. F., *et al*. (1999b). A nude mouse model of human osteosarcoma lung metastases for evaluation of new therapeutic strategies. *Clinical and Experimental Metastases*, **17**, 501–506.

Khana, C., *et al*. (2001). An orthotopic model of murine osteosarcoma with clonally related variants differing in pulmonary metastatic potential. *Clinical Cancer Research*, in press.

Kuroyanagi, Y., *et al*. (1996). Cytotoxicity tests for antimicrobial agents using cultured skin substitutes fixed at interface of air and culture medium. *Journal of Biomaterials Sciences, Polymer Edition*, **7**, 1005–1015.

Ley, R. D., *et al*. (1989). Ultraviolet radiation-induced malignant melanoma in *Monodelphis domestica*. *Photochemistry and Photobiology*, **50**, 1–5.

Lu, C. and Kerbel, R. (1994). Cytokines growth factors and loss of negative growth controls in the progression of human cutaneous malignant melanoma. *Current Opinions in Oncology*, **6**, 212–220.

MacEwen, E. G., *et al*. (1989). Therapy for osteosarcoma in dogs with intravenous injection of liposome-encapsulated muramyl tripeptide. *Journal of the National Cancer Institute*, **81**, 935–938.

Meier, F., *et al*. (2001). Human melanoma progression in skin reconstructs: biological significance of bFGF. *American Journal of Pathology*, in press.

Meyers, W. H., *et al*. (1990). Development and characterization of pediatric osteosarcoma xenograft. *Cancer Research*, **50**, 2781–2785.

Millikan, L. E., *et al*. (1974). Melanoma in Sinclair swine: a new animal model. *Journal of Investigative Dermatology*, **62**, 20–30.

Pollak, M., *et al*. (1992). Inhibition of metastatic behavior of murine osteosarcoma by hypophysectomy. *Journal of the National Cancer Institute*, **84**, 966–971.

Ponec, M. and Kempenaar, J. (1995). Use of human skin recombinants as an *in vitro* model for testing the irritation potential of cutaneous irritants. *Skin Pharmacology*, **8**, 49–59.

Radinsky, R., *et al*. (1994). Terminal differentiation and apoptosis in experimental lung metastases of human osteogeneic sarcoma cells by wild type p53. *Oncogene*, **9**, 1877–1883.

Romerdahl, C. A., *et al*. (1989). The role of ultraviolet radiation in the induction of melanocytic skin tumours in inbred mice. *Cancer Communications*, **1**, 209–216.

Satyamoorthy, K., *et al*. (1999). Human xenografts, human skin and skin reconstructs for studies in melanoma development and progression. *Cancer and Metastasis Reviews*, **18**, 401–405.

Setlow, R. B., *et al*. (1989). Animal model for ultraviolet radiation-induced melanoma: platyfish-swordtail hybrid. *Proceedings of the National Academy of Sciences of the USA*, **86**, 8922–8926.

Simonetti, O., *et al*. (1995). Visualization of diffusion pathways across the stratum corneum of native and *in vitro* reconstructed epidermis by confocal laser scanning microscopy. *Archives of Dermatological Research*, **287**, 465–473.

Soballe, P. W., *et al*. (1996). Carcinogenesis in human skin grafted to SCID mice. *Cancer Research*, **56**, 757–764.

Takayama, H., *et al*. (1997). Diverse tumourigenesis associated with aberrant development in mice overexpressing hepatocyte growth factor/scatter factor. *Proceedings of the National Academy Sciences of the USA*, **94**, 701–706.

Worth, L. L., *et al*. (1997). Phase II study of recombinant interleukin-1-alpha (IL-1α) and etoposide in patients with relapse osteosarcoma. *Clinical Cancer Research*, **3**, 1721–1729.

Worth, L. L., *et al*. (2000). Intranasal therapy with an adenoviral vector containing the murine interleukin-12 gene eradicates osteosarcoma lung metastases. *Clinical Cancer Research*, **6**, 3713–3718.

Xie, S., *et al*. (1997). Expression of *MCAM/MUC18* by human melanoma cells leads to increased tumour growth and metastasis. *Cancer Research*, **57**, 2295–2303.

FURTHER READING

Killion, J. J., *et al*. (1999). Orthotopic models are necessary to predict therapy of transplantable tumors in mice. *Cancer and Metastasis Reviews*, **17**, 279–284.

Kleinerman, E. S. (1999). Biological therapy for osteosarcoma using liposome-encapsulated muramyl tripeptide. *Hematology/Oncology Clinics of North America*, **9**, 927–938.

MacEwen, E. G. (1996). Canine osteosarcoma. *Controversies in Clinical Oncology*, **26**, 123–133.

Satyamoorthy, K., *et al*. (1999). Human xenografts, human skin and skin reconstructs for studies in melanoma development and progression. *Cancer and Metastasis Reviews*, **18**, 387–400.

Chapter 73

Models for CNS Malignancies

Laura E. Crotty, John H. Sampson, Gary E. Archer and Darell D. Bigner

Duke University Medical Center, Durham, NC, USA

CONTENTS

- Introduction
- Principles of Experimental Model Applications
- *In Vivo* Experimental Techniques
- Interpretation of Experimental Results
- Conclusion

INTRODUCTION

Central nervous system (CNS) neoplasms must be approached differently from other neoplasms because of their location in the brain. While designing brain tumour therapy studies, experimenters must be cognisant of the sensitivity of CNS tissue to radiation, chemotherapy, cytokine therapy, and autoimmune reactions (experimental allergic encephalitis (EAE)), and experimental design must take into account special features of the brain such as its compartmentalization. The brain is in a compartment separated from the rest of the body by the blood–brain barrier (BBB) and this compartmentalization creates special challenges. One challenge is generating an immune response to a brain tumour when there are no lymphatics in the brain and very few immune system cells allowed into brain parenchyma. This chapter covers the methods used to work around these sensitivities and challenging features and describes in-depth the best models for brain tumour experimentation.

The chapter first presents a discussion of specific characteristics of CNS tumours, metabolism, morphology and cytokine secretion. Next, special considerations regarding establishment of human brain tumour cell lines, *in vitro* culture of CNS tumour cells, and *in vivo* models using human brain tumour cells are reviewed. Syngeneic animal tumour models are discussed, with a focus on characteristics of an ideal *in vivo* brain tumour model. The final section illustrates how results in animal models can lead to new therapeutic strategies utilized in the clinical setting. This chapter is designed to be a guide for the novice brain tumour investigator such that appropriate experimental models can be selected, implemented and interpreted in terms of clinical trial possibilities.

Characteristics of CNS Tumours

Primary CNS tumours can arise from any nervous system cell type, producing a spectrum of approximately 50 different tumour types. The most prevalent primary CNS tumours are those arising from neuroepithelial tissue, such as gliomas: astrocytoma, oligodendroglioma and ependymoma. The second most common tumours are those derived from cranial and spinal nerves, such as schwannoma and neurofibroma. Other sources of CNS tumour are the meninges, anterior pituitary and embryonal tissue (**Table 1**).

CNS tumours have a wide range of neoplastic expression patterns, involving characteristics such as anaplasia, growth rate, antigenic expression and genetic differentiation. The degree of anaplasia in CNS tumours ranges from mature cell types with specific antigenic expression patterns to cells with a lack of any normal cytological features. Growth patterns range from the fast-growing glioblastoma multiforme (GBM) to the slow-growing pilocytic astrocytoma that can take years to develop. Prevalent antigenic expression patterns in human brain tumours include synaptophysin, neurofilament proteins, glial fibrillary acidic protein (GFAP) and S100 (surface Ag). Common genetic changes are the loss of chromosome 10 (a high incidence in GBM, while uncommon in low-grade gliomas, suggesting its importance for tumour progression); epidermal growth factor receptor (EGFR) mutants; gains of chromosome 7; structural abnormalities of 9p, 19q and 22; loss of 1p, 9p, 10q, 17p, 19q and 22q (frequent in gliomas); loss of IFNα and -β genes (high-grade gliomas); overexpression of EGFR mutants, EGFR, PDGFR and MDM-2; p53 mutations; and p16 deletions. Loss of tumour-suppressor genes is fundamental for the development of many CNS tumours, especially the loss of *Rb*, *TP53*, *WT-1*, *WAF1*, *MTS1*, *NF1*, *NF2*, *DCC* and *MCC*. An example of how some of these chromosomal changes come together to cause malignancy is the transformation of glial cells to astrocytoma, anaplastic astrocytoma (AA) and finally GBM. This progression is caused by a specific series of chromosomal changes. The transformation to astrocytoma often involves the loss of 17p (p53 mutation), PDGF-α receptor overexpression and loss

Table 1 Primary brain tumours

Tumours derived from	Tumour type	Typical WHO Grade	Usual location
Neuroepithelial tissue	A. Pilocytic astrocytoma	I	Cerebellum, third ventricle and optic nerve
	Astrocytoma	II	Cerebral hemispheres, optic pathways, third ventricle, cerebellum and brainstem
	Anaplastic astrocytoma	III	Cerebral hemispheres
	Glioblastoma	IV	Cerebral hemispheres
	B. Oligodendroglioma	II	
	Anaplastic oligodendroglioma	III	
	C. Ependymoma	II	Floor of the fourth ventricle, central canal of the spinal cord, lateral or third ventricles and cerebellopontine angle
	D. Mixed oligoastrocytoma	II, III	
	E. Choroid plexus tumour	Papilloma: I Carcinoma: III	Choroid plexus
	Pineocytoma	II	
	Pineoblastoma	IV	
Embryonal tissue	Medulloblastoma	IV	Cerebellum or brainstem
	Neuroblastoma		
	Ependymoblastoma	I, II, III	
	Primitive neuroectodermal tumour	I, II, III	Supratentorial compartment
Cranial and spinal nerves	A. Schwannoma	I	Meatus acusticus internus or externus in the cerebello-pontine angle
	B. Neurofibroma	III, IV	
	C. Malignant peripheral nerve sheath tumour	III, IV	
	Malignant schwannoma	III, IV	
	Neurogenic sarcoma	III, IV	
	Anaplastic neurofibroma	III, IV	
Meninges	Meningioma	I	Dura
	Atypical meningioma	II	Dura
	Anaplastic meningioma	III	Dura and brain invasion
Sellar region	Pituitary adenoma	I	Sellar region
	Pituitary carcinoma	II	Brain invasion from sellar region, metastasis
	Craniopharyngioma	I	Sellar and suprasellar regions
Outside the CNS	Metastasis	IV	

of 22q. Progression to anaplastic astrocytoma involves the loss of 13q (Rb mutation), loss of 19q, loss of 9p (IFNα/β and p16^{INK4} mutation), and *CDK4* and *MDM2* gene amplification. Final progression to GBM involves *EGFR* gene amplification and rearrangement, and the loss of 10p and 10q.

CNS tumours have been shown to secrete cytokines to modulate the immune system. GBM tumours have been demonstrated to secrete TGF-β, IL-1α and -1β, IL-6, IL-10 and GM-CSF. Within a single tumour type, vast heterogeneity can be observed. For example, the malignant gliomas AA and GBM are each heterogeneous at morphological, biological, genetic and antigenic levels (Bigner *et al*., 1981). This heterogeneity is believed to be a consequence of the plasticity of neoplastic glial transformation.

No matter what the specific type or expression pattern, CNS tumours are expanding lesions that produce symptoms primarily by taking up space within the skull. As they grow through local diffuse infiltration or within distinct encapsulated areas of the brain, CNS tumours increase the mass of tissue residing within the rigid skull. CNS tumours

may also increase intracranial mass by blocking the flow of CSF, causing hydrocephaly. Increased mass leads to increased intracranial pressure, which causes compression and displacement of the brain and spinal cord. These intracranial changes may ultimately lead to herniation of the brainstem, followed by death. New brain tumour therapies should be tested thoroughly in experimental *in vivo* intracranial models to prove that the treatment targets the tumour without exacerbating intracranial pressure.

Tumour location has a large impact on symptoms and treatment possibilities. Tumours in the frontal lobe tend to cause personality changes; tumours in the left temporal lobe can lead to haemianopia and aphasia. Tumours in the cerebral hemispheres can often be resected; those in the brainstem often cannot. For these reasons, it is important to investigate therapeutic options through *in vivo* models that study each tumour type where it would be found in humans.

Although brain tumours have the capacity to divide rapidly, invade locally and metastasize within the CNS, it is very rare for a brain tumour to metastasize outside the CNS. It is not clear why brain tumours do not metastasize outside the CNS, but there are several fields of thought. Aggressive brain tumours, such as AA and GBM, may kill the patient before metastases can be detected. The fact that extraneural metastases have been related to previous surgery or radiation therapy supports this theory, as these patients have an increased length of survival, and thus an increased window for metastases to spread and develop. However, this correlation between surgery and extraneural metastases also supports the theory that brain tumour cells cannot cross the intact BBB. Surgery can give tumours direct access to extrameningeal tissues, bypassing the BBB. Researchers should keep this in mind when designing therapy that involves prolonged BBB disruption, as it is possible that this disruption may allow tumour cells to metastasize extraneurally. The BBB is a selective membrane, blocking the entrance of blood-borne molecules >180 Da and large, complex biological molecules such as cells and viruses from the brain. The theory of the BBB as a CNS tumour blockade is substantiated by the rare observation of brain tumour invasion of blood vessels, but weakened by the common occurrence of extra-CNS tumours metastasizing across the BBB into the CNS. Extra-CNS tumours may be able to bind the markers on the endothelial surface of brain tumour vasculature, whereas CNS tumours cannot bind the markers on the endothelial surface of non-brain target organ vasculature, which may be causing a one-way metastasis tide. This is supported by experiments performed in the 1960s where malignant gliomas were transplanted into subcutaneous tissue as an immunization procedure. The glioma cells grew but did not metastasize, which indicates that primary brain tumours have properties that make them fundamentally different from tumours originating elsewhere in the body and that the BBB is not the sole barrier to extra-CNS metastasis.

Virtually any malignant solid tumour in the body can spread to the CNS and involve the cerebrum, cerebellum, spinal cord, leptomeninges, dura or pituitary. The most common primary tumour sites are the lung and breast, followed by melanoma of the skin, the urinary tract, prostate, kidney, thyroid and gastrointestinal tract. Tumour metastasis to the CNS is present in approximately 20–40% of all cancer patients, and these metastases kill 10 times the number of patients dying from primary brain tumours. For this reason, experimental models for metastases to the brain are very important and are discussed in the animal *in vivo* models section.

Special Considerations for Researchers

The BBB is especially challenging for the clinician to overcome in the treatment of brain tumours. The intact BBB prevents the passage of ionized water-soluble drugs with a molecular mass >180 Da. Many of the currently available effective chemotherapeutic agents have a molecular mass between 200 and 1200 Da. The intact BBB also blocks the passage of large, complex biological molecules such as viruses and cells. Because it is so selective, the BBB must be bypassed or overcome to deliver therapy in a manner that causes no adverse effects to the neural tissue. Bypassing the BBB to generate an immune response is an important consideration in immunotherapy, as there is no lymphatic system to collect antigen and present it to the immune system. Because the BBB also blocks the entrance of many cells of the immune system to the brain, the immune response is often based solely on microglia cells, the specialized immune cells of the CNS. As a result, it is difficult to generate a strong immune response to CNS malignancies, including tumours. Immune cells can be activated outside the brain through *in vivo* or *in vitro* methods and then given access to the brain through analogous transplantation directly into the brain, or through infusion of the activated cells with concurrent breakdown of the BBB with certain drugs. Some of these techniques are described in the *in vivo* experimental models section.

Another special concern is the sensitivity of the CNS to inflammatory responses. Inflammatory responses cause swelling through the accumulation of inflammatory cells and fluid in the damaged site, and this swelling causes increased intracranial pressure and potentially herniation of the brain. The BBB normally protects the brain from harmful inflammation, but therapies that use cytokines, radiation or drugs can reduce the protective nature of the BBB. One particularly troublesome inflammatory process that can be induced through therapy is experimental allergic encephalomyelitis (EAE). The brain is normally protected from developing EAE because there is no route for antigens from ordinary brain parenchyma to be presented to the peripheral immune system. However, EAE can be induced through the use of whole tumour cell lysates in vaccines or therapy models because of the

normal and fetal CNS antigens expressed on glioma cell lines and brain tumour tissues. This has been shown experimentally in the ability of human GBM tissue to induce lethal EAE in nonhuman primates. Researchers must be aware of this potentially fatal complication and test their therapies thoroughly.

For work with animal models for CNS malignancies, the safety evaluation of a treatment or therapy for humans should include testing in two animal models. One of the animal models should be susceptible to EAE (e.g. Lewis rats) to permit more extensive evaluation of the safety of the experimental immunotherapeutic protocol, including the potentially detrimental autoimmune responses that could be stimulated by a cell-mediated immunotherapy protocol. Formally evaluating the risk of immunologically based therapies using this model should be considered in the preclinical testing of all such protocols. Additional animal models and their susceptibilities to EAE are discussed in the *in vivo* animal model section.

PRINCIPLES OF EXPERIMENTAL MODEL APPLICATIONS

Many experimental models for brain tumour research exist. First, human tumour cell models are discussed, including methods for establishing cell lines and for growing them in nude animals for *in vivo* studies. Second, animal tumour cell models are discussed, with a focus on the strengths and weaknesses of each model. Finally, the rapidly expanding field of transgenic and knockout animals and their applicability to brain tumour research is covered.

Human Brain Tumours

Human brain tumours can be studied *in vitro* by obtaining specimens from resections or biopsies or by using established cell lines. In addition, human CNS tumour xenografts are grown *in vivo* in laboratory animals.

Establishment of Cell Lines

Human glial tumour cells suitable for *in vitro* culture are often obtained from viable areas of gross tumour resections, but specimens may also be obtained through an ultrasonic aspirator, open or stereotactic biopsies or from CSF. Specific methods of growing explanted human CNS tumour cells have been described by Liwnicz *et al*. (1986), Onda *et al*. (1988) and Bjerkvig *et al*. (1990).

One major dilemma of primary explantation cell culture is determining whether the growing cells are malignant or have been overgrown by contaminating normal cellular elements. Our laboratory confirms the neoplastic genotype through formal karyotype analysis supplemented with DNA ploidy analysis, fluorescent *in situ* hybridization to determine loss of chromosome 10 and flow cytometry analysis using specific antibodies recognizing one of the EGFR mutations commonly found on human gliomas. Less specific techniques of selecting for neoplastic cells involve growing the newly explanted cells without CO_2, at cooler temperatures (34 °C), in nutrient-poor media, or in a spinner flask to select for cells capable of generating spheroids. One of the most reliable methods for proving that a population of cells is neoplastic is to implant them into an immunodeficient rodent and confirm their tumorigenicity by the assessment of progressive growth and serial transplantation. This method requires a large number of initial cells, sometimes $>5 \times 10^8$ cells, and the period until tumour growth may be as long as 6 months.

Previously Established Lines

A number of widely used human malignant glioma lines, such as U87MG, U105MG, U118MG, U138MG, U178MG, U251MG, U343MG, U373MG, U410MG and U1231MG, were derived by Ponten and Westermark at the University of Uppsala in Sweden. Many other gliomas are available from brain tumour research laboratories at academic sites such as Duke University. The medulloblastoma lines derived and characterized at Duke University and those derived by Tamura *et al*. (1989) demonstrate true medulloblastoma profiles with neuronal differentiation and positive staining for synaptophysin and neurofilament proteins, and not for GFAP. The Daoy medulloblastoma cell line described by Jacobsen *et al*. (1985) does demonstrate glioma-associated markers, in addition to being positive for synaptophysin. TE-671 was originally designated as a medulloblastoma, but has subsequently been shown to be a rhabdomyosarcoma.

Maintenance

Once a human CNS neoplastic cell line has been established, strict protocols must be established to prevent the development of artifacts. Some potential artifacts that can impact cell culture experiments significantly are (1) contamination with other cell lines carried within the same laboratory, (2) overgrowth with a competing cell type contained in the cell line, such as normal glial cells, endothelial cells, macrophages or cells of leptomeningeal origin, (3) contamination with *Mycoplasma*, a bacterial species that can alter cell growth drastically, and has been the cause of incorrect therapeutic responses, (4) tumour cell differentiation with alteration of certain biological properties including GFAP expression, (5) selection of subpopulations that have an *in vitro* growth advantage or (6) loss of tumorigenic capacity after serial passage *in vitro*, the mechanism of which is unknown. Testing for each of these potential artifacts must be done on a regular basis. *Mycoplasma* contamination is tested with high-performance liquid chromatographic analysis of cell

culture supernatants and is determined by the inability of conditioned media from *Mycoplasma*-free cell lines to convert deoxyadenosine to adenosine or thymidine to thymine. Changes in tumorigenicity may be detected through alterations in *in vitro* morphological and growth characteristics, which can be tested for through microscopic observation and growth curves. Changes in the predominant cell type of a cell line can be detected through karyotype and restriction fragment length polymorphism (RFLP) analysis, expression of GFAP, S100 and other surface antigens, presence of glucose-6-phosphate dehydrogenase isoenzymes and response to cyclic adenosine monophosphate and various standard chemotherapeutic agents. These combined tests create a specific profile for each cell line such that contamination or changes over time can be detected.

An alternative to standard cell culture techniques has been developed to reproduce more accurately the three-dimensional architectural pattern of human gliomas *in situ*. This technique uses multicellular spheroids prepared either directly from biopsies digested with collagenase and plated onto agar-coated flasks or from established cell lines. The production of extracellular matrix glycoproteins is promoted in this system. This technique is especially applicable in the testing of cytotoxic and cytostatic drugs.

In Vivo Models

Many human CNS tumour xenografts have been grown subcutaneously or intracranially and serially passaged in immunodeficient animals. These xenografts have primarily been GBM, in addition to some medulloblastomas, a few ependymomas and a myxopapillary ependymoma. Xenografts have also been grown in the intrathecal space of the rat, providing a model of neoplastic meningitis. The xenografts usually maintain the morphological features of the original tumour, including reactivity for GFAP and expression of tumour-associated antigens. The morphological features that characterize human GBMs, including necrosis, can be seen in these xenografts, although endothelial proliferation is less common. As in serial passage *in vitro*, serial passage *in vivo* may influence the subpopulation dynamics in the tumour, possibly creating a tumour that is histologically and biologically different from the initial tumour. As a result, it is important to confirm routinely that the tumours growing in the animals are the same as the original tumour, and not a differentiated tumour or a murine tumour induced by the presence of the xenografts.

Tumour cells can be given intracranially through stereotactic injection to specific sites in the brain. Tumour cells can also be injected into the carotid artery to set up intracranial growth in a less controlled manner, similar to the haematogenous route that tumour cell metastases take in humans to reach the brain. These intracranial models better simulate the clinical situation in the pattern of tumour growth and the barriers of the brain to therapy. However, the therapeutic efficacy may be evaluated solely on the basis of animal survival, which may not directly correlate with tumour growth. An alternative model is injection of tumour cells subcutaneously, for which evaluation of therapeutic efficacy is based on tumour size. The subcutaneous model can provide an assessment of tumour susceptibility to a specific therapy without the confounding variables that come with the intracranial model – BBB and therapeutic efficacy in the CNS. A common approach is to complete preliminary testing in the subcutaneous model before testing in an intracranial model.

Animal Models

Many transplantable animal brain tumour models are available, each with unique attributes that may be advantageous or disadvantageous for obtaining accurate and clinically relevant answers to specific experimental questions. Animal tumour cell lines can be used to study CNS malignancies *in vitro* and *in vivo*. The use of animal cell lines provides a more natural model for *in vivo* experiments, as the animals have functional immune systems and are receiving allogenic transplants apposed to the immunodeficient mice used in human xenotransplantation studies. The ideal animal glial tumour model is a tumour that (1) arises spontaneously or is derived from a tumour that arose spontaneously, (2) displays evidence of glial differentiation histologically and at a molecular level, (3) is biologically and genetically stable and (4) requires few tumour cells to initiate (a) predictable tumour growth *in vitro* and (b) reproducible and fatal growth intracerebrally and subcutaneously *in vivo* in syngeneic animals.

Spontaneous Brain Tumour Models

Spontaneous CNS tumours are not unique to humans, but they are more prevalent in humans than in lower species. Therefore, one of the main difficulties in the use of animal models is to overcome the relatively rare incidence of spontaneous tumours in most animals, even in older animals. The prevalence of spontaneous CNS tumours is reported to be as low as 0.01%. However, some spontaneous tumours have been isolated and established as cell lines. The most widely used spontaneous tumours that have been found in inbred rodent strains and established as cell lines include spontaneous gliomas (astrocytomas) from VM mice, neuroblastoma from A/J mice and melanoma from C57BL/6 mice. No spontaneous brain tumour models in rats have been established. The major limitation to the use of mouse models in the evaluation of immunotherapy protocols is the resistance to induction of EAE. Mice in general, and particularly mouse strains such as C57BL/6 and VM/Dk with the H-2^b MHC background, are notoriously resistant to the induction of EAE.

VM Gliomas

A number of transplantable cell lines that are uniformly tumorigenic in syngeneic VM mice, both intracranially and subcutaneously, that appear to be glial in origin to a number of observers, and that express glial markers such as GFAP and glutamine synthetase (Sampson *et al*., 1997) have subsequently been derived from VM gliomas. Two of these cell lines have been extensively characterized. The SMA-497 cell line has been extensively studied in chemotherapeutic assays both *in vitro* as multicellular tumour spheroids and *in vivo* in syngeneic VM mice. The SMA-560 murine astrocytoma cell line is not a cloned population; thus, it maintains representation of some of the heterogeneity of the original tumour and has recently been shown to secrete biologically active TGF-β, an immunosuppressant secreted by many human malignant gliomas (Sampson *et al*., 1997). Because these tumours arose spontaneously in an inbred mouse strain, are highly tumorigenic, are clearly of an astrocytic lineage and secrete an immunosuppressant commonly secreted by human gliomas, they serve as an excellent animal model of human gliomas, especially when immunotherapeutic protocols are being assessed.

4C8 Glioma

A more recently established glioma cell line is the 4C8 glioma. This cell line was derived from a spontaneous glioma-like tumour that arose in mice transgenically modified to express the neu oncogene (c-*neu*) under the transcriptional control of the myelin basic protein (MBP) gene promoter. 4C8 has been used to establish subcutaneous and intracranial tumours in syngeneic B6D2F1 mice.

C1300 Neuroblastoma

Another spontaneously arising tumour that may be a useful model for nervous system tumours is the C1300 neuroblastoma that arose in the abdominal cavity of an inbred A/J mouse. This cell line has a typical neuroblastic appearance and can be grown intraperitoneally or subcutaneously in syngeneic A/J mice.

B16 Melanoma

The B16 melanoma is syngeneic in C57BL/6 mice. Because of its spontaneous origin and wide use, the B16 melanoma is a practical model of secondary tumours that grow within the nervous system. B16 melanoma is also valuable as a primary CNS tumour model because it arises from the same cells: neural crest cells. Characteristics of its growth and therapy have been well established outside the CNS, making it an easy model with which to work.

Induced Brain Tumour Models

An alternative to spontaneous tumour models is induced animal tumours. Physical, viral or chemical agents can induce nervous system tumours in experimental animals. Although therapeutic ionizing radiation is accepted as the only neurocarcinogen in humans, viruses and chemicals have been successful at inducing nervous system tumours in various species of experimental animals. Although these models were initially directed at uncovering the aetiology of human brain tumours, the rapid progress that has been made in the last decade in understanding the molecular basis of human nervous system tumours directly may mean that these models will not fulfil their promise in this area. Many of them may remain useful, however, in evaluating novel approaches to brain tumour therapeutics.

Physical Agents

Ionizing and ultraviolet radiation, low-frequency electric and magnetic fields and hyperthermia have been considered potential physical neurocarcinogens. Ionizing radiation is the only established human neurocarcinogen, but radiation has been used only infrequently for the induction of nervous system tumours in animals. Rats have been susceptible to the induction of GBMs after the direct intracerebral implantation of ^{60}Co, transplacental intravenous administration of ^{32}P or total body irradiation. Total body irradiation was also successful in inducing tumours in Rhesus monkeys (*Macaca mulatta*) that possess all the histopathological features of human GBM. Unfortunately, none of these tumours have been established appropriately for use in an experimental transplantation model.

Viral Agents

The foundation for the discovery of proto-oncogenes and tumour-suppressor genes can be attributed to early studies in viral carcinogenesis. Vázquez-López (1936), using cell-free extracts of the Rous sarcoma to induce intracranial sarcomas in chickens, is generally credited with the first induction of an intracranial tumour by a virus. Intracranial injection of a number of different viruses has since been used to induce a large variety of nervous system tumours in lower animals and nonhuman primates. In addition to the large number of positive yields, the method has the advantage over chemically induced tumour models (discussed in the next section) of producing tumours within a relatively short time, often within a few weeks. Newborn animals show the greatest degree of susceptibility, the yield depending on the type of virus used, the exact age of the animal at the time of inoculation and the precise dose and concentration of the virus injected. A good deal of correlation exists between the type of virus used and the tumour type produced, although this is not absolute: AA and GBM have most often been obtained with the oncogenic RNA viruses such as the avian, murine and simian sarcoma viruses. Choroid plexus papillomas are induced by some of the simian and human types of papovaviruses, and medulloblastomas and pineocytomas by isolates of the human JC papovaviruses. Finally, neuroblastomas and

retinoblastomas have been reliably induced with the human adenovirus type 12.

Oncogenic RNA viruses have been the most serviceable in the development of useful animal models of cerebral tumours. A technique for reproducibly inducing glial tumours with a uniform mortality distribution in large numbers of rats was established by using the avian sarcoma virus (ASV). An important feature of these tumours is that syngeneic rats bearing ASV-induced tumours have exhibited a depression in their cell-mediated immune responses not unlike those found in humans with malignant glial tumours. A permanent cell line, S635 (clone 5), derived from an AA induced in a neonatal female F344 rat by intracerebral inoculation with concentrated Schmidt–Ruppin-D strain of this virus, has provided a useful rodent brain tumour model. Cells from this line cultured *in vitro* and tumours grown *in vivo* from this cell line have a consistent architecture composed of interlaced, elongated cells that express GFAP, even after more than 100 passages *in vitro*. This cell line, *in vitro,* has a population doubling time of 18.82 h and a colony formation efficiency of 90.8%. *In vivo*, these cells are highly tumorigenic, with as few as 100 cells being uniformly lethal within 1 month when injected intracerebrally into syngeneic rats.

Chemical Agents

The early establishment of an association between tobacco snuff use and oral cancers and between chimney sweeping and scrotal cancer laid the foundation for the study of chemicals as carcinogenic agents. Since that time, in an effort to establish models for testing brain tumour therapies, a number of different compounds that consistently induce nervous system tumours in experimental animals have been identified. The relative ease with which nervous system tumours could be selectively obtained in experimental animals after the administration of these agents, particularly polycyclic aromatic hydrocarbons (PAHs) and the *N*-nitroso compounds, resulted in a vigorous exploration of these animal models in various basic and applied aspects of neuro-oncology research. Many of these tumours have been subsequently passaged *in vitro* and *in vivo* and are now widely used to evaluate potential therapeutic agents.

Polycyclic Aromatic Hydrocarbons and Derivatives. PAHs and their derivatives have been used to induce gliomas, meningiomas and malignant schwannomas. PAHs may also be capable of transforming fetal mouse brain cells *in vitro*. Several chemicals from this large group of compounds have been shown to be neurocarcinogens, with 3- and 20-methylcholanthrene (MCA), benzo[*a*]pyrene and dimethylbenz[*a*]anthracene being the most effective and widely studied. Mice, rats and Syrian hamsters are generally the only animal species vulnerable to neuro-oncogenesis from use of these compounds, and tumour induction generally requires direct implantation of these compounds in the form of powder or pellets into the brain. Malignant tumours of the sciatic nerve in mice have also been produced by injecting PAHs locally.

Methylcholanthrene-induced murine tumours

The ependymoblastoma series. A series of murine gliomas were induced in the 1940s and 1950s by direct implantation of MCA pellets intracerebrally. The tumours derived by this technique, especially the Ep, EpA, GL261 and GL26, have become the most widely used models of murine primary intracranial glial neoplasms in the USA and are often referred to as the ependymoblastoma series because of their early histological classification. These tumours can all be grown subcutaneously and intracerebrally *in vivo*, and some have been established as cell lines *in vitro*. The histological picture of these tumours is similar to that of the rare childhood human ependymoblastoma, but these tumours really demonstrate no signs of ependymal differentiation and on closer examination consist only of undifferentiated cells of uncertain origin. Hence, if they model human gliomas at all, at best they represent a rare form of childhood glioma and not the more common malignant astrocytic tumours generally seen in humans.

Despite their extensive use, these tumours have been poorly characterized, and their derivation is not clearly documented. For example, induction of the Ep tumour is generally attributed to Zimmerman and Arnold (1941), who injected 20-MCA pellets intracranially in male C3H mice. Although a number of tumours of different histologies were obtained by using this technique, none were initially classified as ependymoblastoma. Apparently, one of the mixed gliomas induced in these experiments emerged with the histological appearance of ependymoblastoma after serial subcutaneous passage. Subsequent researchers, using the Ep cell line from Zimmerman and referencing this original paper (Zimmerman and Arnold, 1941), usually transplanted this tumour into C57BL/6 mice. C57BL/6 mice carry the H-2^b MHC haplotype and are therefore incompatible with the H-2^k MHC haplotype of C3H mice from which the tumour was originally derived. Even Zimmerman himself, when referring to this ependymoblastoma, references this original paper (Zimmerman and Arnold, 1941), which used C3H mice (H-2^k) exclusively, but states that the tumour has been propagated in C57BL/6 mice (H-2^b). The EpA tumour, a subline of the Ep tumour, suffers from the same problems. Although the details are unclear, the Glioma 261 cell line (GL261) also appears to have been induced in male C3H mice (H-2^k) by the implantation of 20-MCA pellets, yet many authors transplant this tumour into C57BL/6 mice (H-2^b). Similarly, the derivation of the Glioma 26 (GL26) tumour, also induced by Zimmerman and Arnold, cannot be deduced accurately from the published literature. However, reports of the presence of H-2^b MHC class I and II molecules on the cell lines of the ependymoblastoma series suggest that they may have indeed been derived from C57BL mice.

The failure to document stringently that the tumour lines used by a particular investigator are syngeneic with the inbred mouse strain used will allow potential histocompatibility differences to influence experimental results greatly. Overall, these findings suggest that extensive characterization of cell lines within the ependymoblastoma series is needed prior to any meaningful experimentation and that favourable responses to various therapeutic modalities using these cell lines must, therefore, always be interpreted with caution and preferably confirmed in a second experimental system. Finally, because all these tumours did not arise spontaneously but were induced artificially with chemicals, these tumours cannot, by virtue of their origin, accurately portray the intrinsic qualities of spontaneous human tumours.

L1210 Leukaemia. L1210 leukaemia was induced by the application of 0.2% MCA in diethyl ether cutaneously on an inbred female DBA mouse (subline 212) and may serve as a very useful model of tumours metastatic to the nervous system. After systemic injection, these leukaemic cells disseminate and infiltrate peripheral blood, lymph nodes, liver, spleen and brain. The tumour grows primarily along the leptomeninges as in humans with CNS leukaemia. L1210 leukaemia has been used to demonstrate the efficacy of a number of chemotherapeutic agents and has been extensively used to demonstrate the protective barrier that the intracranial compartment provides against many of these agents.

N-Nitroso Compounds and Other Alkylating Agents. Some alkylating agents, especially the nitrosourea derivatives, have a striking neurotropicity and are more powerful and more reliable inducers of experimental nervous system tumours than are PAHs. Although a large number of different alkylating agents have been employed for inducing nervous system neoplasms in experimental animals, *N*-methyl-*N*-nitrosourea (MNU) and especially *N*-ethyl-*N*-nitrosourea (ENU) remain the most widely used of these compounds, and they are particularly effective in rats. Importantly, nonhuman primates are also susceptible to the neurocarcinogenic action of these substances when exposed transplacentally. Mice, however, are relatively resistant to the neuro-oncogenic action of these compounds, so most murine brain tumour models have been derived from PAH-induced tumours. The relative activity of MNU or ENU in the rat is highly dependent on the age of the animal at the time of injection. MNU is the most active neuro-oncogenic substance in adult animals, whereas the higher alkyl homologues, including ENU, are significantly less effective in adult rats. In very young animals, on the other hand, the relative oncogenicity of MNU and ENU is curiously reversed, and a very high incidence of nervous system tumours is induced by a single subcutaneous or intracerebral injection of ENU in newborn animals. This marked susceptibility of younger animals to tumour induction with ENU provided a convenient transplacental approach to nervous system tumour induction. In this protocol, a single low-dose intravenous injection of ENU is given to gravid rats on the fifteenth or sixteenth day of gestation, which produces tumours in virtually 100% of the offspring, with a latency period of approximately 300 days. In addition to brain tumours, tumours of the trigeminal nerve, spinal cord and peripheral nervous system (PNS) can also be induced.

Nitrosourea administration usually results in multiple subependymal tumours that morphologically resemble mixed gliomas, usually with a mixture of oligodendroglia and astrocytes, which may involve the cerebrum, cerebellum or spinal cord. Typical oligodendrogliomas, which more frequently appear in the hemispheric subcortical white matter, and also typical astrocytomas are produced with these compounds, but these gliomas hardly ever show the typical histological features of GBMs. PNS tumours also frequently develop after administration of the *N*-nitroso compounds and are generally anaplastic schwannomas. The cranial nerve roots within the cerebellopontine angle and the spinal nerve roots are most frequently involved.

Most of the commonly used rat glioma cell lines have been induced by systemic or transplacental injection of nitrosoureas. The 9L gliosarcoma and the C6 glioma are the most commonly used cell lines derived by these techniques.

9L Gliosarcoma. The 9L gliosarcoma cell line was derived from a tumour originally induced in inbred male Fischer 344 (F344) rats by weekly intravenous injections of 5 mg/kg MNU. It has acquired an extensive sarcomatous component, however, after serial passage, and differs considerably from the original tumour, which had the histological appearance of an AA. These changes demonstrate the phenotypic and genetic instability associated with this cell line and sublines derived from it. Thus, this cell line should be obtained from the original source when possible, and the source of any sublines should always be identified. Favourable therapeutic outcomes with this tumour must always be viewed in the light of its genetic instability, which potentially could produce confounding immune responses against this tumour, even in theoretically syngeneic F344 rats.

C6 Glioma. The C6 glioma is a very frequently used experimental model of nervous system neoplasia. However, it is a completely inappropriate model whenever histoincompatibility may alter experimental results, which is almost always. The major limitation to using the C6 cell line is its uncertain genetic background. The reference usually cited for the induction of this cell line (Benda *et al.*, 1968) does not report the rat strain from which the tumour was derived. However, Figure 1 in that paper is identical with Figure 10 in a subsequent paper by the same authors (Benda *et al.*, 1971), which indicates that the C6 glioma was induced by weekly intravenous injections of 5 mg/kg MNU in random-bred male Wistar–Furth rats. Because this cell line was derived in

a randomly bred rat, no syngeneic host exists for this cell line. It has also been discovered recently that the C6 cell line carries an MHC haplotype different from the RT1.A^d MHC haplotype found on BDIX and BDX rats, a strain often used in experiments with the C6 cell line. Therefore, therapeutic responses determined with this cell line in immunocompetent rats must always be evaluated in the context of the allogeneic responses that may be stimulated against this line, even in inbred Wistar–Furth rats, which clearly differ genetically from the cell line itself. Thus, although the C6 cell line remains a good model for *in vitro* glioma studies, for *in vivo* studies, especially ones those where immunological responses may play an important role, results should always be confirmed in more appropriate model systems.

CNS-1 Glioma. A more recent addition to the wide variety of nitrosourea-induced rat glioma cell lines, the CNS-1 cell line, has one particular feature that may make it extremely useful. The CNS-1 cell line was induced in inbred Lewis rats by weekly injection of MNU. Because Lewis rats are highly susceptible to the induction of EAE, this cell line permits more extensive evaluation of the safety of experimental immunotherapeutic protocols, including the potentially detrimental autoimmune responses that could be stimulated by cell-mediated immunotherapy protocols. In light of the documented expression of normal and fetal CNS antigens on glioma cell lines and brain tumour tissues and the ability of human GBM tissue to induce lethal EAE in non-human primates, consideration should be given to evaluating formally the risk of immunologically based therapies using this model in the preclinical testing of all such protocols.

Transgenic and Knockout Models

The development of gene targeting with the ability to alter or modify genes has opened new doors for neuro-oncology researchers. Most of the original neuro-oncology experiments with transgenic animals were designed to study the role of different genes in brain tumorigenesis. Common genes targeted in this research include *p53*, *Rb* and *Nf1* (**Table 2**). The development of transgenic mice using tissue-specific promoters such as GFAP and MBP has also been extensively used to study neural tumour formation. The combination of transgene expression in knockout mice has been used in order to study the effects of multiple alterations on tumorigenesis, thereby mimicking the process of tumour formation more accurately. All of these experiments and their results are discussed below.

Knockout mice are created through homologous recombination to express only one copy of a gene (heterozygous gene deletion) or no copies of the gene (null; homozygous gene deletion). *Ptc*, *Rb*, *p53* and *Nf1* are four genes that have been examined in knockout mice for their effects on brain tumorigenesis. Homozygous inactivation of *Ptc* in mice is fatal during embryogenesis, whereas heterozygous inactivation of *Ptc* in mice leads to medulloblastoma development. Homozygous *Rb1* knockout mice are not viable and heterozygous *Rb1* knockout mice develop pituitary adenocarcinomas at an increased rate. Homozygous *p53* knockout mice have been shown to accelerate tumorigenesis, developing lymphomas, sarcomas and gliomas at a higher rate. These mice have a decrease in gamma radiation and doxorubicin-induced apoptosis, causing increased tumorigenesis. Conversely, Jacks *et al*. (1994) demonstrated homozygous *p53* knockout mice not to develop brain tumours, and heterozygous *p53* knockout mice to develop only rare ependymomas. Homozygous *Nf1* knockout mice are predisposed to tumour formation. Phaeochromocytoma is a tumour derived from the adrenal medulla (neural crest derived tissue) and is one of the predominate tumour types found in *Nf1* null mice.

Yin *et al*. (1997) used *Rb* knockout mice to study the ability of *Bax* to act as a tumour suppressor in the *p53*-mediated apoptosis. Through taking mice with *Rb* inactivation in brain epithelium and crossing them with *Bax*- and/or *p53*-deficient mice, mice were created that lacked *Rb* and *Bax* expression in brain epithelium. These mice developed tumours at a higher rate than *Rb* knockout mice alone, which indicates that *Bax* is required for a full *p53*-mediated response, and therefore that *Bax* acts as a tumour suppressor.

The same *Rb* knockout model was used by Liao *et al*. (1999) to study the role of *Atm* in *p53*-dependent apoptosis and tumour suppression. Mice with *Rb* inactivation in brain epithelium were crossed with *Atm*-deficient mice to create mice that lacked *Rb* and *Atm* expression in brain epithelium. In this study, *p53*-dependent apoptosis, transactivation and tumour suppression were unaffected by *Atm* deficiency, which indicates that signalling in the DNA damage pathway is distinct from that in the oncogene-induced pathway. It was further shown that *Atm* deficiency has no effect on tumorigenesis and tumour progression in this model.

Marino *et al*. (2000) induced medulloblastomas in *p53* null mice by inactivating *Rb* in the external granular layer cells of the cerebellum. Because *p53* homozygous knockout (null) mice do not develop brain tumours, this research indicates that the loss of *Rb* function is essential for medulloblastoma development.

Transgenic mouse models are typically made to express an additional gene ubiquitously under transcriptional control of a gene expressed only in the target tissue. The glial cell specific GFAP promoter and the MBP, expressed only in myelinating cells, have been used as anchor genes for the introduction of transgenes in the nervous system. Through this technique, ERBB2 and the SV 40 large T antigen have been introduced specifically into murine nervous systems. Transgenic ERBB2 mice develop oligodendrogliomas, and transgenic SV 40 large T antigen mice develop anaplastic astrocytomas.

Table 2 Transgenic and knockout mouse models

Gene	Expression	Tumour type	Reference
Ptc	Heterozygous (+/−) knockout	Medulloblastoma	Goodrich *et al.* (1997); Zurawel *et al.* (2000); Wetmore *et al.* (2000)
Rb1	Heterozygous (+/−) knockout	Pituitary carcinoma	Jacks *et al.* (1992)
Rb1 and *bax*	Rb1 (+/−) and Bax (−/−) knockout	Pituitary carcinoma	Yin *et al.* (1997)
Rb1 and *Atm*	Rb1 (+/−) and Atm (−/−) knockout	Pituitary carcinoma	Liao *et al.* (1999)
p53	Homozygous (−/−) knockout	None	Jacks *et al.* (1994)
	Heterozygous (+/−) knockout	Ependymoma	Jacks *et al.* (1994)
p53 and *Rb1*	TP53 (−/−) knockout and Rb1 (−/−) inactivation	Medulloblastoma	Marino *et al.* (2000)
Nf1	Homozygous (−/−) knockout	Phaeochromocytoma	Jacks *et al.* (1994)
EGFR, mutant	Glial cells	Glioma	Holland (2000)
v-src	GFAP-expressing cells	Astrocytoma, GBM	Theurillat *et al.* (1999)
erbB2	MBP-expressing cells	Oligodendroglioma	Hayes *et al.* (1992)
SV 40 large T antigen	GFAP-expressing cells	Anaplastic astrocytoma	Danks *et al.* (1995)
	Tyrosine hydroxylase promoter–enhancer	PNET	Suri *et al.* (1993)
	Tryptophan hydroxylase promoter	Pineal tumours	Son *et al.* (1996)
	MSV enhancer–natural promoter	PNET	Korf *et al.* (1990); Theuring *et al.* (1990); Gotz *et al.* (1992)
	LH β-subunit promoter	PNET; retinoblastoma	Windle *et al.* (1990a); Marcus *et al.* (1991)
	Interphotoreceptor retinoid-binding protein promoter	PNET	al-Ubaidi *et al.* (1992)
	PNMT promoter	Retinoblastoma	Baetge *et al.* (1988); Hammang *et al.* (1993)
	Schwann cell-specific P_0	Schwannoma	Messing *et al.* (1994)
	LVP or SV40 enhancer–promoter	Choroid plexus tumour	Chen and Van Dyke (1991)
	CF transmembrane conductance regulator–promoter	Malignant intraventricular ependymomas and choroid plexus carcinomas	Perraud *et al.* (1992)
	GnRH promoter	Neuroblastoma	Mellon *et al.* (1990); Radovick *et al.* (1991)
	Olfactory marker protein regulatory elements	Neuroblastoma	Largent *et al.* (1993)
	Glycoprotein hormone α-subunit promoter–enhancer	Gonadotrope anterior pituitary adenomas	Windle *et al.* (1990b)
SV40 early region	Metallothionein promoter	Choroid plexus carcinoma	Brinster *et al.* (1984); Palmiter *et al.* (1985)
JCV (T-ag)	SV 40 regulatory region	Choroid plexus papilloma	Ressetar *et al.* (1993)
JCV (Mad-1)	JCV early region promoter	Neuroblastoma	Small *et al.* (1986)
JCV (CY)	JCV early region promoter	Medulloblastoma, PNET	Krynska *et al.* (1999)
Adenovirus E1A and E1B	MMTV LTR	Neuroblastoma	Koike *et al.* (1990)
Polyomavirus large T antigen	Polyoma early region promoter	Pituitary adenoma	Helseth *et al.* (1995)

Table 2 (Continued)

Gene	Expression	Tumour type	Reference
Polyomavirus middle T antigen	Polyoma early region promoter	Pituitary adenoma	Bautch *et al.* (1987)
	Thymidine kinase promoter	Neuroblastoma	Aguzzi *et al.* (1990)
Papillomavirus 16 *E6*, *E7* and *ORF*	Human β-actin promoter	Ependymoma, pituitary carcinoma, choroid plexus carcinoma	Arbeit *et al.* (1993)
HTLV-1 TAT	HTLV-1 LTR	Peripheral nerve sheath tumours	Hinrichs *et al.* (1987)
MYC and *TAX*	HTLV-1 LTR	Brain tumours (not specified)	Benvenisty *et al.* (1992)

Table 3 Rodent brain tumour models[a]

Species	Cell lines	Syngeneic strain	Induction technique	Histological type	GFAP staining
Mouse	P560	VM/Dk	Spontaneous[b]	Anaplastic astrocytoma	++
Mouse	GL261	Unclear	MCA	Ependymoblastoma	No data
Mouse	GL26	Unclear	MCA	Ependymoblastoma	+
Mouse	Ep, EpA	Unclear	MCA	Ependymoblastoma	No data
Mouse	L1210	DBA	MCA	Leukaemia	N/A
Mouse	B16	C57BL/6	Spontaneous	Melanoma	N/A
Rat	S635-clone 5	Inbred CD Fischer 344	ASV (Schmidt–Ruppin)	Anaplastic astrocytoma	++
Rat	9L	Inbred CD Fischer 344	MNU	Gliosarcoma	−
Rat	RG-2	Inbred CD Fischer 344	ENU	Glioma	−[c]
Rat	CNS-1	Inbred Lewis	MNU	Glioma	+
Rat	C6	None (random-bred Wistar–Furth)	MNU	Glioma	+[d]

[a]Abbreviations: GFAP, glial fibrillary acidic protein; MNU, *N*-methyl-*N*-nitrosourea; ENU, *N*-ethyl-*N*-nitrosourea; MCA, 3- and 20-methylcholanthrene; N/A, not applicable.
[b]VM/Dk mice have a high incidence of spontaneous tumours. These lines were derived from spontaneous astrocytomas in this mouse strain.
[c]Some GFAP-positive cells have been noted in intracerebral xenografts, but these could be reactive cells only.
[d]Inducible by dbcAMP, norepinephrine and organ culture.

The GFAP-v-src transgenic mouse model developed by Weissenberger *et al*. (1997) was used to study the triggering of angiogenesis during tumour progression. GFAP-v-src mice develop low-grade astrocytomas early in their lives. These low-grade tumours develop into highly malignant, hypervascularized GBM. By studying VEGF levels at different tumour stages, Theurillat *et al*. (1999) found that expression of angiogenic signals is an early event during tumour progression and precedes hypervascularization. Because of the similarity in tumour progression between GFAP-v-src and human gliomas, these transgenic mice may be an excellent model for anti-angiogenic research.

Holland (2000) created transgenic EGFR mutant mice. These mice express a constitutively active, mutant form of EGFR in cells of glial lineage. The presence of the mutant EGFR induces glioma-like lesions, with characteristics of increased cell density, vascular proliferation and GFAP and nestin expression. These transgenic mice will allow for testing of target-specific therapies against EGFR mutants in a model that is very similar to glioma development in humans.

Summary

A number of transplantable animal brain tumour models are available (**Table 3**), each with unique attributes that may be advantageous or disadvantageous in finding accurate and clinically relevant answers to specific experimental questions. The primary use of experimental models is to predict therapeutic responses in human tumours. In this sense, because very few treatment modalities have shown efficacy against human primary brain tumours clinically, it is difficult to prove that one model is necessarily better than another. Nevertheless, it is probably prudent to avoid basing clinical decisions on information obtained from use of models that can be predicted to provide potentially misleading results, for example, those with inherent histoincompatibilities.

IN VIVO EXPERIMENTAL TECHNIQUES

In vivo experimental techniques include implanting tumours intracranially, use of a metastasis model,

techniques for overcoming the BBB, radiation and treatment through injection and infusion and treatment through viral transfection (gene therapy).

Implanting Tumours Intracranially

Intracranial injection requires the use of a stereotactic frame, such as the Kopf stereotactic frame (Kopf Instruments, Tujunga, CA, USA), and highly accurate syringes, such as Hamilton syringes. Cells for injection should be gently harvested, rinsed well in sterile Dulbecco's phosphate-buffered saline, counted and resuspended in 10% (v/v) methylcellulose in zinc option medium. The number of cells required can be determined by injecting increasing numbers of cells (30, 100, 300, 1000, etc.) until 100% of animals develop tumours. This mixture is loaded into a 250-μL Hamilton syringe for mice or a 500-μL syringe for larger rodents such as rats. The syringe is positioned in the stereotactic frame, the animals are anaesthetized and the needle is positioned at bregma. To implant tumour cells in the right caudate nucleus of mice, the needle is moved 2 mm to the right of bregma and 4 mm below the surface, and 5 μL of cells are given. These coordinates were determined by using *A Stereotaxic Atlas of the Albino Mouse Forebrain* (Slotnick and Leonard, 1975). Coordinates for rats can be determined with *A Stereotaxic Atlas of the Rat Brain* (Pelligrino *et al*., 1979) or *The Rat Brain: a Stereotaxic Atlas of the Forebrain and Lower Parts of the Brain Stem* (Konig and Klippel, 1974). Stereotactic atlases for pig, monkey (*Macaca mulatta*, *Macaca fuscata*, *Cebus apella*, *Pan satyrus*, *Saimiri sciureus*), cat, pigeon, golden hamster, chick (*Gallus domesticus*), dog and baboon are also available.

Metastasis Model

To emulate brain metastases in an animal model, researchers have directly injected tumour cells into the brain and also injected tumour cells intracarotid. The direct intracranial injection excludes the possibility of metastases in additional organs, keeping the model simple. The intracarotid method mimics the development of brain metastases through haematogenous spread. The intracarotid model produces metastases in different regions of the brain, according to where the tumour cells grow the best (meninges versus parenchyma).

Cells from brain metastases have a lower growth rate and have a lower metastatic potential than tumour cells from visceral metastases, which indicates that brain metastases originate from a unique subpopulation of cells within the primary neoplasm. Thus, the intracarotid injection of tumour cells may be a more accurate depiction of metastases to the brain because of the added selection step of tumour cells crossing the BBB to enter the brain parenchyma.

Overcoming the BBB

The BBB is a formidable shield against drug or antibody therapy given intravenously, as only very small molecules (<180 Da) can cross the barrier freely. One common method of overcoming the BBB is giving high doses of drug or antibody systemically, such that enough can cross the barrier to be efficacious against malignancy in the CNS. Unfortunately, systemic toxicity is often reached before an effective concentration is achieved in the CNS. This method is also expensive and wasteful because of the amount of drug or antibody needed. Breaking down the BBB immediately before giving the therapy gives greater access to the brain, and therefore decreases the dose, systemic toxicity and price of treatment. Mannitol is the most frequently used BBB disruption agent. Mannitol is a hypertonic sugar solution that is introduced to the internal carotid or vertebral artery circulation immediately before the administration of intra-arterial chemotherapy. In rats, mannitol can be infused cephalad at 25% in 37 °C water via a catheter in the external carotid artery, at a rate of 0.12 mL/s, by using a constant-flow pump. The treatment is then immediately given through the same catheter.

Direct delivery of drugs and antibody to the tumour through stereotactic injection may be a better route. Direct injection requires less drug or antibody and only the tumour and the tissue immediately surrounding the tumour are exposed to the treatment, which eliminates systemic toxicity. The number of times drugs or antibody can be delivered through this route is limited because of the invasiveness of the technique.

The BBB can also hamper the development of an immune response to tumours in the CNS, as there is restricted movement of regular immune system cells into the CNS, and the CNS immune cells (microglia) may have different processes of antigen gathering and recognition. Methods to overcome this difficulty have been explored, and one strategy involves surgically removing part or all of an intracranial tumour, homogenizing the tumour, pulsing antigen-presenting cells (APCs) with the homogenate and injecting the pulsed cells systemically. It has been shown that an efficacious cellular response to CNS tumours can be induced through these pulsed APCs (Heimberger *et al*., 2000a). This demonstrates that these CNS tumours are not incapable of generating an immune response, but that the tumours are shielded from generating a response by their location in the brain.

Radiation

In vivo experiments involving the effects of external radiation on intracranial tumours have been done in rat and mouse models. Radiation should be confined to the head to prevent radiation damage to the rest of the body. This can be done by anaesthetizing the animals and positioning them such that only their heads are exposed

during external beam radiation. Alternatively, for use with wide-field (whole-body) radiation, the animals can be anaesthetized and placed in restraining containers that restrict radiation to specific sites on the body, such as the brain, through shielding with 1 mm of Pb. Drug, cytokine or antibody treatment is often combined with radiation to test for synergy and to determine the appropriate schedule of dosing.

Treatment Through Injection and Infusion

Intracranial tumours can be treated through direct injection of drug, antibody or immune cells into the tumour. Rat and mouse experiments have been done where the use of a stereotactic frame allows for precise targeting of the tumour (Sampson *et al*., 2000). Small amounts of liquid (5–20 μL) can be injected at one time into the tumour. This method has two advantages over systemic treatment of brain tumours. First, it uses 1% of the amount of drug needed for systemic treatment. Second, the drug or antibody does not have to be altered such that it can pass through the BBB, which might possibly alter its efficacy against the tumour at the same time. The disadvantage is that direct injection is very invasive and cannot be done repeatedly.

Infusion has the same advantages as direct injection plus the ability to give larger doses, cover the tumour more extensively and treat over a longer period of time (Heimberger *et al*., 2000b). Infusion involves the surgical implantation of a tube into the centre of the brain tumour and attaching the tube to an infusion pump placed subcutaneously. Infusion of drug is a better method than simple injection, as infusion pushes the fluid through the tumour, giving greater penetration into the tumour and exposing tumour cells to more drug. Infusion pumps can be acquired commercially (ALZET Osmotic Pumps, DURECT, Cupertino, CA, USA) and used for up to 4-week intervals. A long infusion time should be used, unless the drug half-life at body temperature makes it necessary to use a shorter interval, as a longer infusion time will decrease the amount of pressure generated, lessening the amount of drug that is wasted as it is expelled from the brain along the tubing.

Treatment Through Viral Transfection

Gene therapy of brain tumours is a potential therapeutic modality. The four main approaches to gene therapy of brain tumours are (1) transfer of drug susceptibility with a suicide gene, (2) transduction with a toxic gene, (3) transduction with antisense cell-cycle genes and (4) adoptive immunotherapy. Gene therapy is accomplished through recombinant viral vectors that deliver genes specifically into dividing cells. Herpesvirus and adenovirus are two vectors commonly used, and the thymidine kinase gene is a common choice to transfer drug susceptibility. Thymidine kinase transduction makes tumour cells sensitive to the antiviral agent ganciclovir, a nontoxic prodrug. To deliver these viral vectors to the brain tumour, investigators typically use direct stereotactic inoculation into the tumour tissue. Other routes have been investigated, such as injecting the viral vectors intra-arterially after osmotic BBB disruption with mannitol.

The use of suicide genes in the therapy of brain tumours has many positive caveats. Because the brain consists mostly of nondividing cells, there is very little toxicity to nontumour cells. The 'bystander effect' extends the therapeutic effect of the viral vectors by extending cytotoxic effects to untransduced cells. Transfer of genes that repress growth through antiangiogenesis (interferon gamma) is another current field of research. *p53* gene replacement is strongly indicated as a therapeutic modality, and more efficacious modes of delivery are being explored.

Combination gene therapy is a promising concept for brain tumour treatment. By transferring more than one gene into a tumour, a heterogeneous mixture of tumour cells can be hit. This is important in brain tumour therapy because of the known heterogeneity of cells within a single glioma tumour (Bigner *et al*., 1981). The introduction of genes encoding multiple agents with different mechanisms of action through combination gene therapy may induce synergistic effects. Along that same line of thought, combining gene therapy with chemotherapy and radiation will most likely produce a superior response to one that is achievable through gene therapy alone.

INTERPRETATION OF EXPERIMENTAL RESULTS

A substantial amount of research has been done in the brain tumour field to date, and we have greatly advanced our understanding of brain tumour genetic and biological bases. However, little improvement has been made in the clinical outcome of brain tumour patients with highly malignant tumours. One reason for this discrepancy is that there are very few good models for glioma research. And, a general assumption is that unless the therapy works extremely well in a well-established, syngeneic animal model, it probably will not be efficacious against human brain tumours clinically. Another reason for the lag in improving clinical outcomes is that it is difficult to extrapolate positive results in animal models to positive results in human brain tumour patients. Human brain tumour patients are much more complex than any animal or *in vitro* brain tumour model, and it is therefore difficult to interpret experimental findings in animal models as they would apply to human patients.

An additional problem in extrapolating results from animal models to humans is that many enzymes and

proteins differ from species to species. Consequently, a therapy that works through alterations in cellular enzymes or proteins and is effective in mice may have little or no specificity for the same enzymes or proteins in humans, and thus little efficacy.

CONCLUSION

There are a number of brain tumour models currently available, encompassing a wide variety of *in vitro* and *in vivo* systems. Many chemotherapy and immunotherapy agents have been tested experimentally in these systems and subsequently approved for clinical use. The mainstays of current brain tumour therapy are surgical resection, radiation, and chemotherapy. However, more specific and efficacious therapies are now being introduced: gene therapy, tumour-specific antibodies given intrathecally or intratumourally, antiangiogenesis factors and dendritic cells being activated with tumour-specific lysate and reintroduced. These new therapies will be added to the regimen of surgery, radiation and chemotherapy and are expected to improve outcomes.

There are still major issues to be resolved in the clinical treatment of brain tumours, such as safe, effective delivery of genetic material in gene therapy, safe, effective delivery of immunotherapy and chemotherapy and the best method, timing, and combination of surgery with other treatment modalities. To resolve these issues, the interpretation of results in experimental animal models and the extrapolation to the treatment of humans clinically must be worked out further.

REFERENCES

Aguzzi, A., *et al*. (1990). Sympathetic hyperplasia and neuroblastomas in transgenic mice expressing polyoma middle T antigen. *New Biology*, **2**, 533–543.

al-Ubaidi, M. R., *et al*. (1992). Bilateral retinal and brain tumors in transgenic mice expressing simian virus 40 large T antigen under control of the human interphotoreceptor retinoid-binding protein promoter. *Journal of Cell Biology*, **119**, 1681–1687.

Arbeit, J. M., *et al*. (1993). Neuroepithelial carcinomas in mice transgenic with human papillomavirus type 16 E6/E7 ORFs. *American Journal of Pathology*, **142**, 1187–1197.

Baetge, E. E., *et al*. (1988). Transgenic mice express the human phenylethanolamine *N*-methyltransferase gene in adrenal medulla and retina. *Proceedings of the National Academy of Sciences of the USA*, **85**, 3648–3652.

Bautch, V. L., *et al*. (1987). Endothelial cell tumors develop in transgenic mice carrying polyoma virus middle T oncogene. *Cell*, **51**, 529–537.

Benda, P., *et al*. (1968). Differentiated rat glial cell strain in tissue culture. *Science*, **161**, 370–371.

Benda, P., *et al*. (1971). Morphological and immunochemical studies of rat glial tumors and clonal strains propagated in culture. *Journal of Neurosurgery*, **34**, 310–323.

Benvenisty, N., *et al*. (1992). Brain tumours and lymphomas in transgenic mice that carry HTLV-I LTR/c-myc and Ig/tax genes. *Oncogene*, **7**, 2399–2405.

Bigner, D. D., *et al*. (1981). How heterogeneous are gliomas? *Cancer Treatment Reports*, **65**, Suppl. 2, 45–49.

Bjerkvig, R., *et al*. (1990). Multicellular tumor spheroids from human gliomas maintained in organ culture. *Journal of Neurosurgery*, **72**, 463–475.

Brinster, R. L., *et al*. (1984). Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. *Cell*, **37**, 367–379.

Chen, J. D. and Van Dyke, T. (1991). Uniform cell-autonomous tumorigenesis of the choroid plexus by papovavirus large T antigens. *Molecular and Cellular Biology*, **11**, 5968–5976.

Danks, R. A., *et al*. (1995). Transformation of astrocytes in transgenic mice expressing SV40 T antigen under the transcriptional control of the glial fibrillary acidic protein promoter. *Cancer Research*, **55**, 4302–4310.

Goodrich, L. V., *et al*. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science*, **277**, 1109–1113.

Gotz, W., *et al*. (1992). Midline brain tumors in MSV-SV 40-transgenic mice originate from the pineal organ. *Acta Neuropathologia*, **83**, 308–314.

Hammang, J. P., *et al*. (1993). Oncogene expression in retinal horizontal cells of transgenic mice results in a cascade of neurodegeneration. *Neuron*, **10**, 1197–1209.

Hayes, C., *et al*. (1992). Expression of the neu oncogene under the transcriptional control of the myelin basic protein gene in transgenic mice: generation of transformed glial cells. *Journal of Neuroscience Research*, **31**, 175–187.

Heimberger, A. B., *et al*. (2000a). Bone marrow-derived dendritic cells pulsed with tumor homogenate induce immunity against syngeneic intracerebral glioma. *Journal of Neuroimmunology*, **103**, 16–25.

Heimberger, A. B., *et al*. (2000b). Temozolomide delivered by intracerebral microinfusion is safe and efficacious against malignant gliomas. *Clinical Cancer Research*, **6**, 4148–4153.

Helseth, A., *et al*. (1995). Endocrine and metabolic characteristics of polyoma large T transgenic mice that develop ACTH-producing pituitary tumors. *Journal of Neurosurgery*, **82**, 879–885.

Hinrichs, S. H., *et al*. (1987). A transgenic mouse model for human neurofibromatosis. *Science*, **237**, 1340–1343.

Holland, E. C. (2000). A mouse model for glioma: biology, pathology, and therapeutic opportunities. *Toxicology and Pathology*, **28**, 171–177.

Jacks, T., *et al*. (1992). Effects of an Rb mutation in the mouse. *Nature*, **359**, 295–300.

Jacks, T., *et al*. (1994). Tumour predisposition in mice heterozygous for a targeted mutation in Nf1. *Nature Genetics*, **7**, 353–361.

Jacobsen, P. F., *et al*. (1985). Establishment of a human medulloblastoma cell line and its heterotransplantation into nude mice. *Journal of Neuropathology and Experimental Neurology*, **44**, 472–485.

Koike, K., *et al*. (1990). Activation of retrovirus in transgenic mice: association with development of olfactory neuroblastoma. *Journal of Virology*, **64**, 3988–3991.

Konig, J. F. R. and Klippel, R. A. (1974). *The Rat Brain: a Stereotaxic Atlas of the Forebrain and Lower Parts of the Brain Stem* (Williams and Wilkins, Baltimore).

Korf, H. W., *et al*. (1990). S-antigen and rod-opsin immunoreactions in midline brain neoplasms of transgenic mice: similarities to pineal cell tumors and certain medulloblastomas in man. *Journal of Neuropathology and Experimental Neurology*, **49**, 424–437.

Krynska, B., *et al*. (1999). Human ubiquitous JCV(CY) T-antigen gene induces brain tumors in experimental animals. *Oncogene*, **18**, 39–46.

Largent, B. L., *et al*. (1993). Directed expression of an oncogene to the olfactory neuronal lineage in transgenic mice. *Journal of Neuroscience*, **13**, 300–312.

Liao, M. J., *et al*. (1999). Atm is dispensable for p53 apoptosis and tumor suppression triggered by cell cycle dysfunction. *Molecular and Cellular Biology*, **19**, 3095–3102.

Liwnicz, B. H., *et al*. (1986). Continuous human glioma-derived cell lines UC-11MG and UC-302MG. *Journal of Neuro-Oncology*, **3**, 373–385.

Marcus, D. M., *et al*. (1991). Primitive neuroectodermal tumor of the midbrain in a murine model of retinoblastoma. *Investigative Ophthalmology and Visual Science*, **32**, 293–301.

Marino, S., *et al*. (2000). Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum. *Genes and Development*, **14**, 994–1004.

Mellon, P. L., *et al*. (1990). Immortalization of hypothalamic GnRH neurons by genetically targeted tumorigenesis. *Neuron*, **5**, 1–10.

Messing, A., *et al*. (1994). Hypomyelinating peripheral neuropathies and schwannomas in transgenic mice expressing SV40 T-antigen. *Journal of Neuroscience*, **14**, 3533–3539.

Onda, K., *et al*. (1988). Correlation of DNA ploidy and morphological features of human glioma cell cultures with the establishment of cell lines. *Acta Neuropathologia*, **76**, 433–440.

Palmiter, R. D., *et al*. (1985). SV40 enhancer and large-T antigen are instrumental in development of choroid plexus tumours in transgenic mice. *Nature*, **316**, 457–460.

Pellegrino, L. J., *et al*. (1979). *A Stereotaxic Atlas of the Rat Brain*, 2nd edn (Plenum Press, New York).

Perraud, F., *et al*. (1992). The promoter of the human cystic fibrosis transmembrane conductance regulator gene directing SV40 T antigen expression induces malignant proliferation of ependymal cells in transgenic mice. *Oncogene*, **7**, 993–997.

Radovick, S., *et al*. (1991). Migratory arrest of gonadotropin-releasing hormone neurons in transgenic mice. *Proceedings of the National Academy of Sciences of the USA*, **88**, 3402–3406.

Ressetar, H. G., *et al*. (1993). Expression of viral T-antigen in pathological tissues from transgenic mice carrying JC-SV40 chimeric DNAs. *Molecular and Chemical Neuropathology*, **20**, 59–79.

Sampson, J. H. and Bigner, D. D. (1998). Experimental tumors and the evaluation of neurocarcinogens. In: Bigner, D. D., *et al*. (eds), *Russell and Rubinstein's Pathology of Tumors of the Nervous System*, 6th edn 167–230 (Arnold, London).

Sampson, J. H., *et al*. (1997). Characterization of a spontaneous murine astrocytoma and abrogation of its tumorigenicity by cytokine secretion. *Neurosurgery*, **41**, 1365–1373.

Sampson, J. H., *et al*. (2000). Unarmed, tumor-specific monoclonal antibody effectively treats brain tumors. *Proceedings of the National Academy of Sciences of the USA*, **97**, 7503–7508.

Slotnick, B. M. and Leonard, C. M. (1975). *A Stereotaxic Atlas of the Albino Mouse Forebrain*. DHEW Publication No. (ADM) 75–100 (US Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, Rockville, MD).

Small, J. A., *et al*. (1986). Early regions of JC virus and BK virus induce distinct and tissue-specific tumors in transgenic mice. *Proceedings of the National Academy of Sciences of the USA*, **83**, 8288–8292.

Son, J. H., *et al*. (1996). Immortalization of neuroendocrine pinealocytes from transgenic mice by targeted tumorigenesis using the tryptophan hydroxylase promoter. *Molecular Brain Research*, **37**, 32–40.

Suri, C., *et al*. (1993). Catecholaminergic cell lines from the brain and adrenal glands of tyrosine hydroxylase-SV40 T antigen transgenic mice. *Journal of Neuroscience*, **13**, 1280–1291.

Tamura, K., *et al*. (1989). Expression of major histocompatibility complex on human medulloblastoma cells with neuronal differentiation. *Cancer Research*, **49**, 5380–5384.

Theurillat, J. P., *et al*. (1999). Early induction of angiogenetic signals in gliomas of GFAP-v-src transgenic mice. *American Journal of Pathology*, **154**, 581–590.

Theuring, F., *et al*. (1990). Tumorigenesis and eye abnormalities in transgenic mice expressing MSV-SV40 large T-antigen. *Oncogene*, **5**, 225–232.

Vázquez-López, E. (1936). On the growth of Rous sarcoma inoculated into the brain. *American Journal of Cancer*, **26**, 29–55.

Weissenberger, J., *et al*. (1997). Development and malignant progression of astrocytomas in GFAP-v-src transgenic mice. *Oncogene*, **14**, 2005–2013.

Wetmore, C., *et al*. (2000). The normal patched allele is expressed in medulloblastomas from mice with heterozygous germ-line mutation of patched. *Cancer Research*, **60**, 2239–2246.

Windle, J. J., *et al*. (1990a). Retinoblastoma in transgenic mice. *Nature*, **343**, 665–669.

Windle, J. J., *et al*. (1990b). Cell lines of the pituitary gonadotrope lineage derived by targeted oncogenesis in transgenic mice. *Molecular Endocrinology,* **4**, 597–603.

Yin, C., *et al*. (1997). Bax suppresses tumorigenesis and stimulates apoptosis *in vivo*. *Nature*, **385**, 637–640.

Zimmerman, H. M. and Arnold, H. (1941). Experimental brain tumors. I. Tumors produced with methylcholanthrene. *Cancer Research*, **1**, 919–924.

Zurawel, R. H., *et al*. (2000). Evidence that haploinsufficiency of Ptch leads to medulloblastoma in mice. *Genes*, *Chromosomes and Cancer*, **28**, 77–81.

FURTHER READING

Levin, V. A. (1999). Neuro-oncology: an overview. *Archives of Neurology*, **56**, 401–440.

Markert, J. M., *et al*. (2000). Genetically engineered HSV in the treatment of glioma: a review. *Reviews in Medical Virology*, **10**, 17–30.

Parney, I. F., *et al*. (2000). Glioma immunology and immunotherapy. *Neurosurgery*, **46**, 778–791.

Pollack, I. F., *et al*. (2000). Exploitation of immune mechanisms in the treatment of central nervous system cancer. *Seminars in Pediatric Neurology*, **7**, 131–143.

Prados, M. D., *et al*. (1998). Primary central nervous system tumors: advances in knowledge and treatment. *CA: A Cancer Journal for Clinicians*, **48**, 331–360.

Qureshi, N. H. and Chiocca, E. A. (1999). A review of gene therapy for the treatment of central nervous system tumors. *Critical Reviews in Oncogenesis*, **10**, 261–274.

Sampson, J. H. and Bigner, D. D. (1998). Experimental tumors and the evaluation of neurocarcinogens. In: Bigner, D. D., *et al.* (eds). *Russell and Rubinstein's Pathology of Tumors of the Nervous System*, 6th edn 167–230 (Arnold, London).

Chapter 74

Models for Endocrine Cancer: Uterine Carcinoma as an Example

Retha R. Newbold

National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

CONTENTS

- Biological Origin and Tumour Specifics
- Principles of Model Applications
- Potential Problems of Research Models
- Perspectives

BIOLOGICAL ORIGIN AND TUMOUR SPECIFICS

Carcinoma of the uterine endometrium is one of the most frequently diagnosed malignancies of the female reproductive tract and accounts for approximately 34 000 new cases reported each year in the United States (Miller *et al*., 1993). Although the incidence of endometrial cancer varies widely throughout the world, industrialized countries in Europe and North America report incidences 4–5 times higher than in developing countries of Asia, Africa and South America and also Japan; world-wide, approximately 150 000 cases are diagnosed each year, making endometrial neoplasia a leading cancer in women (Harras, 1996). Epidemiological and clinico-pathological evidence suggest that there are two forms of endometrial carcinoma, based on histopathological characteristics, hormone responsiveness, aggressiveness and differentiation of the lesion, and patient history, especially exposure to unopposed oestrogens (Deligdisch and Holinka, 1987; Kurman *et al*., 1994); most endometrial tumours can be categorized as either hormone-responsive and well differentiated (Type I), or hormone unresponsive and poorly differentiated (Type II), with the former neoplasms usually occurring in perimenopausal women and the latter usually occurring in older women (>60 years) (Kurman *et al*., 1994).

Oestrogens are known risk factors for Type I uterine carcinoma in humans, whereas Type II tumours are usually unrelated to oestrogen (Kurman *et al*., 1994). Clinical studies evaluating oestrogen replacement therapy (ERT) in postmenopausal women have been especially useful in establishing the association between treatment with unopposed oestrogens and Type I endometrial cancer (Persson *et al*., 1989). The role of unopposed oestrogen stimulation in endometrial carcinogenesis is so strong that most factors associated with increased oestrogen levels such as chronic anovulation or obesity increase the risk of uterine disease. In contrast, factors that lower the proliferative effects of oestrogen on the endometrium, such as the addition of progesterone or its derivatives (collectively called progestins) to hormone replacement therapies (HPT), lower the risk of endometrial cancer (Kurman *et al*., 1994).

Studies reporting a high incidence of both endometrial carcinoma and breast cancer among mothers, sisters and aunts of patients diagnosed with endometrial cancer suggest a genetic factor in the development of a small number of these lesions (Harras, 1996). Endometrial cancer and breast cancers share some other risk factors such as early age at menarche and late age at menopause. Unlike breast cancer, late age at first birth has not been found to be a risk for endometrial cancer. However, multiple births have been associated with decreased risk of endometrial cancer while women who have never had children, particularly those with a clinical history of infertility, are at greater risk of developing the disease. Reports of women with Stein–Leventhal syndrome, a rare condition characterized by polycystic ovaries, excessive oestrogen production and infertility, have noted a particularly high risk for endometrial carcinoma. Obesity, which is often accompanied by increased levels of endogenous oestrogens, is recognized as another risk factor; in spite of the evidence linking obesity to development of endometrial carcinoma, few studies have investigated the potential role of dietary factors in developing the disease (Levi *et al*., 1993). Surveying all of these reported risk factors suggests a common link between hormonal imbalances especially excess oestrogen production (Miller *et al*., 1993).

An increasing number of reports suggest that women treated with tamoxifen, a therapeutic compound with antioestrogenic (antagonist) activity in the breast but

oestrogenic (agonist) activity in the uterus, may have an increased risk of developing endometrial cancer (Magriples *et al*., 1993; van Leeuwen *et al*., 1994; Fisher *et al*., 1994; Friedl and Jordan, 1994; King, 1995). In fact, the World Health Organisation's International Agency for Research on Cancer (IARC) and the US National Toxicology Program (NTP) published their evaluation of the carcinogenic risks of tamoxifen, supporting the hypothesis that this pharmaceutical drug is associated with increased risk of endometrial cancer (IARC, 1996; NTP, 2000). Further experimental animal studies give additional support to tamoxifen's carcinogenic potential in the rodent uterus (Newbold *et al*., 1997). Although tamoxifen is reported to exert its effects by competing with oestrogen for binding to the oestrogen receptor (ER), thereby blocking the transcriptional activity of the receptor, the exact mechanisms by which it decreases breast cancer yet increases uterine cancer remain unclear. Its complex pharmacological behaviour suggests that tamoxifen's antagonistic or agonistic activities are dose, species, tissue and cell type dependent (Mahfoudi *et al*., 1995). Data continue to accumulate, however, to suggest that various drugs and environmental chemicals with hormonal activity may represent aetiological factors for human endometrial disease, including cancer.

PRINCIPLES OF MODEL APPLICATIONS

Given the documented role of oestrogens in the development of uterine disease and the increasing use of hormones to alleviate symptoms of menopause and other gynaecological problems, it is important to have appropriate *in vivo* experimental animal models to study the ontogeny and progression of uterine carcinoma and the mechanisms associated with this disease. Further, it becomes increasingly important to have an experimental animal model of uterine carcinoma since many environmental chemicals have been reported to possess oestrogenic activity and exposure, therefore, is possibly widespread. Since the carcinogenic potential of these environmental chemicals for the human uterus is unknown, this demands careful attention.

Although carcinogenic responses to hormones have been noted (e.g. liver, testis and pituitary tumours in mice treated with oestrogen), the power of experimental animal models can be greatly enhanced with the development and use of target organ-specific models. However, so far there are few experimental models for which oestrogens can be shown to induce uterine cancer. In fact, adult rodents given long-term treatment with the potent synthetic oestrogen, diethylstilboestrol (DES), did not show a significant increase in uterine tumours (Highman *et al*., 1980). In a long-term feeding study, oestrogenic compounds given to adult mice was not associated with increased incidence of uterine carcinogenesis; endometrial hyperplasia was seen, suggesting oestrogen stimulation, but uterine cancers were lower than 1% after lifetime exposures to high doses (Highman *et al*., 1980). However, rodents, i.e. mice or hamsters treated neonatally on days 1–5 with DES, show a high incidence of uterine cancers (Leavitt *et al*., 1981, 1982; Newbold and McLachlan, 1982; Newbold *et al*., 1990). As an example, **Figure 1** shows the prevalence of uterine tumours in mice following neonatal treatment with DES on days 1–5 (2 μg per pup per day). By 18 months of age, neonatal DES-exposed mice had a 90% incidence of uterine carcinoma (Newbold *et al*., 1990); this increased to >95% in animals older than 24 months. These DES-induced uterine tumours were oestrogen dependent; when DES-treated mice were ovariectomized before puberty, no uterine tumours developed (**Figure 1**). To demonstrate neoplasia, uterine tumours were transplanted and carried in nude mice. The transplanted tissue retained some differentiated uterine gland structure and function and, in addition, required oestrogen supplementation for maintenance (Newbold *et al*., 1990). Another verification of neoplasia was the fact the uterine tumour cells grew on soft agar, an assay commonly used to demonstrate neoplasia potential. Other compounds with oestrogenic activity, i.e. hexoestrol and tetrafluorodiethystilboestrol (TF-DES) also induced a high prevalence of uterine adenocarcinoma following developmental treatment with the compounds (Newbold *et al*., 1990); 17β-oestradiol induced uterine carcinoma but the prevalence of lesions was lower than that induced by DES (Newbold *et al*., 1990). Further, recent studies describe the induction

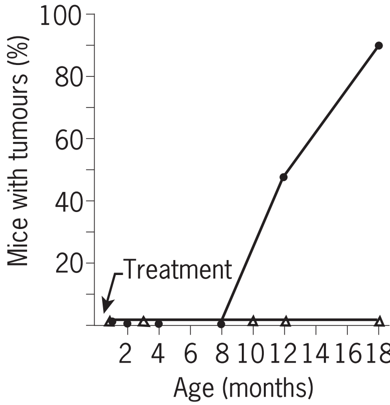
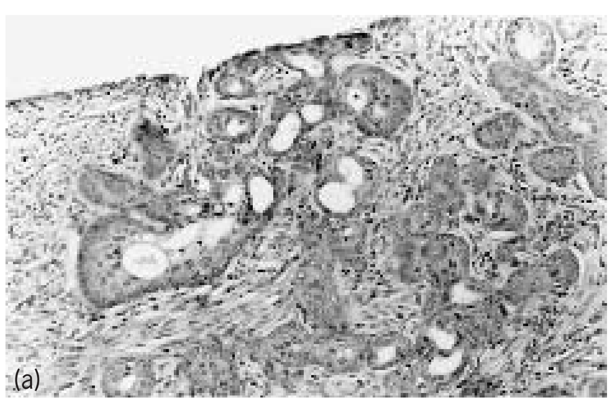


Figure 1 Incidence of uterine carcinoma in developmentally oestrogenized mice. Neonatal mice were treated with DES (2 μg per pup per day) on days 1–5 and killed at various ages. One group of mice was left intact (•) and the other group was prepubertally ovariectomized (△) on day 17. Ovariectomy before puberty blocks the development of uterine carcinoma later in life. In intact mice, the prevalence of tumours was 90% at 18 months. In animals older than 18 months, the lesion progressed to involve more of the uterine structure and the lesions became more undifferentiated. (From Newbold *et al*., 1990, *Cancer Research*, **50**, 7677–7681.)

of uterine adenocarcinoma following neonatal treatment of mice with catechol oestrogens (Newbold and Liehr, 2000); mice treated with 17β-oestradiol and with 17α-ethinyloestradiol had a total uterine tumour incidence of 7% and 43% respectively. 2-Hydroxyoestradiol induced tumours in 12% of the mice but 4-hydroxyoestradiol was the most carcinogenic oestrogen, with a 66% incidence of uterine adenocarcinoma. The high tumour incidence with 4-hydroxyoestradiol supports the postulated role of this metabolite in inducing DNA damage which results in hormone-associated cancers (Newbold and Liehr, 2000).

These oestrogen-induced neoplasms in mice are histologically similar to those reported in humans (Newbold and McLachlan, 1982; Newbold *et al*., 1990; Harras, 1996; Newbold and Liehr, 2000). **Figure 2** shows a representative histological picture of the murine uterine carcinomas associated with neonatal treatment with DES. Cell lines have been derived from these uterine carcinomas (Hebert *et al*., 1992) and biochemical and molecular alterations studied (Endo *et al*., 1994; Risinger *et al*., 1994). Similarly, in the neonatal oestrogen-exposed hamster, uterine carcinoma occurs at a high frequency following developmental exposure to DES (Leavitt *et al*., 1981, 1982). Further studies in the hamster conclude that imbalances in the oestrogen-regulated uterine expression of c-*jun*, c-*fos*, c-*myc*, *bax*, *bcl-2* and *bcl-x* proto-oncogenes probably play a role in the molecular mechanisms by which neonatal DES treatment ultimately induces epithelial neoplasia in the rodent uterus (Zheng and Hendry, 1997). Thus, the oestrogen-treated neonatal rodent provides an excellent model for further studies on uterine carcinoma. (For additional endocrine models see chapter on *Mammary Tumour Induction in Animals as a Model for Human Breast Cancer*.)



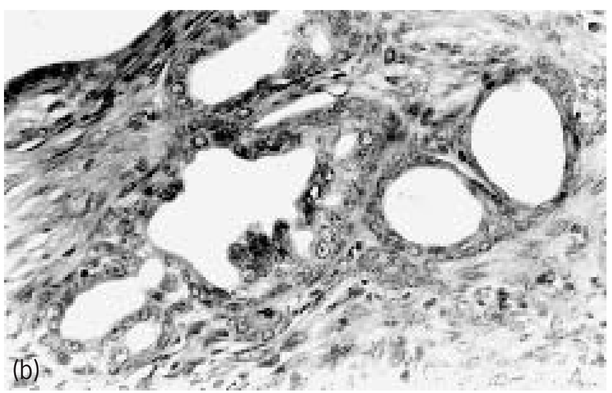


Figure 2 Uterine adenocarcinoma in a 18-month-old mouse treated neonatally on days 1–5 with DES (2 μg per pup per day). (a) Low-power magnification: areas of neoplastic cells form relatively well-defined glandular patterns. (b) High-power magnification: endometrial glands are crowded together with little intervening stroma. (From Newbold *et al*., 1990, *Cancer Research*, **50**, 7677–7681.)

POTENTIAL PROBLEMS OF RESEARCH MODELS

It is interesting that examination of numerous chemical carcinogenesis databases reveals few studies in which uterine tumours in rats or mice are associated with chemical treatment. However, the design of these studies aimed at identifying chemicals for carcinogenic potential starts exposures when animals are mature (approximately 5–8 weeks of age at the earliest). The timing of treatment may not be adequate to detect many uterine carcinogens if exposure of relatively undifferentiated cell types is important for carcinogenic transformation. This may be an important missing factor in terms of demonstrating organ-specific tumours, in particular, with the human uterus representing one of the most sensitive sites for oestrogen-associated neoplasia, and exposure models using undifferentiated cells should be considered. The undifferentiated uterine cells of the neonatal rodent have been demonstrated to be a sensitive model of oestrogen-induced carcinogenesis.

PERSPECTIVES

The newborn rodent is exposed to oestrogens in a period that is especially sensitive to neoplastic transformation; compared with mature tissues, it is relatively deficient in the ‘classical’ uterine epithelial oestrogen receptor (ERα) and mature ER–oestrogen mechanisms. These uterine epithelial cells, however, have the ability to respond to oestrogens by cell division whether through ERα in the stroma (Bigsby *et al*., 1990) or through various growth factor pathways (McLachlan *et al*., 1994). The primate endometrium following menses resembles the neonatal mouse uterine epithelium, i.e. low epithelial ER, ability to respond to oestrogens by cell division and similar growth factor pathways. In other words, the newly re-epithelialized human endometrium may functionally recapitulate the neonatal mouse uterine epithelium on a monthly basis.

Thus, whereas the adult mouse, rat or hamster uterus appears refractory to the carcinogenic potential of

oestrogenic chemicals, developing uterine tissues are highly susceptible to perturbation by oestrogens. Differentiation of the uterine epithelial target cell may render the tissue less sensitive. This suggests an important biological principle which can be tested. Rather than comparing animal models with humans according to age, whenever possible, similar biological stages should be compared (**Figure 3**). The newborn mouse uterine epithelium has low levels of ER, but will respond to oestrogens; it is also the biological period in which oestrogens can neoplastically transform the cell. Likewise, each month following menstruation, the subhuman primate uterus undergoes re-epithelialization; the new epithelium is deficient in ER yet responds to oestrogen (McClellan *et al*., 1986). Therefore, it is possibly susceptible to neoplastic transformation each month. Likewise, the human uterus undergoes the same cyclic re-epithelialization process and fluctuations in ER levels (Brenner and West, 1975; Katzenellenbogen, 1980; Lessey *et al*., 1988; Bern, 1992). Furthermore, the primate uterus at postmenopause is characterized by low levels of ER but a high susceptibility to perturbation by oestrogens (Kurman *et al*., 1994).

In summary, the newborn mouse is not only useful in determining the developmental toxicity of oestrogens (Bern, 1992), but additionally, it has the potential to predict the cellular risk (undifferentiated cellular phenotype) of the adult human uterus. It is feasible that the undifferentiated human endometrium, which has the potential to divide, proliferate and differentiate during each menstrual cycle, puts a woman at increased risk, whereas this risk is

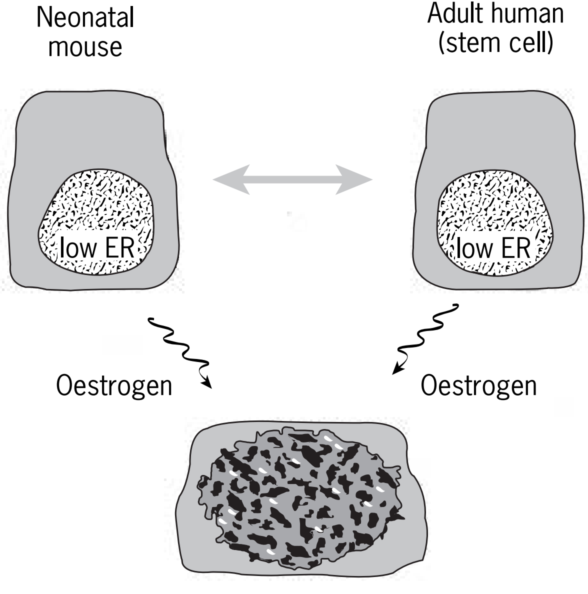


Figure 3 Schematic representation of hormonally responsive uterine epithelial cells from a neonatal mouse and an adult human comparing similar cellular phenotypes.

only apparent for the rodent in early development. Nevertheless, the neonatal treated rodent provides a new model to focus on cellular risks that are important in identifying potential carcinogens and additional studies designed to investigate the mechanism involved in these hormone-associated lesions. Further use of the animal model to investigate treatment and prevention strategies for endocrine-related cancer should follow.

REFERENCES

Bern, H. A. (1992). The development of the role of hormones in development – a double remembrance. *Endocrinology*, **131**, 2037–2038.

Bigsby, R. M., *et al*. (1990). Strain differences in the ontogeny of oestrogen receptors in murine uterine epithelium. *Endocrinology*, **126**, 2592–2596.

Brenner, R. M. and West, N. B. (1975). Hormonal regulation of the reproductive tract in female mammals. *Annual Review of Physiology*, **37**, 273–302.

Deligdisch, L. and Holinka, C. F. (1987). Endometrial carcinoma: two diseases? *Cancer Detection and Prevention*, **10**, 237–246.

Endo, S., *et al*. (1994). Cytogenetic analysis of murine cell lines from diethylstilboestrol-induced uterine endometrial adenocarcinomas. *Cancer Genetics and Cytogenetics*, **74**, 99–103.

Fisher, B., *et al*. (1994). Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. *Journal of the National Cancer Institute*, **86**, 527–537.

Friedl, A. and Jordan, V. C. (1994). What do we know and what don't we know about tamoxifen in the human uterus. *Breast Cancer Research and Treatment*, **31**, 27–39.

Harras, A. (1996). *Cancer Rates and Risks*, 4th edn (National Cancer Institute, Bethesda, MD).

Hebert, C. D., *et al*. (1992). Characterization of murine cell lines from diethylstilboestrol-induced uterine endometrial adenocarcinomas. *In Vitro Cell Development and Biology*, **28A**, 327–336.

Highman, B., *et al*. (1980). Neoplastic and preneoplastic lesions induced in female C3H mice by diets containing diethylstilboestrol or 17-beta-oestradiol. *Journal of Environmental Pathology and Toxicology*, **4**, 81–95.

IARC (1996). *Monographs on the Evaluation of Carcinogenic Risks to Humans: Some Pharmaceutical Drugs*, Vol. 66 (International Agency for Research on Cancer, Lyon).

Katzenellenbogen, B. S. (1980). Dynamics of steroid hormone receptor action. *Annual Review of Physiology*, **42**, 17–35.

King, C. M. (1995). Tamoxifen and the induction of cancer. *Carcinogenesis*, **16**, 1449–1454.

Kurman, R. J., *et al*. (1994). Endometrial carcinoma. In: Kurman, R. J. (ed.) *Blaustein's Pathology of the Female Genital Tract*, 4th edn. 439–486 (Springer, New York).

Leavitt, W. W., *et al*. (1981). Etiology of DES-induced uterine tumours in the Syrian hamster. *Advances in Experimental Medicine and Biology*, **138**, 63–86.

Leavitt, W. W., *et al*. (1982). Antioestrogen and progestin action in diethylstilboestrol-induced endometrial abnormalities in the Syrian hamster. In: Sutherland, R. L. and Jordan, V. C. (eds), *Non-steroidal Antioestrogens. Molecular Pharmacology and Antitumour Activity*. 165–175 (Academic Press, Sydney).

Lessey, B. A., *et al*. (1988). Immunohistochemical analysis of human uterine oestrogen and progesterone receptors throughout the menstrual cycle. *Journal of Clinical Endocrinology and Metabolism*, **67**, 334–340.

Levi, F., *et al*. (1993). Dietary factors and the risk of endometrial cancer. *Cancer*, **71**, 3575–3581.

Magriples, U., *et al*. (1993). High-grade endometrial carcinoma in tamoxifen-treated breast cancer patients. *Journal of Clinical Oncology*, **11**, 485–490.

Mahfoudi, A., *et al*. (1995). Specific mutations in the oestrogen receptor change the properties of antioestrogens to full agonists. *Proceedings of the National Academy of Sciences of the USA*, **92**, 4206–4210.

McClellan, M., *et al*. (1986). Immunocytochemical localization of oestrogen receptors in the macaque endometrium during the luteal–follicular transition. *Endocrinology*, **119**, 2467–2475.

McLachlan, J. A., *et al*. (1994). Control of uterine epithelial growth and differentiation: implications for oestrogen-associated neoplasia. In: Li, E. A. (ed.), *Hormonal Carcinogenesis*. 51–57 (Springer, New York).

Miller, B. A., *et al*. (1993). *SEER Cancer Statistics Review 1973–1990* (National Cancer Institute, Bethesda, MD).

Newbold, R. R. and Liehr, J. G. (2000). Induction of uterine adenocarcinoma in CD-1 mice by catechol oestrogens. *Cancer Research*, **60**, 235–237.

Newbold, R. R. and McLachlan, J. A. (1982). Vaginal adenosis and adenocarcinoma in mice exposed prenatally or neonatally to diethylstilboestrol. *Cancer Research*, **42**, 2003–2011.

Newbold, R. R., *et al*. (1990). Uterine adenocarcinoma in mice following developmental treatment with oestrogens: a model for hormonal carcinogenesis. *Cancer Research*, **50**, 7677–7681.

Newbold, R. R., *et al*. (1997). Uterine carcinoma in mice treated neonatally with tamoxifen. *Carcinogenesis*, **18**, 2293–2298.

NTP (2000). *Report on Carcinogens*, 9th edn (US Department of Health and Human Services, Public Health Service, National Toxicology Program, Research Triangle Park, NC).

Persson, I., *et al*. (1989). Risk of endometrial cancer after treatment with oestrogens alone or in conjunction with progestogens: results of a prospective study. *British Medical Journal*, **298**, 147–151.

Risinger, J. I., *et al*. (1994). Use of representational difference analysis for the identification of *mdm2* oncogene amplification in diethylstilboestrol-induced murine uterine adenocarcinomas. *Molecular Carcinogenesis*, **11**, 13–8.

van Leeuwen, F. E. (1994). Risk of endometrial cancer after tamoxifen treatment of breast cancer. *Lancet*, **343**, 448–452.

Zheng, X. and Hendry, W. J. III (1997). Neonatal diethylstilboestrol treatment alters the oestrogen-regulated expression of both cell proliferation and apoptosis-related proto-oncogenes (c-*jun*, c-*fos*, c-*myc*, *bax*, *bcl-2*, and *bcl-x*) in the hamster uterus. *Cell Growth and Differentiation*, **8**, 425–434.

FURTHER READING

Bern, H. (1992). The fragile fetus. In: Colborn, T. and Clement, C. (eds), *Chemically-induced Alterations in Sexual and Functional Development: the Wildlife/Human Connection* (Princeton Scientific Publishing, Princeton, NJ).

Colborn, T., *et al*. (1996). *Our Stolen Future* (Penguin Books, New York).

Herbst, A. L. and Bern, H. A. (1981). *Developmental Effects of Diethylstilbestrol (DES) in Pregnancy*. (Thieme-Stratton, New York).

Huff, J., *et al*. (1996). *Cellular and Molecular Mechanisms of Hormonal Carcinogenesis: Environmental Influences* (Wiley-Liss, New York).

Newbold, R. (1995). Cellular and molecular effects of developmental exposure to diethylstilbestrol: implications for other environmental estrogens. *Environmental Health Perspectives*, **103** (Suppl. 7), 83–87.

Newbold, R. R. and McLachlan, J. A. (1996). Transplacental hormonal carcinogenesis: diethylstilbestrol as an example. *Progress in Clinical Biological Research*, **394**, 131–147.

Websites

http://www.niehs.nih.gov.
http://www.nci.nih.gov.
http://ehis.niehs.nih.gov.
http://e.hormone.tulane.edu.

Chapter 75

Plain Film Radiography

Anne M. Silas and J. Marc Pipas

Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA

CONTENTS

- Introduction
- Musculoskeletal Neoplasms
- Abdominal Neoplasms
- Lung Neoplasms
- Complications of Cancer – Disease and Therapy
- Conclusion

INTRODUCTION

Initial imaging of the chest, abdomen and musculoskeletal systems has traditionally consisted of plain film radiology for the evaluation of suspected pathology. Plain radiographs are readily available, require relatively basic patient cooperation (i.e. a long breath hold, such as is often necessary for computed tomography (CT) or magnetic resonance imaging (MRI), is not necessary) and are fairly low in cost. On the other hand, the diagnostic yield for the detection and follow-up of cancer using plain film imaging techniques may be low. As a screening modality, plain films are often nonspecific. While the basic supposition of screening is the detection of early-stage disease in asymptomatic individuals, where the intent is to reduce mortality, plain radiographs often lack both sensitivity and specificity, and therefore must be relied on with caution. The aim of this chapter is to review the most common uses of general radiography with plain films in the initial evaluation and follow-up of cancer, and to present the particular areas of real utility and the limitations of this modality.

Chest radiography, in conjunction with sputum analyses, was used in the past to screen for lung cancer with little proven benefit. Traditional chest radiography has typically demonstrated larger parenchymal lesions, but lacked soft tissue contrast necessary to evaluate small or central lesions, or lymphadenopathy in the hilar regions or mediastinum. Evolving digital radiography techniques may increase plain film sensitivity for the detection of cancerous lesions.

Plain film examination of the abdomen is best used to exclude perforation or obstruction. As a screening modality in the setting of mild or nonspecific complaints, abdominal radiography has been shown to yield little diagnostic information. Cross-sectional imaging modalities such as CT and MRI show greater soft tissue resolution and provide more diagnostic detail when evaluating for soft tissue or solid organ abnormalities.

High spatial resolution of radiographs in addition to high contrast of bony detail is beneficial for the plain film evaluation of the bony skeleton. Suspicious findings involving aberration of cortical integrity and morphology can usually be seen accurately. While musculoskeletal plain film surveys are useful for the detection of internal lesion characteristics and matrices, soft tissue and neurovascular involvement cannot be distinguished with this modality. Surveys are considered useful for certain malignancies with a high incidence of bone involvement which may be asymptomatic (e.g. prostate or breast cancer or multiple myeloma). For the most part, musculoskeletal X-rays are most useful when obtained in conjunction with symptoms or signs, or with radionuclide studies.

Multiplanar modalities such as ultrasound, CT and MRI are capable of greater spatial and soft tissue resolution than plain film radiography. As such, these modalities yield more specific diagnostic information than plain films and are essential imaging techniques for initial diagnosis and follow-up of cancer.

This chapter will briefly discuss the role of plain film radiography in the diagnosis and follow-up of lung, abdominal and musculoskeletal neoplasms. There is a limited role in dealing with other neoplasms, such as those of the central nervous or haematopoietic systems, a role confined largely to specific situations which are beyond the scope of this chapter.

MUSCULOSKELETAL NEOPLASMS

Traditional X-ray evaluation has served as the initial imaging tool for the evaluation of bone lesions. X-rays are often the initial diagnostic test, and are focused on a specific symptomatic area. Radiographs provide information regarding the appearance of osseous structures, such as bone density, morphology and cortical integrity. Internal characteristics of benign and malignant lesions such as ossification or chondroid matrices, tumour margins and rate and aggressiveness of growth patterns seen on serial studies may also be assessed (Peabody *et al*., 1998). These radiographic features, along with clinical information such as patient age, provide the basis for formulating the initial differential diagnosis and diagnostic plan (Jaovisidha *et al*., 1998).

Plain film radiographs have to been used to follow and evaluate response to treatment and also to show side effects of treatment and disease such as osteopenia, soft tissue calcification and pathological fractures. Radiographs also complement newer imaging modalities such as MRI by contributing imaging data of the highest spatial resolution to help resolve cross-sectional data that may be nonspecific (Feldman, 2000). While digital radiography techniques have augmented the contrast resolution of the plain film examination, these still fall short of many advanced imaging features of other available modalities such as the multiplanar capability of MRI, MRI evaluation of bone marrow and CT detection of subtle matrix presence and involvement of cancellous bone.

Primary benign bone lesions are characterized at plain film examination by well-demarcated and intact tumour margins and precise zones of transition from involved to uninvolved segments.

Calcification of internal bony tumoural architecture is detectable at plain film examination. Elements that calcify may be of osteoid (bone) or chondroid (cartilage) origin and are referred to as osteoid or chondroid matrices when identified. These findings however, are nonspecific as these features are found in both benign and malignant lesions. Additional characteristics such as patient age and lesion multiplicity and location are invoked to aid in categorizing these findings.

Many benign bone lesions such as aneurysmal bone cysts, fibrous dysplasia and giant cell tumour are readily identifiable at plain film examination. Typical location, patient age and radiographic features provide sufficient diagnostic information. Following excision, these lesions may be followed radiographically for recurrence, characterized by a change in the cemented appearance of a curreted lesion or prosthesis, or by increased lysis of surrounding bone (Davies and Vanel, 1998). Soft tissue masses are late signs of recurrence and are not reliably identifiable on plain film examination. While MRI displays greater soft tissue contrast for the identification of soft tissue masses, cortical bone signal alterations secondary to treatment may not be distinguishable from those due to recurrence. That is, the increased specificity of tissue delineation inherent in MRI may be lost in the postoperative state, when normal healing may not be distinguishable from recurrent tumour. In this situation, changes in the appearance over time on plain X-rays may be definitive.

The absence of well-defined lesion margins and a wider transition zones, as well as serial film demonstration of rapid growth, suggesting more aggressive lesions, should increase the concern for malignancy.

Malignant processes such as multiple myeloma have radiographic features of generalized osteopenia and expansile, lytic bony lesions, as well as pathological fractures (Blake *et al*., 1999) (**Figure 1**). The skeletal survey has traditionally been considered the initial diagnostic imaging modality for patients with proven or suspected myeloma, since these lesions are usually not detected at scintigraphy. This is because radionuclide bone scans demonstrate increased uptake of the tracer in areas of

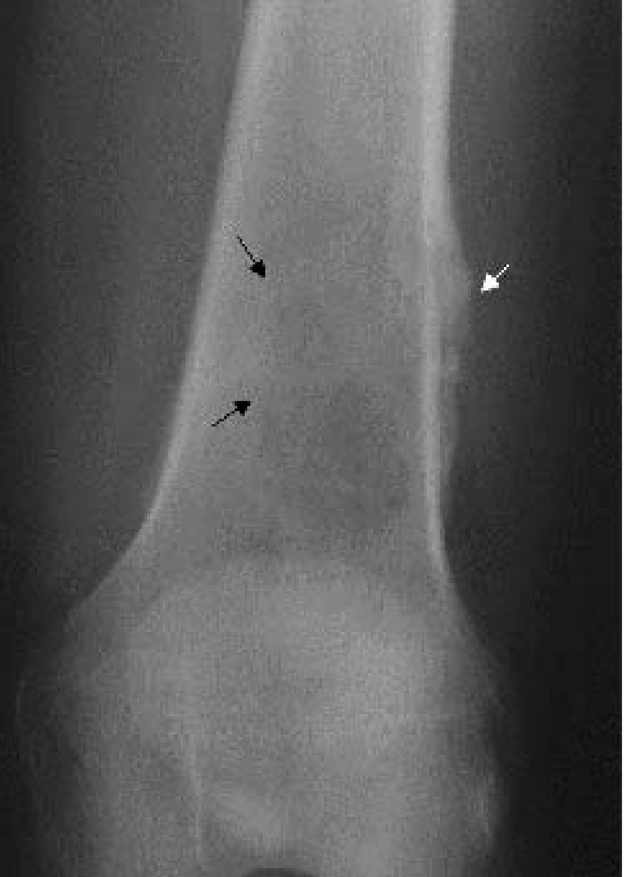


Figure 1 Metastatic bone lesion. 75-year-old female with painful distal femur. Plain film radiograph of distal femur shows permeative, lytic lesion (black arrows) with cortical periosteal reaction (white arrow) seen in distal femur.

bone turnover – that is, destruction and repair. In multiple myeloma, there is typically focal destruction without repair, and therefore the bone scan may be negative in the presence of widespread bony involvement. A significant percentage of bone density must be lost, however, for these lesions to be visible on plain films.

Other primary malignant bone lesions such as chondrosarcoma and osteosarcoma are characterized radiographically by periosteal new bone formation, lifting of the cortex and cortical destruction and chondroid or osteoid matrix. Subtle internal matrices are often better appreciated with CT scanning (Nathan *et al*., 1999). Treatment of malignant sarcomas often incorporates surgical resection and chemo- and radiation therapy. Resection usually involves most of the involved bone and as such recurrence is found in surrounding soft tissues, at the site of resection or at the junction of an implanted prosthesis (Davies and Vanel, 1998). Displaced cement or altered prosthesis position, in comparison with baseline X-rays, may be indicative of underlying recurrence. MRI has been considered the best imaging modality for the evaluation of musculoskeletal neoplasms since soft tissue detail, neurovascular involvement and joint extension can be demonstrated in multiplanar fashion. Again, however, the specificity of MRI may be lost in the postoperative situation.

Soft tissue and neurovascular infiltration, haematoma, local extension and distant metastases and local host response to chemotherapy cannot be reliably identified by plain film radiography (Lindner *et al*., 1998).

Distant metastases and multiplicity of disease may be identified with advanced imaging techniques including radionuclide scintigraphy, CT and MRI. Patterns of bony lysis or sclerosis tend to replicate themselves when imaged by plain film skeletal survey and may be used to identify skeletal metastases, be they due to synchronous or metachronous disease. Whereas radiographs are used for baseline screening for pathological fractures and pulmonary metastases, chest CT is more sensitive for the detection of small, multiple lung nodules. Additional sites of metastatic involvement including the brain and soft tissues of the abdomen and retroperitoneum are not well assessed at plain film radiography and require cross-sectional evaluation (Bearcroft and Davies, 1999). Scintigraphy plays a role in the identification of skeletal metastases but is insensitive to the presence of very small lesions, and to lesions with very rapid bone destruction.

ABDOMINAL NEOPLASMS

Plain film examination of the abdomen has largely been replaced by CT and ultrasound for screening and detection of neoplasms. Plain film screening for cancer, in the patient with mild or nonspecific symptoms, yields little diagnostic information. Use in specific situations, such as bowel perforation or obstruction, has been shown to be very useful (William, 2000).

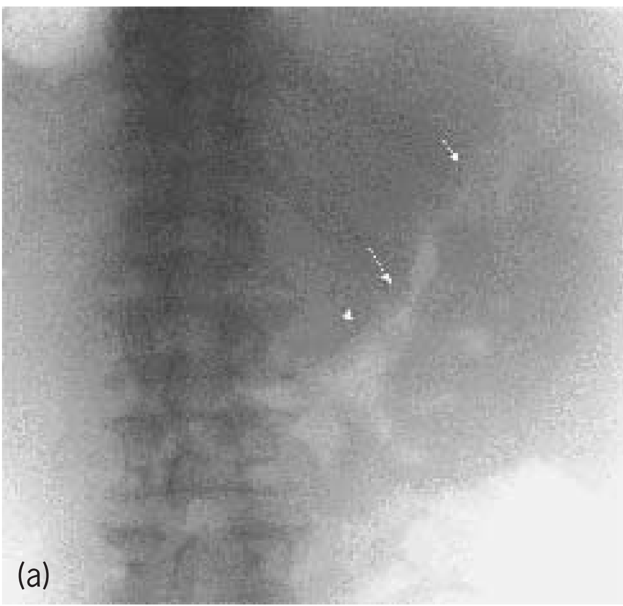
Some basic principles of X-ray attenuation are helpful in clarifying the reasons for this. Intra-abdominal soft tissue contrast is created by the differential attenuation of X-rays by fat, water and air. Retroperitoneal spaces are defined by fat surrounding structures such as the kidneys and retroperitoneal portions of small and large bowel loops. The psoas muscle is a readily identifiable retroperitoneal landmark, since this muscle is bordered by fat. Inferior edges of intraperitoneal structures such as the liver and spleen are also bordered by fat making these visible on X-rays. Indirectly, the organ edges may be infered by their mass effect on adjacent air-filled loops of bowel. Enlargement of these solid organs effaces these fat planes and displaces bowel loops. Intraperitoneal soft tissue masses or ascites (i.e. processes behaving like water density projected against water-density liver or spleen on plain films) also obscures soft tissue planes (Messmer, 2000).

Lymphomas, classified as Hodgkin disease (HD) or extranodal non-Hodgkin lymphoma (NHL), are soft tissue neoplasms involving either lymph nodes, spleen or extranodal sites. Plain film radiography is generally of little use; diagnosis and follow-up require imaging using a modality with greater soft tissue contrast resolution, such as ultrasound or CT scanning (Bragg, 1987) (**Figure 2**).

Hollow viscera such as small and large bowel and stomach contain variable amounts of gas, detectable at radiography. Normal bowel gas acts as contrast material on abdominal plain films. Gas may be seen within ulcerating soft tissue masses or outlining foci of bowel wall thickening associated with neoplastic masses (**Figure 3**). Secondary signs of carcinoma such as bowel obstruction or perforation are usually evident at plain film examination. Liquid barium contrast (administered by mouth or per rectum) is used with insufflated gas to perform double-contrast radiographic examinations of the upper and lower gastrointestinal tract (Bradford *et al*., 2000). These radiographic techniques may be used for duodenal and colonic cancer detection. Staging of neoplastic processes involves the use of additional imaging modalities such as MRI, CT scan and endosonography (Thoeni, 1997).

Dense calcifications are readily apparent on plain films, although intra-abdominal calcifications related to cancer are rare. Typically calcified solid masses include uterine leiomyomas and lymph nodes. Dystrophic and psammomatous calcifications may be seen with fibrolamellar hepatocellular carcinomas, renal cell carcinomas, ovarian cystadenocarcinomas and metastases (Baker and Cho, 2000).

Abdominal plain films, then, are most useful in acute situations, such as when there is a question of bowel perforation or obstruction. They are very useful in combination with barium to define abnormalities in the bowel, but they have little role for routine evaluation or for screening.



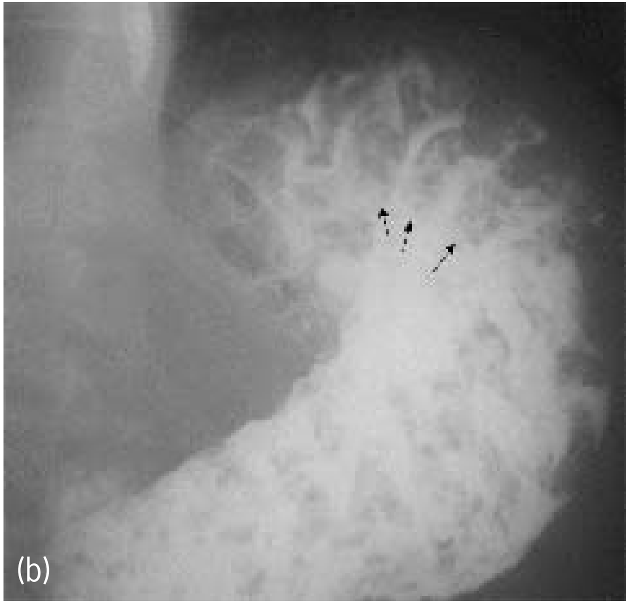


Figure 2 Gastric lymphoma. 36-year-old male with early satiety. (a) Plain film of the abdomen shows upper abdominal soft tissue mass in expected region of stomach with narrowed, lobulated, air-filled luminal contour (white arrows) suggesting gastric wall mass. (b) Double contrast upper gastrointestinal barium examination confirms thickened gastric folds with irregular mucosal contours (black arrows).

LUNG NEOPLASMS

Lung cancer is the leading cause of cancer deaths among men and women in the United States (Miettinen, 2000). Although considerable debate exists in the current literature regarding the role of chest X-ray screening for lung cancer, improved prognosis is clearly associated with detection (by whatever means) of early-stage, potentially curable lesions in high-risk populations (Caro *et al*., 2000; Gavelli and Giampalma, 2000).

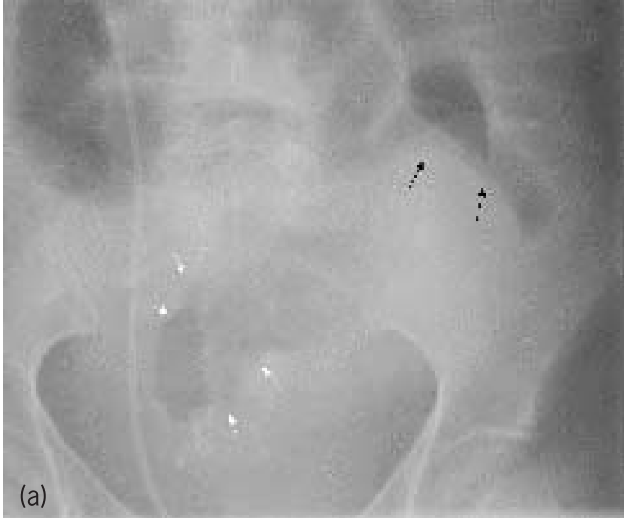
Chest radiography has traditionally served as the first-line imaging modality for lung cancer screening. Plain film radiography of the chest is widely available and of relatively low cost, compared with CT, for the purposes of lung cancer screening and has been recommended as a screening modality for high-risk populations (Dominioni *et al*., 2000) (**Figure 4**). While traditional chest X-rays have been described as having poor diagnostic value for the detection of small lung cancers (Yang *et al*., 2001), various factors continue to influence detection. The most common lung cancer cell type is now adenocarcinoma, which has been described as more slowly growing compared to squamous cell carcinoma, thereby widening the window of beneficial detection (Austin *et al*., 2001). Solid lung tumours, those not maintaining air-filled alveolar spaces, are better seen on chest radiographs (Yang *et al*., 2001). Technological advances in computer-aided conventional and digital chest examinations may improve the diagnostic potential of chest radiography, increasing the yield of nonmetastatic disease (Reeves and Kostis, 2000). The usual chest X-rays are analogue in format. Digital chest radiographs may have somewhat reduced inherent resolution, but the ability to post-process, to re-examine the image with magnification and with alteration of the relative densities, may add substantially to the sensitivity and utility of plain chest X-rays.

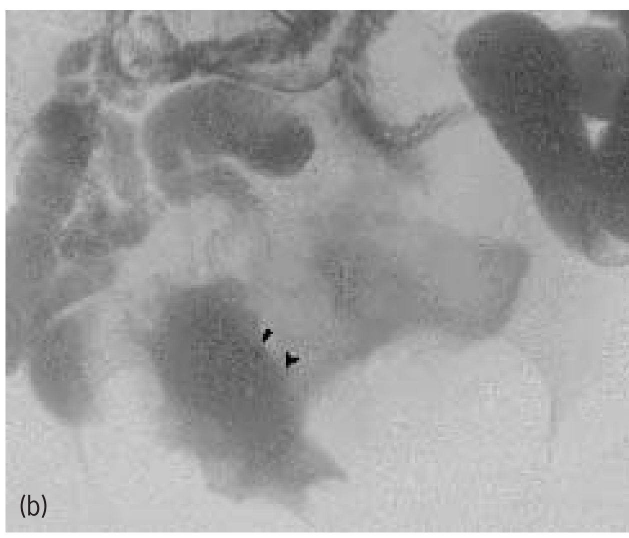
Recent technical advances in the use of low-dose chest CT have proved superior for early-stage lung cancer detection. A recent review of lung cancer detection and screening trials shows that chest CT is more sensitive for detecting small lung nodules that may be cancerous and that CT detects more lung cancers than plain film radiography (Miettinen *et al*., 2001). These findings imply that screening chest CT will detect more cases of resectable early-stage lung cancer, although effects on reducing mortality secondary to the disease are still unknown (Boiselle *et al*., 2000; Patz *et al*., 2000).

Low-dose screening CT will also detect small indeterminate nodules, the majority of which will be benign. The workup of these nodules, however, will not be insignificant in terms of cost and time. The efficiency, then, of low-dose chest CT for lung cancer screening requires further evaluation (MacMahon and Vyborny, 1994).

COMPLICATIONS OF CANCER – DISEASE AND THERAPY

Deleterious effects of cancer chemotherapy and steroid therapy include osteoporosis and osteonecrosis. Proposed mechanisms for these side effects include decreased





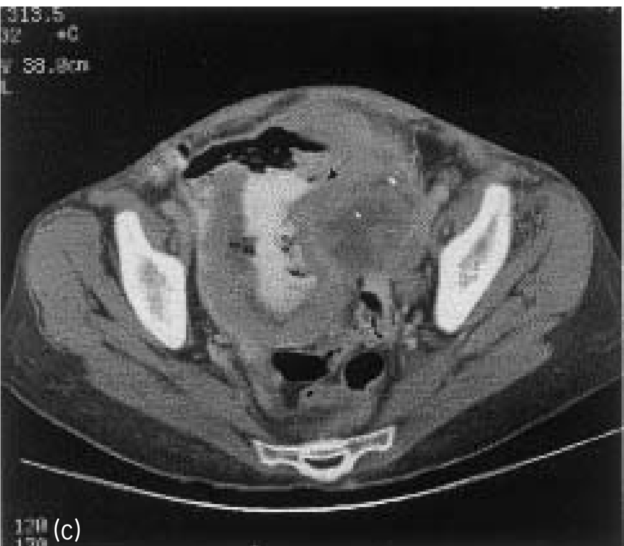
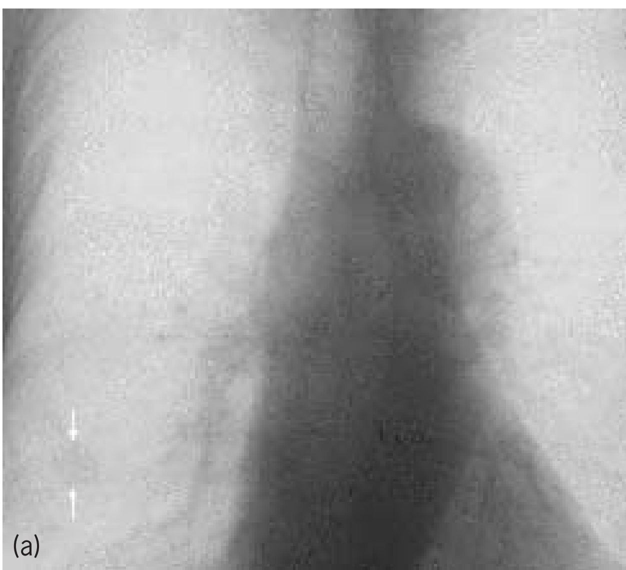


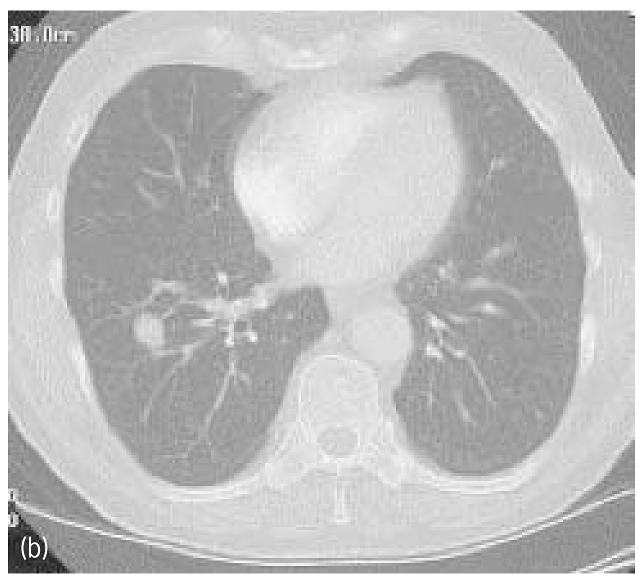
Figure 3 Metastatic melanoma. 53-year-old female with known melanoma and abdominal pain. (a) Abdominal radiograph shows distended, air-filled loops of bowel in abdomen that are displaced upward by pelvic soft tissue mass (black arrows) containing unusual air pattern (white arrows). (b) Double-contrast barium examination of the small bowel shows unusual collection of barium (arrowheads) conforming to configuration of unusual air pattern seen on plain film (a). (c) CT scan confirms luminal contrast filling ulcerated pelvic soft tissue mass (black arrow).

bone formation characterized by deficient osteoblasts and osteoclasts, and also focal bone segment death characterized by cell death in cortical bone (Weinstein *et al*., 1998). The role of radiographs for surveillance of these deleterious side effects is controversial as radiographic diagnosis of osteoporosis is subject to wide interpreter variability. Furthermore, the positive predictive value of plain film detection of bony erosions for the purpose of predicting clinically significant joint disease has recently been reported as 40% (Falbo *et al*., 1999).

Osteoporosis may result from cancer treatment drugs, such as methotrexate, as well as the tumour itself. Osteoporosis is a process affecting the microarchitecture of bony tissue, causing loss of bone mass which can lead to fractures. These fractures, and diminished bone density, may be seen on plain films, although subtle or incomplete fractures and processes involving bone marrow are more sensitively imaged with MRI. The role of plain film radiographic evaluation may lie in identifying those patients at risk of the effects of osteoporosis and who may benefit from the institution of prophylaxis or treatment. An example is the use of orthopaedic procedures to prevent pathological fracture due to a metastatic deposit, or the institution of pharmacotherapy to treat severe systemic osteoporosis.

Methotrexate is one chemotherapeutic agent that has been recognized as associated with osteoporosis (Pfeilschifter and Fiel, 2000). Methotrexate osteopathy is probably secondary to the combined processes of increased bone resorption and inhibited bone formation. Other agents associated with osteoporosis include cyclophosphamide, doxorubicin and corticosteroids. While bone densitometry is considered the study of choice for the evaluation of osteopenia, plain film evaluation is performed as a first-line imaging modality. While significant loss of bone mass is required before radiographic evidence of loss of density is evident, plain films, particularly if prior films are available for comparison, are useful for screening purposes, and





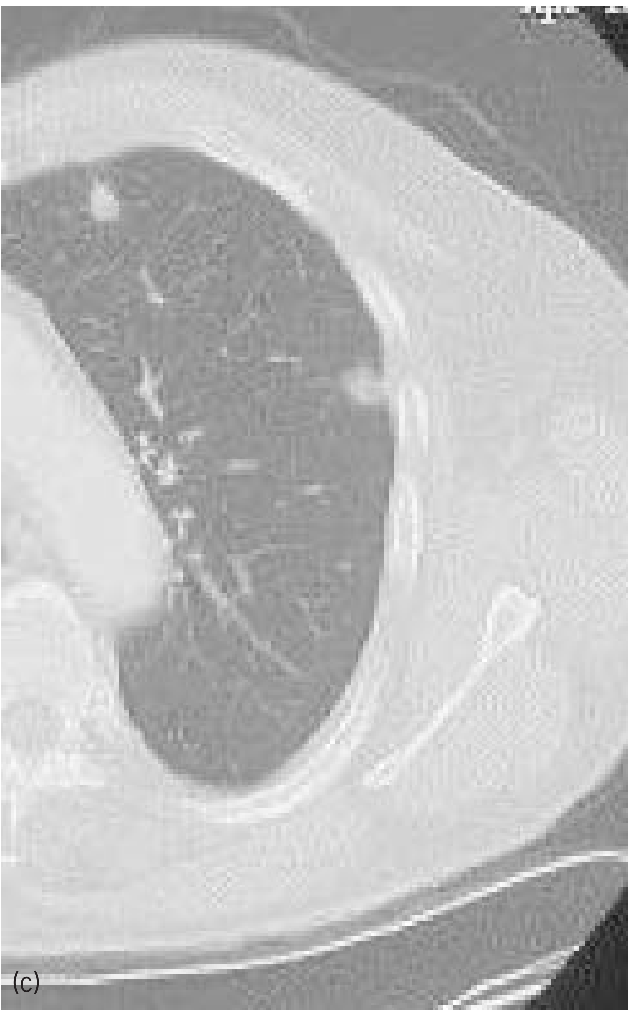


Figure 4 Metastatic rectal cancer. 75-year-old male with resected rectal cancer. (a) Follow-up chest X-ray shows new right lung nodule (white arrows). (b) CT scan shows right lung nodule seen at plain film examination of chest. (c) CT scan of left chest shows additional nodules.

also demonstrate specific abnormalities associated with osteoporosis, such as vertebral body wedge compression fractures and late signs of femoral osteonecrosis.

Hypertrophic pulmonary osteoarthropathy (HPO) may be secondary to an underlying process such as lung cancer. Typical findings of HPO on plain film examination include periostitis involving diaphyseal regions of long bones, becoming thicker and more extensive with time, as well as soft tissue ossification (Viola *et al*., 2000). Lung cancer may be detected secondarily following the identification of HPO on plain films. Radiographic follow-up for resolution of bone changes following resection of lung cancer is controversial. Failure of bone abnormalities to resolve postresection was previously thought to favour poor prognosis, although a recent report showed that in two patients, the absence of reversal of periosteal changes did not correlate with poor long-term outcome (Orts *et al*., 2000).

CONCLUSION

Plain film radiography is an imaging modality that is widely available, low in cost, has essentially no associated morbidity unless used excessively and frequently serves as the first-line imaging modality in the workup of many disease processes.

Bone radiographs identify features of bony disease that may be unique to the diagnosis in question, and are useful for evaluation of musculoskeletal neoplasms both initially

as the sole modality, for follow-up of lesions and in conjunction with additional modalities such as CT and radionuclide scanning and MRI.

The use of chest X-rays for the detection of lung cancer is controversial as the effect of dedicated screening programmes on long-term survival is still unknown. Extensive studies to evaluate this question, that is, whether early detection actually alters survival overall, or whether it merely discloses very slow-growing lesions that would have little impact on long-term outcome, are currently under way. Recent studies have suggested that the use of CT scanning for the identification of lung cancer, in particular small and multifocal lung lesions, is beneficial. Whether or not screening of high-risk patients with CT will reduce mortality secondary to lung cancer, compared with screening with chest radiographs or with no routine radiographic screening, has not yet been conclusively shown.

Finally, while plain film examination of the abdomen yields information regarding bowel obstruction and perforation, and also bowel wall thickening and displacement of bowel loops, this imaging modality lacks the contrast resolution necessary to distinguish soft tissue detail in the abdomen and pelvis. Abdominal radiographs, then, are mainly of use in patients with particular symptoms or signs. They play no role in the evaluation or follow-up of patients with known or suspected malignancies.

Plain X-rays, then, play a major role for evaluation of malignancies in certain organ systems. They have the advantages of ready availability in most facilities around the world, relatively low cost, little if any associated morbidity if obtained properly and high utility in specific situations. They have significant limitations, but have the added advantage that these limitations are well delineated and widely recognized. Plain radiography, as initially developed by Roentgen in 1895, remains a central tool in the diagnosis and treatment of malignancies.

REFERENCES

Austin, J. H. M., *et al*. (2001). Screening for lung cancer. *New England Journal of Medicine*, **344**, 935.

Baker, S. R. and Cho, K. C. (2000). Abdominal calcifications. In: Gore, R. M. and Levine, M. S. (eds), *Textbook of Gastrointestinal Radiology*, 2nd edn. 178–187 (W. B. Saunders, Philadelphia).

Bearcroft, P. W. and Davies, A. M. (1999). Follow-up of musculoskeletal tumors. 2. Metastatic disease. *European Radiology*, **9**, 192–200.

Blake, M. A., *et al*. (1999). Musculoskeletal case of the day. Multiple myeloma. *American Journal of Roentgenology*, **173**, 796–800.

Boiselle, P. M., *et al*. (2000). Lung cancer detection in the 21st century: potential contributions and challenges of emerging technologies. *American Journal of Roentgenology*, **175**, 1215–1221.

Bradford, D., *et al*. (2000). Early duodenal cancer: detection on double-contrast upper gastrointestinal radiography. *American Journal of Roentgenology*, **174**, 1564–1566.

Bragg, D. G. (1987). Radiology of the lymphomas. *Current Problems in Diagnostic Radiology*, **16**, 177–206.

Caro, J. J., *et al*. (2000). Could chest X-ray screening for lung cancer be cost-effective? *Cancer*, **89** (Suppl.), 2502–2505.

Davies, A. M. and Vanel, D. (1998). Follow-up of musculoskeletal tumors. 1. Local recurrence. *European Radiology*, **8**, 791–799.

Dominioni, L., *et al*. (2000). Screening for lung cancer. *Chest Surgery Clinics of North America*, **10**, 729–736.

Falbo, S. E., *et al*. (1999). Clinical significance of erosive azotemic osteodystrophy: a prospective masked study. *Skeletal Radiology*, **28**, 86–89.

Feldman, F. (2000). Musculoskeletal radiology: then and now. *Radiology*, **216**, 309–316.

Gavelli, G. and Giampalma, E. (2000). Sensitivity and specificity of chest X-ray screening for lung cancer: review article. *Cancer*, **89** (Suppl.), 2453–2456.

Jaovisidha, S., *et al*. (1998). An integrated approach to the evaluation of osseous tumors. *Orthopedic Clinics of North America*, **29**, 19–39.

Lindner, N. J., *et al*. (1998). Local host response in osteosarcoma after chemotherapy referred to radiographs, CT, tumour necrosis and patient survival. *Journal of Cancer Research and Clinical Oncology*, **124**, 575–580.

MacMahon, H. and Vyborny, C. (1994). Technical advances in chest radiography. *American Journal of Roentgenology*, **163**, 1049–1059.

Messmer, J. M. (2000). Gas and soft tissue abnormalities. In: Gore, R. M. and Levine, M. S. (eds), *Textbook of Gastrointestinal Radiology*, 2nd edn. 157–177 (W.B. Saunders, Philadelphia).

Miettinen, O. S. (2000). Screening for lung cancer. *Radiology Clinics of North America*, **38**, 479–486.

Miettinen, O. S., *et al*. (2001). Screening for lung cancer. *New England Journal of Medicine*, **344**, 935.

Nathan, R. C., *et al*. (1999). Musculoskeletal case of the day. Clear cell chondrosarcoma. *American Journal of Roentgenology*, **173**, 795–799.

Orts, D., *et al*. (2000). Hypertrophic osteoarthropathy in lung cancer: are the radiographic bone changes reversible after curative resection? *Monaldi Archives of Chest Diseases*, **55**, 122–123.

Patz, E. F., *et al*. (2000). Current concepts: screening for lung cancer. *New England Journal of Medicine*, **343**, 1627–1633.

Peabody, T. D., *et al*. (1998). Evaluation and staging of musculoskeletal neoplasms. *Journal of Bone and Joint Surgery*, **80**, 1204–1218.

Pfeilschifter, J. and Fiel, I. J. (2000). Osteoporosis due to cancer treatment: pathogenesis and management. *Journal of Clinical Oncology*, **18**, 1570–1593.

Reeves, A. P. and Kostis, W. J. (2000). Computer-aided diagnosis for lung cancer. *Radiology Clinics of North America*, **38**, 497–509.

Thoeni, R. F. (1997). Colorectal cancer. Radiologic staging. *Radiology Clinics of North America*, **35**, 457–485.

Viola, I. C., *et al*. (2000). Primary hypertrophic osteoarthropathy. *Journal of Rheumatology*, **27**, 1562–1563.

Weinstein, R. S., *et al*. (1998). Inhibition of osteoblastogenesis and promotion of apoptosis of osteoblasts and osteocytes by glucocorticoids. Potential mechanisms of their deleterious effects on bone. *Journal of Clinical Investigation*, **102**, 274–282.

Williams, S. (2000). Abdomen: normal anatomy and examination techniques. In: Gore, R. M. and Levine, M. S. (eds), *Textbook of Gastrointestinal Radiology*, 2nd edn. 144–156 (W. B. Saunders, Philadelphia).

Yang, Z. G., *et al*. (2001). Visibility of small peripheral lung cancers on chest radiographs: influence of densitometric parameters, CT values and tumour type. *British Journal of Radiology*, **74**, 32–41.

Chapter 76

Computed Tomography

Marlene L. Zawin, Catherine M. Camputaro and Arthur T. Rosenfield
Yale University Medical Center, New Haven, CT, USA

CONTENTS

- Introduction
- Basic CT Features and Techniques
- Clinical Applications
- Future Directions

INTRODUCTION

Computed tomography (CT) has widespread oncological applications. CT produces transverse sections that provide a three-dimensional display of the distribution of X-ray attenuation within the body. Its basic principle is that the internal structure of an object can be reconstructed from multiple projections of the object. Highly collimated X-rays are focused on pre-selected transverse sections of a patient. The data acquired from the multiple detectors subsequently undergo computer analysis, reconstructing the precise attenuation information into an image for display on a monitor. CT can detect minute differences in tissue X-ray attenuation and provides highly accurate quantitative information about the X-ray attenuation properties of the imaged tissues. All CT scanners have an X-ray generator, scanning gantry, computer system, operator and physician's viewing consoles and a hard copy camera (**Figure 1**) (Ter-Pogossian, 1977). The scanning gantry supports the high-heat-capacity X-ray tube, collimators, detector array with its associated electronics and scanning drive motor. Recent advances in both scanner hardware, with faster image acquisition, and in software applications, with multiple viewing options and interactive capabilities, have produced new and exciting CT applications.

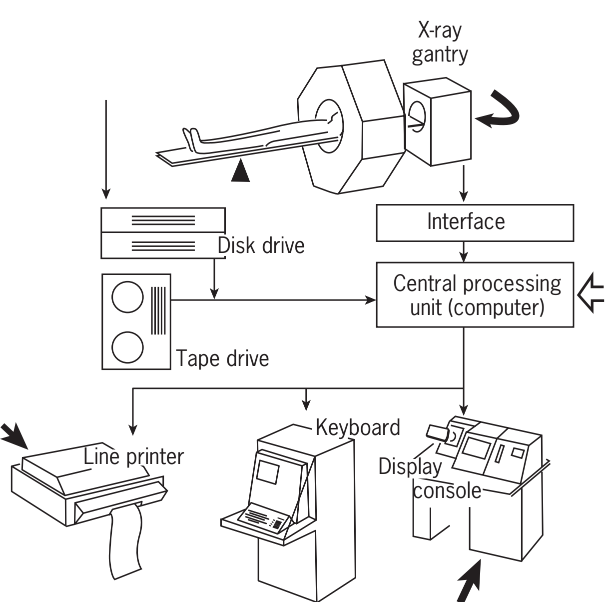


Figure 1 Basic components of a CT scanner. The scanning equipment consists of the CT table (arrowhead) and the X-ray gantry which consists of the X-ray tube, generator, collimators and detectors (curved arrow). These are interfaced with the main computer (central processing unit (CPU)) (open arrow), which in turn interfaces with the viewing elements which may include: technologists' and radiologists' display consoles (broad arrow), tape and disk drives (long arrow) and printer (short arrow). (From Ter-Pogossian, 1977, *Seminars in Roentgenology*, **12**, 13–25.)

BASIC CT FEATURES AND TECHNIQUES

Data Acquisition

In conventional radiography, a large proportion of available information is lost in attempting to display it from a three-dimensional patient on to a two-dimensional film as all of the structures are superimposed on top of each other. Individual structures can be distinguished only if they differ significantly in density from one another, thus enabling them to stand out clearly against the tissues in front and behind them. With standard CT, single sections of a patient are imaged sequentially. This is accomplished by obtaining a series of angular projections or views of the section and reconstructing a two-dimensional image from this series of one-dimensional projections. The X-ray beam

is confined to a selected section of the patient (the width of which is determined by the prepatient collimators close to the X-ray tube). This improves image detail by eliminating the superimposition of unwanted structures. Post-patient collimators located close to the X-ray detectors limit scatter radiation which is also a cause of image degradation in conventional radiography. The detectors measure the intensity of the X-ray beam exiting the patient and their signals are used by the central processing unit (CPU) to calculate 'profiles' of the X-ray attenuation by the tissues traversed that section (Ter-Pogossian, 1977).

Image Reconstruction

The information obtained by the detectors is first stored in quantitative form. The scanned transverse section of the patient is subdivided into a series of small individual blocks called 'voxels.' Each voxel is assigned a value proportional to the degree that it attenuated the X-ray beam. The linear attenuation coefficient (μ) is used to quantitate the attenuation. It is determined by the following equation:

$$N = N_0 e^{\mu - x} \quad [1]$$

where N_0 is the number of initial photons, N the number of transmitted photons and x the slice thickness. Thousands of these equations, derived from linear transmission readings taken from different directions (projections), must be solved to determine the linear attenuation coefficients of all the voxels.

The linear attenuation coefficients then undergo another calculation, which permits the computer to present the information as a picture with a large grey scale. The reconstructed CT image is a two-dimensional array of quantized grey-scale values or picture elements (pixels). The pixel values or CT numbers are expressed as the ratio of attenuation coefficients between a given material and water, and are directly related to the linear attenuation coefficients at corresponding positions within the slice:

$$\begin{aligned}&\text{pixel value (Hounsefield units)}\\&= 1000\,(\mu - \mu^{\text{water}})/\mu^{\text{water}}\end{aligned} \quad [2]$$

where (μ is the average linear attenuation coefficient of the voxel represented by the pixel and μ^{water} is the linear attenuation coefficient of water (both are evaluated for the effective energy of the beam exiting the patient) (Barnes *et al*., 1979). Thus, water has a CT number of zero, and a region with a CT number of 100 Hounsefield units (HU) has a linear attenuation coefficient that is 10% greater than that of water. CT images usually have 12 bits per pixel, and they are employed to represent numbers ranging from −1000 to 3095 HU, or 4096 grey scale values. Air has a CT number of −1000 HU, fat ranges from −50 to −350 HU, soft tissue structures from 10 to 80 HU and bone from 400 to 3000 HU.

In order to display this wide range of CT numbers on a monitor with generally 256 shades of grey, a CT number ('window level') is selected which is the estimated average of the tissue of interest. The computer is then instructed to assign one shade of grey to each of the 128 CT numbers below and each of the 128 CT numbers above the window level ('window width'). Both the window level and width can be adjusted in accordance to the structure being evaluated. Multiple window levels and window widths are employed to maximize the information that can be obtained from a single CT examination (**Figure 2**).

CT Scanner Evolution

Godfrey N. Hounsefield of EMI Ltd in England introduced the first clinical CT scanner in 1971 (Hounsefield, 1973). Thereafter, there have been four generations or models of the standard two-dimensional CT scanner. Their scanner designs differed in the movement of the X-ray source and both the number and the movement (or lack of, as in the case of the fourth-generation CT) of the X-ray detectors. Developments in both CT hardware and software have been, to a great extent, toward faster scanning. Patient motion, either voluntary (inability to cooperate or understand instructions) or involuntary (due to respiration, cardiovascular pulsations or peristalsis) can result in data misregistration which may cause omission of lesions.

Spiral, helical or continuous-volume CT scanners, with their attendant hardware and software developments, were introduced clinically in 1989 (Kalendar *et al*., 1990; Brink *et al*., 1994). The source–detector assembly is supported by multiple parallel slip-rings rather than electrical cables which permit its continuous rotation. This has decreased interscan delays to less than 5 s and has made volume-acquisition CT scanning possible. Scanning involves simultaneous patient movement at a constant rate through the gantry during continuous rotation of the source–detector assembly such that continuous data acquisition is achieved throughout a volume of interest. The terms spiral or helical CT are derived from the fact that during the scanning process, the X-ray focus describes a spiral or helical path around the patient (**Figure 3**) (Kalendar *et al*., 1990). Scan times for an entire study range from less than 10 s to as long as 30 s, with a data set often obtained within a single breath-hold. Direct reconstruction of images from data obtained over any 360° segment of a spiral CT acquisition will result in motion artifacts due to patient transport. The spiral data set first undergoes a mathematical smoothing process (interpolation) which can be calculated from any arbitrary table position within the scanned volume, followed by conventional image reconstruction. A data volume is generated that may be viewed as transaxial planar images or with multiplanar and 3D methods.

Volumes of interest are scanned continuously, providing seamless or overlapping sets of images. This is of

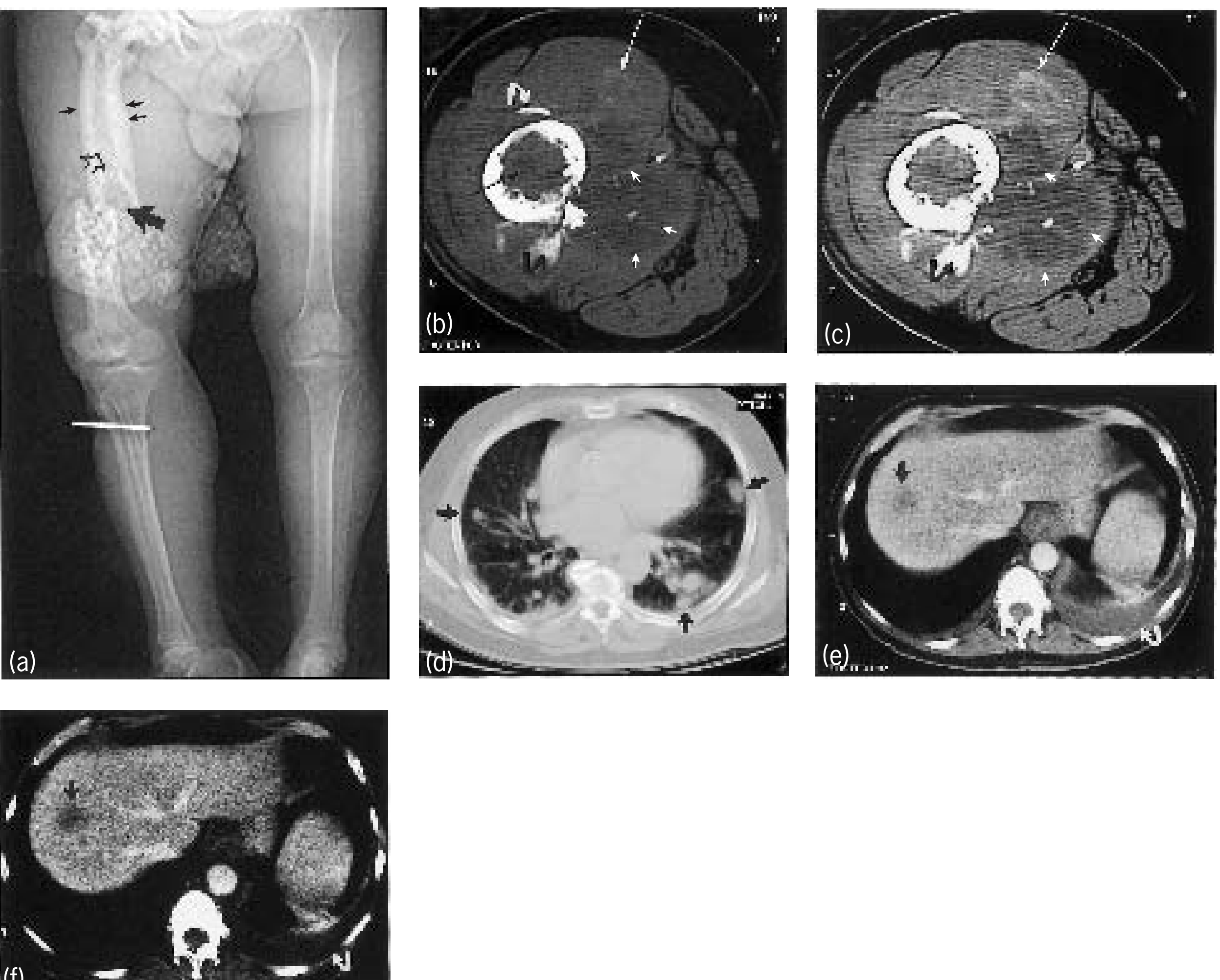


Figure 2 Window settings on CT. Malignant degeneration of a benign lesion. (a) A CT scout image of the legs reveals a pathological fracture of the midiaphysis of the femur. The bone is abnormally enlarged with marked cortical thickening (small arrows) and thickened trabeculae (open arrow) compatible with Paget's disease of bone. A luscency is present at the level of the fracture (large arrow). The scout image is analogous to a conventional radiograph, which is routinely acquired when a patient is first placed on the CT table. It serves to ensure appropriate patient positioning to include the entire area under investigation. It is the reference from which to assign the most superior and inferior locations of the upcoming scan, to select the appropriate field of view (FOV), slice thickness and interslice skip in order to ensure adequate area coverage and sufficient radiographic detail. (b) An axial CT image through the level of the fracture in bone windows confirms the presence of a central lytic lesion (long black arrow) with an associated fracture (large white arrow). Bony fragments (curved arrow) and ill-defined luscent areas (small arrows) representing oedema are present in the adjacent soft tissues. The high attenuation area in this noncontrast CT is compatible with acute blood (long white arrow). (c) Same image as (b) with soft tissue window setting. The fine detail of the bony fragment (curved black arrow) is obscured. The soft tissue changes however, are better appreciated. The luscent areas (small arrows) are more obvious and better defined as is the haemorrhagic area (long arrow). A CT-guided percutaneous biopsy revealed an osteosarcoma. (d) Thoracic CT axial image with lung window setting reveals multiple, bilateral lung nodules compatible with metastatic disease (arrows). Lung window settings optimize visualization of the lung parenchyma at the expense of the skeletal (b) and soft tissue (c) areas. (e) Axial image caudal to (d) at the level of the lower chest and upper abdomen with soft tissue windows reveals a low attenuation hepatic lesion (straight arrow) compatible with a metastasis and a left pleural effusion (curved arrow). (f) Same image as (e) utilizing liver window settings which enhance the conspicuity of the hepatic metastasis (straight arrow). The left pleural effusion is barely visible (curved arrow).

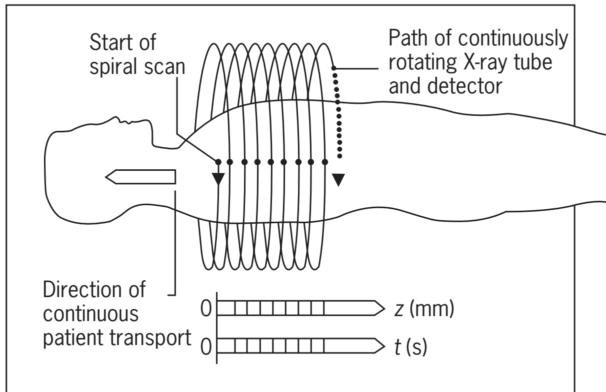


Figure 3 Helical CT scanning. Continuous patient transport and rotating X-ray tube and detectors result in volumetric data acquisition. (From Kalendar *et al.*, 1990, *Radiology*, **176**, 181–183.)

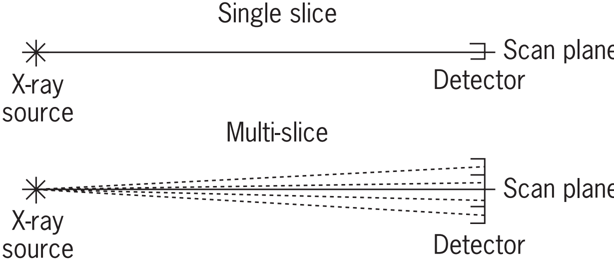


Figure 4 Single-slice vs multi-slice imaging. In single-slice imaging, there is a single X-ray source and single detector. In multi-slice imaging, there are multiple detectors. Thus with multi-slice imaging on a single revolution, multiple slices are generated at once permitting significantly greater speed or resolution. (From Fox *et al.*, 1998, *Neuroimaging Clinics of North America*, **8**, 497–513.)

major importance in scanning anatomical regions, which are subject to patient motion, in which data misregistration from section to section may cause omission of structures or lesions and degradation in the quality of multiplanar and 3D displays. It is also clinically useful in contrast media studies in which complete organs or a large volume need to be measured during different phases of enhancement.

Multidetector-row helical CT is the latest innovation in scanning. Instead of a single row of detectors, multiple rows are used. These rotate simultaneously so that multiple axial images can be obtained at once. In helical mode, four (or more) interweaving helices are created. The ability to acquire multiple images simultaneously can be used for increased speed to limit the effects of cardiac pulsation and peristalsis and to permit a greater area to be covered on one breathold limiting or eliminating misregistration due to breathing. Alternatively, by focusing on a limited region, much higher resolution is possible, particularly for 3D and virtual imaging (**Figure 4**) (Hu *et al.*, 2000).

CLINICAL APPLICATIONS

Lesion Detection

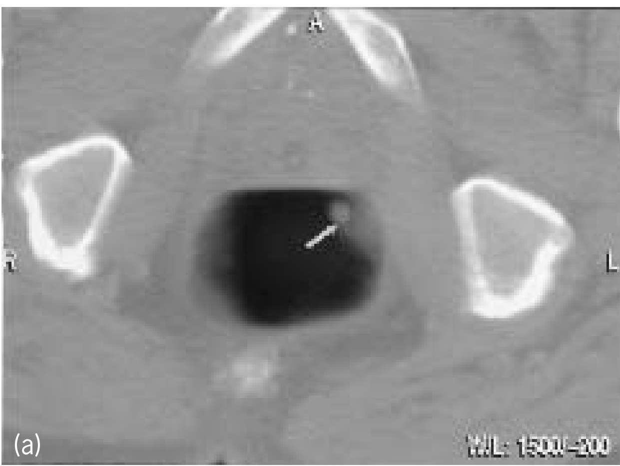
Asymptomatic Patients: Screening

Advances in CT scanner technology have enabled helical CT to be used as a cancer screening modality. Since entire body regions can now be scanned seamlessly and rapidly, often within a single breath-hold, CT examinations are more feasible for a broad range of patients and diseases. An increase in patient throughput due to the increased speed of new scanners permits the performance of a large volume of studies. These factors are crucial to any successful screening programme.

Lung and colon cancers are currently the first and second most common causes of cancer-related deaths in the United States (Landis *et al.*, 1998). The goal of screening in either case is detection and intervention at an early, potentially curable stage. Radiation dose is always a consideration, particularly in long-term screening programmes which require serial examinations every 1–2 years. Fortunately, substantial dose reduction is possible in both cases owing to the high inherent contrast between air-containing organs (the lung and the air-distended colon) and the soft tissue nodules or polyps. Although studies have investigated the accuracy of low-dose CT in the detection of pulmonary nodules (Itoh *et al.*, 2000) and to lesser extent colonic polyps, the controversy concerning technique continues.

Colon cancer is the second leading cause of cancer-related deaths per year in industrialized nations. Most large bowel malignancies arise from pre-existing adenomas; this is substantiated by a decreased rate of cancer occurrence after colonoscopic polypectomy (Winawer *et al.*, 1993). Polyps 1.0 cm in diameter or larger are associated with a higher risk of future malignant change. In individuals older than 50 years, the prevalence of polyps and cancers increases. Therefore, screening is recommended for average-risk individuals in this age group. Screening asymptomatic persons for colorectal cancer is increasingly recognized to result in a 25–50% reduction in cancer mortality. The aim of screening is to detect and remove precancerous polyps and so prevent the development of invasive carcinoma (Winawer *et al.*, 1993). Endoscopy is currently the investigation of choice for the screening, diagnosis, and treatment of colorectal polyps. It is expensive, however, and can have serious complications.

Helical endoluminal CT colonography has shown promise in screening for intraluminal colonic abnormalities such as tumours and polyps. The preparation is identical with that of conventional colonoscopy. The patient is



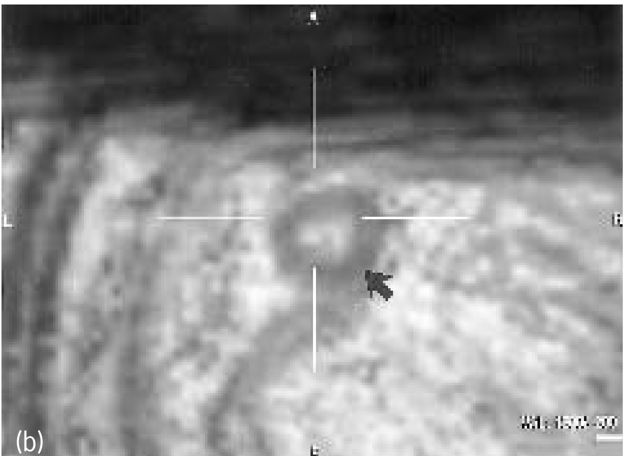


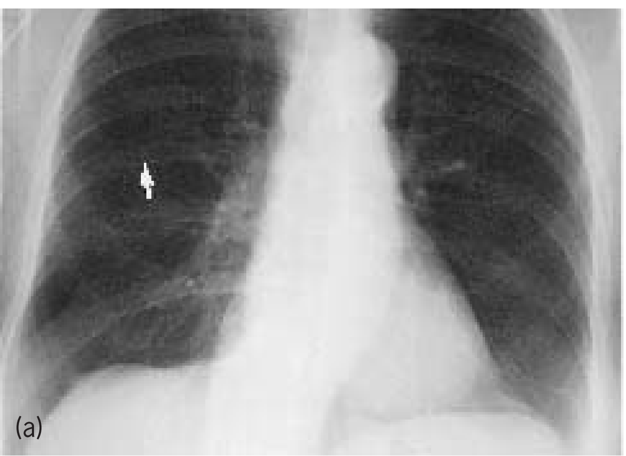
Figure 5 CT or virtual colography (colonography or colonoscopy). (a) Axial CT image at the level of the distended distal sigmoid colon demonstrates a polyp (arrow). (b) Single image from virtual colonoscopic study demonstrates the polyp (black arrow, epicentre of crosshairs).

placed on the CT table and the colon is gently distended by puffs of air to tolerance. Utilizing the subsequent scout image, the entire colon is localized and two breath-hold scans, one supine and the other prone, are obtained. The total examination time is less than 15 min. The acquired data are then viewed as a combination of magnified multiplanar reformatted images and three-dimensional volume-rendered or shaded surface endoluminal images, permitting ante- and retrograde navigation through the colonic lumen (Morrin *et al*., 1999). Preliminary results indicate that the sensitivity for detection of polyps larger than 1 cm is 100%, with a sensitivity of 71–90% for 5–9-mm polyps (**Figure 5**) (Morrin *et al*., 1999).

In addition, CT colonography has demonstrated utility in visualizing the more proximal portions of the colon in cases where conventional colonoscopies which have been incomplete – 5–15% of cases either for technical reasons (Macari *et al*., 1999; Morrin *et al*., 1999) or secondary to an obstructing lesion (Fenlon *et al*., 1999). This is significant because 40% of polyps occur proximal to the splenic flexure. Furthermore, synchronous adenomas are more common in the proximal colon of patients with left-sided polyps. Thirty-one percent of a group of asymptomatic individuals with benign distal adenomas at screening flexible sigmoidoscopy were found to have proximal synchronous neoplasms (Read *et al*., 1997). The CT colography is performed at the same sitting, with the instillation of only a small amount of additional air.

Lung cancer is the leading cause of cancer-related deaths in the United States. The cure rate for lung cancer is 12% with only a slightly higher 5-year survival rate. When stage I cancer is resected, the 5-year survival can be as high as 70% (Henschke *et al*., 1999). Earlier screening programmes utilizing chest X-rays failed to demonstrate any improvement in mortality. In Japan, since the introduction of a national screening programme, survival rates have increased from 33.7% prior to screening to 58.4% (Koike *et al*., 1999). These results were attributed to the relative increased rate of detection of stage I cases – particularly IA peripheral, and also roentgenographically occult lesions. One thousand patients have been enrolled in the Early Lung Cancer Action Project (ELCAP) since its formation in 1992. They are symptom-free individuals, aged 60 years or older, with no history of cancer, who have at least a 10 pack-year history of cigarette smoking and who are medically fit to undergo thoracic surgery (Henschke *et al*., 1999). The protocol consists of a chest X-ray and a low-dose helical CT scan. If a non-calcified nodule is detected, a standard-dose chest CT with high-resolution scanning of the nodule is performed (**Figure 6**). A nodule is classified as benign if its margins are smooth, it measures less than 20 mm and it contains benign calcifications, which were undetected on the low-dose scan. If these criteria are not met, further recommendations are made based upon nodule size (follow-up high-resolution scans if the nodule is 5 mm or less in size, or biopsy if it is 6 mm or larger). The initial results indicate that for low-dose CT, non-calcified nodules were detected three times as commonly as on chest radiography, malignant tumours four times as commonly and stage I tumours six times as commonly. The high frequency of detection of stage I tumours is a strong indication that the cure rate of CT-detected disease is much higher than the current overall cure rate of 12% in the United States (Henschke *et al*., 1999).

Helical CT and software advances now permit three-dimensional volumetric measurements of small pulmonary nodules. These measurements appear to be more accurate than standard two-dimensional measurements in assessing nodule growth and in calculating their doubling times which has been a reasonable, noninvasive technique for prediction of malignancy (Yankelevitz *et al*., 2000).



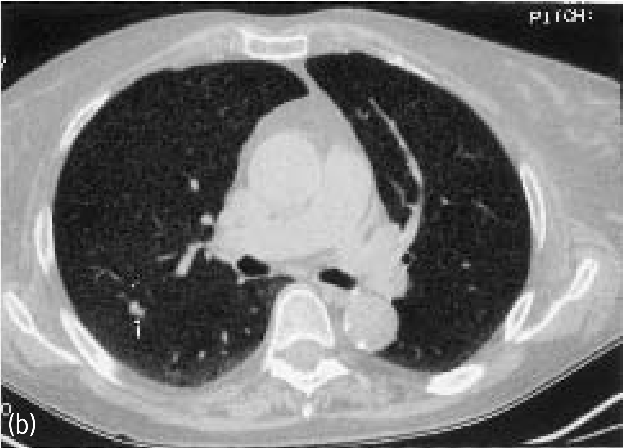
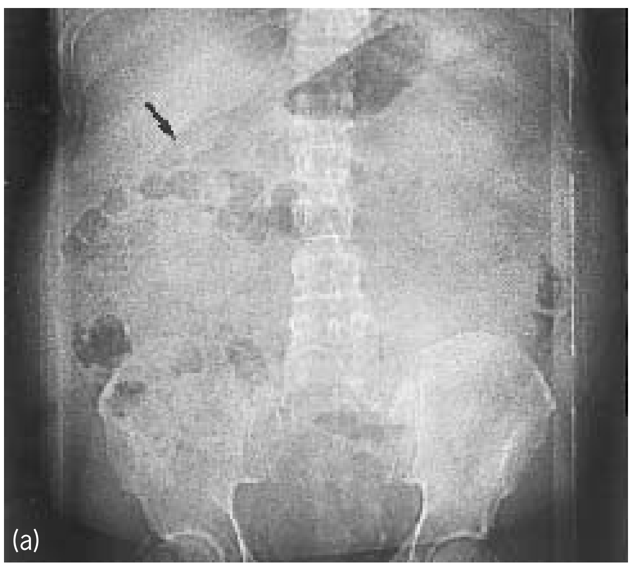


Figure 6 Solitary pulmonary nodule in an asymptomatic smoker. (a) Frontal chest radiograph reveals a tiny, barely perceptible parenchymal nodule (large arrow). (b) A high-resolution axial CT scan confirms the presence of a non-calcified, somewhat lobular parenchymal nodule (small arrow) which was resected and revealed a stage 1 bronchoalveolar cell carcinoma.



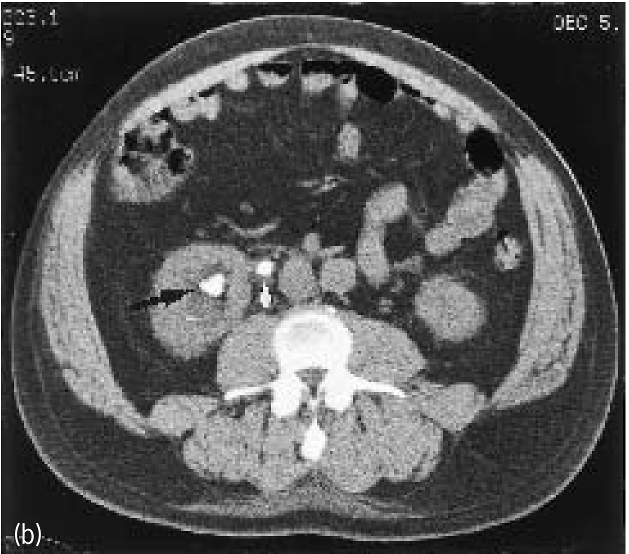


Figure 7 Nephrolithiasis on noncontrast CT, the flank pain protocol. (a) CT scout image of a patient with right flank pain fails to demonstrate any radiopacities in the region of the urinary tract (arrow). (b) Axial slice from a noncontrast CT demonstrates two calculi, one in the renal pelvis (black arrow) and the other in the proximal ureter (white arrow). Noncontrast CT is useful in patients with paraneoplastic syndrome and in cases of suspicious nephrolithiasis to identify ureterolithiasis as the cause of abdominal pain. All stones are opaque on CT except for the rare Indinivir stone in human immunodeficiency virus (HIV) patients receiving this drug.

Symptomatic Patients: Initial Radiographic Workup

CT is widely employed in the diagnostic workup of clinical abnormalities and may, in fact be the examination initially indicated. Abnormalities may be discovered by history such as acute flank pain (**Figure 7**), physical examination as in the case of acute neurological changes suspicious of an intracranial bleed (**Figure 8**) or on laboratory testing. Haematuria detected on routine urine analysis in individuals (particularly males) older than 55 years necessitates evaluation of both the upper and lower urinary tracts. Conventional intravenous pyelography (IVP) has been the standard imaging examination. CT/IVP is a technique consisting of a conventional IVP generally without linear tomography, followed within 2 h by CT scanning of the kidneys and any area of concern on the IVP (Warshauer *et al.*, 1988). This technique provides clarification of IVP findings such as renal

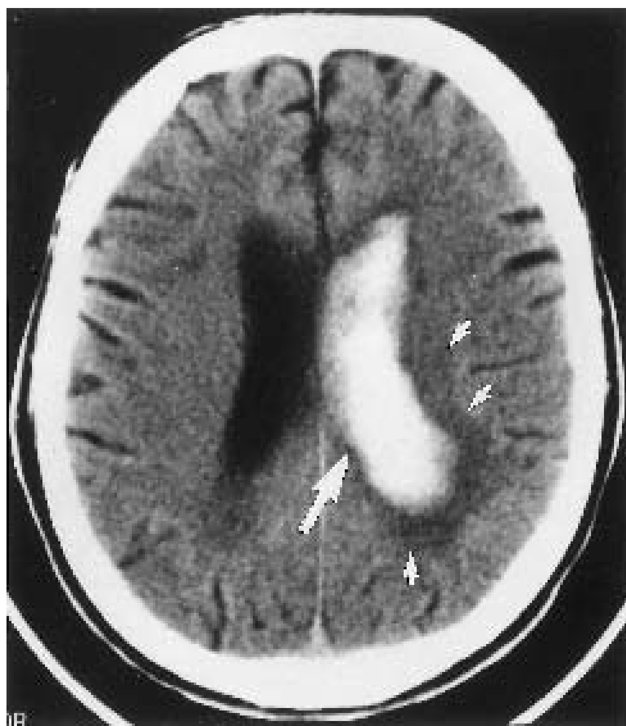


Figure 8 Noncontrast cranial CT. Axial image reveals high attenuation (bright) acute blood within the enlarged left lateral ventricle (large arrow). The adjacent low attenuation (dark) areas are secondary to the accompanying brain oedema (small arrows). The findings were secondary to an acute haemorrhage of a previously undiagnosed brain metastasis in a patient with known primary lung cancer.

malposition, abnormalities in the renal contour, or collecting system (**Figure 9**). It has greater sensitivity for renal lesion detection and specificity in distinguishing simple renal cysts (which require no further workup) from suspected renal neoplasms (**Table 1**) (Bosniak *et al*., 1986).

Occasionally, a CT examination which is indicated in the initial workup of one clinically suspected illness reveals instead a different aetiology, as in the cases of two patients who presented with a palpable right lower quadrant mass (**Figure 10**). A CT scan is frequently requested to investigate further newly discovered radiographic abnormalities (**Figure 11**). Finally, CT has a role in the investigation of an unknown primary malignancy in a patient presenting with radiographic findings indicative of a metastatic lesion (**Figure 12**).

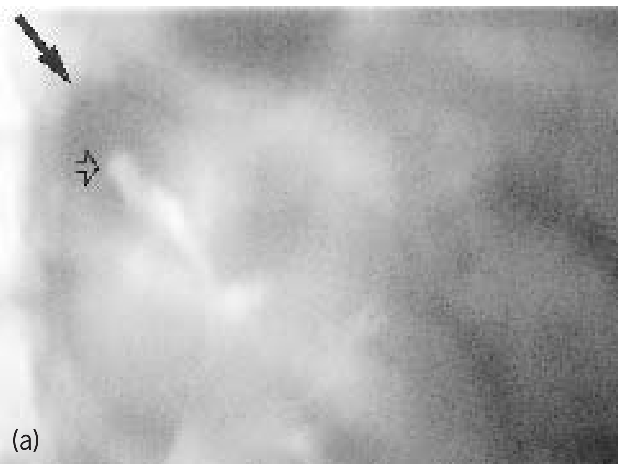
Lesion Characterization

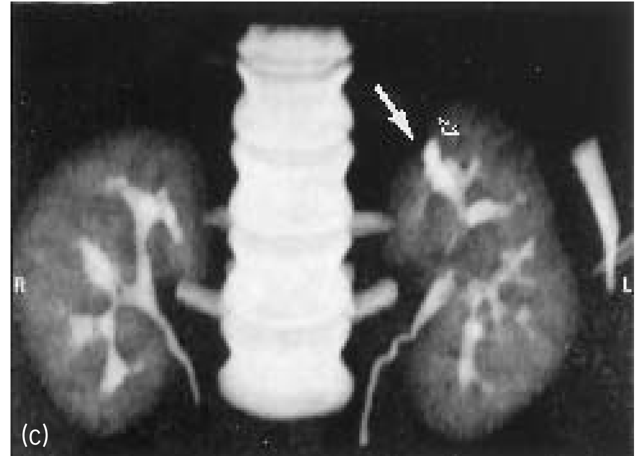
Computed tomography not only possesses increased sensitivity for overall lesion detection over conventional radiography, but in some cases also improved specificity. Noncontrast CT, owing to its sensitivity in discriminating tissues of differing X-ray attenuations, may be sufficient to characterize a lesion measuring at least 1 cm as benign, thus obviating the need for further workup. Detection of fat in a pulmonary mass is pathognomonic for a hamartoma, and in the kidney, an angiomyolipoma. The presence of a smooth-walled completely or centrally calcified pulmonary nodule is diagnostic of a benign granuloma. A round, thin-walled lesion composed of homogeneous fluid (up to 20 HU) is indicative of a benign cyst (see **Figure 21e**). These may occur anywhere within the body and are often discovered incidentally.

Dedicated CT examinations are performed to optimize detection and characterization of masses within the kidney (**Tables 1** and **2**) (Bosniak, 1986, 1997; Aronson *et al*., 1991; Birnbaum *et al*., 1996; Szolar *et al*., 1997; Yuh and Cohan, 1999), lung (**Table 3**) (Yamashita *et al*., 1995; Zhang and Kono, 1997; Swensen *et al*., 2000), adrenal (**Table 4**) (Korobkin *et al*., 1996, 1998; Outwater *et al*., 1996; Boland *et al*., 1997, 1998; Macari *et al*., 1998) and liver (**Table 5**) (Weg *et al*., 1998; Inaba *et al*., 2000; Hollett *et al*., 1995; Baron *et al*., 1996; Miller *et al*., 1998; Chen *et al*., 1999; Murakami *et al*., 2001). Helical CT scanning is used to reduce data misregistration and partial volume averaging with the use of single-breath hold thin-section volumetric acquisitions through an entire organ (liver, kidney) or lesion (pulmonary nodule) prior to and during multiple predetermined phases of intravenous contrast enhancement. The generated data are free of motion artifact and permit comparison of identical levels on scans obtained before and after contrast administration. Partial volume averaging occurs when portions of two adjacent structures are in the same voxel creating an incorrect CT number. This is minimized because a section through the centre of a lesion is assured with helical CT when overlapping sections are reconstructed.

Accurate CT differentiation of a renal neoplasm from a simple (**Table 1**) or a mildly complicated renal cyst (**Table 2**) requires thorough lesion analysis on both pre- and postcontrast scans (Bosniak, 1986, 1997; Aronson *et al*., 1991; Birnbaum *et al*., 1996; Szolar *et al*., 1997; Yuh and Cohan, 1999). Because renal masses do not contain functioning nephrons, enhancement, when present, is generally indicative of neovascularity and malignancy. Enhancement of solid renal masses may be more pronounced on the corticomedullary (vascular) phase which usually occurs between 25 and 80 s after initiation of contrast injection. However, almost all masses have decreased attenuation (and so are more conspicuous) relative to normal renal parenchyma during the other two phases. The nephographic phase usually begins 85–120 s after the onset of injection when contrast filters through the glomeruli and enters the loop of Henle and the collecting ducts. The excretory phase begins when the contrast is first excreted into the calyces, approximately 3–5 min after the onset of injection (**Figure 13**) (Yuh and Cohan, 1999).

A similar protocol is used to distinguish a benign from a malignant solitary pulmonary nodule based upon their





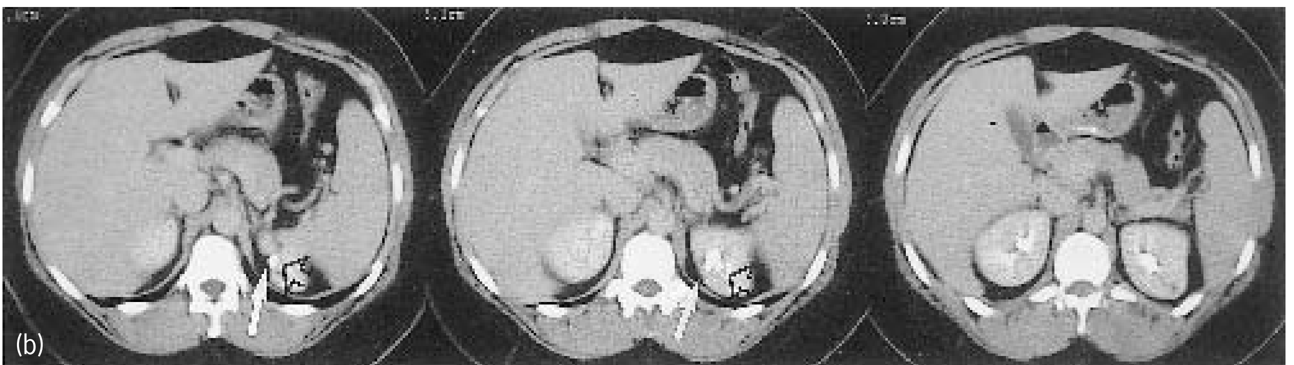


Figure 9 IVP compared with volume-rendered 3D image from the CT scan. (a) Cortical thinning representing a scar in the upper pole of the left kidney (solid arrow) and calyceal blunting (open arrow) typical of reflux nephropathy are seen on this linear coronal nephrotomogram from the conventional IVP. (b) Serial axial CT images also demonstrate scarring in the medial aspect of the upper kidney (solid arrows) but lack the same detail of the collecting system (open arrows). (c) A volume-rendered coronal 3D image displays the same information and detail regarding the renal parenchyma (solid arrow) and collecting system (open arrow) as seen on the IVP. Currently either an IVP combined with a CT or a CT with reconstruction is recommended to evaluate patients with haematuria, or other renal problems, which may affect either the collecting system or the parenchyma rather than the IVP alone.

Table 1 CT characterization of renal cysts vs tumours

Cysts	Malignant tumours
Near water density (0–20 HU)	Greater than water density (regions)
No or thin wall calcification	Coarse wall calcification
Do not enhance	May enhance in regions

differences in vascularity (Yamashita *et al*., 1995; Zhang and Kono, 1997; Swensen *et al*., 2000). To be eligible for this examination a nodule must measure 5 mm or more in diameter, be solid, relatively spherical, homogeneous, without calcification or fat on nonenhanced, thin-section CT images. Helical scanning is performed through the lesion pre and at 1, 2, 3 and 4 min after the onset of contrast injection (**Figure 14; see colour plate section for parts (h) and (i)**). Using enhancement greater than 15 HU as a marker for malignancy, the most recent multicentre study reported 98% sensitivity, 58% specificity and 77% accuracy (Swensen *et al*., 2000).

The initial diagnostic CT scan can provide additional pertinent information for a patient's subsequent workup and treatment plan, including clinical tumour staging (both local extent and the presence or absence of distant metastases) and determination of the best site and modality for histological tissue sampling. It also serves as a baseline examination for comparison with subsequent studies.

Tumour Staging

CT scanning is the established modality for staging virtually all tumours. CT is routinely employed for the preoperative staging of non-small cell lung cancer

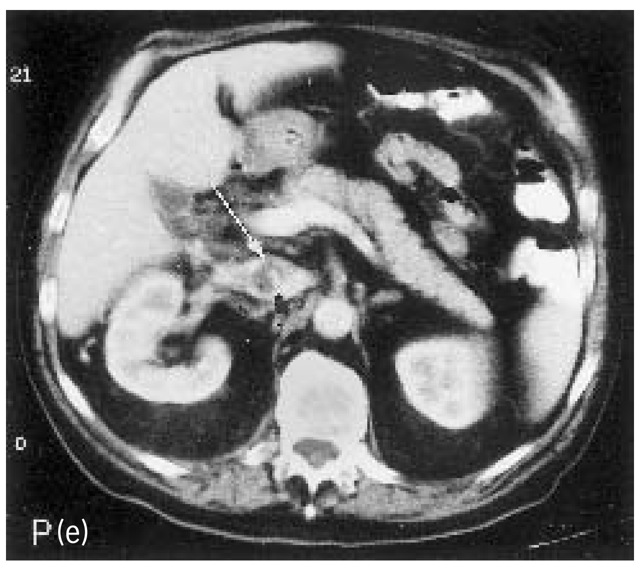
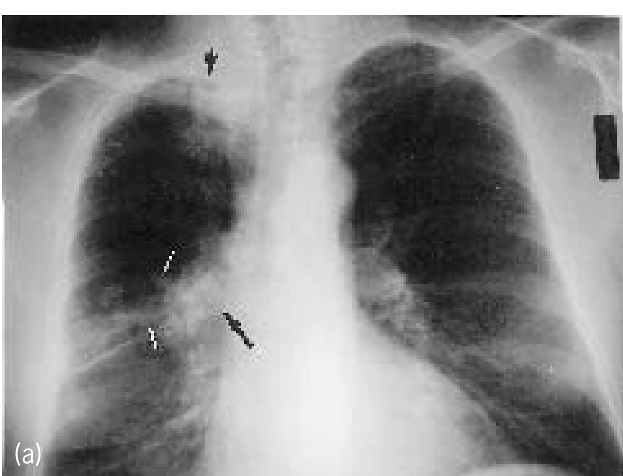
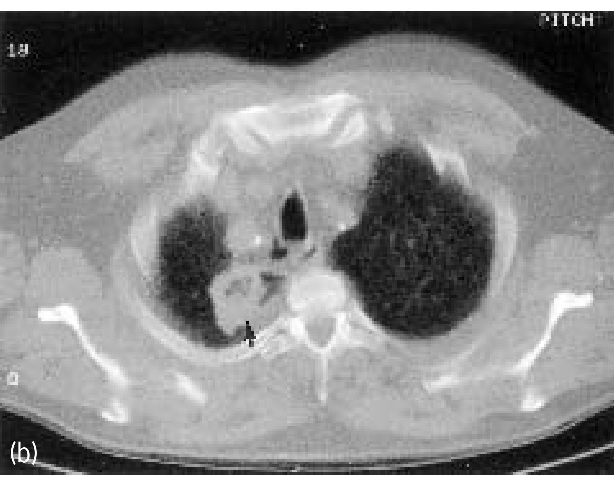


Figure 10 Abdominal CT scans performed for palpable right lower quadrant masses. (a) The caecum is poorly coated by the oral contrast (curved arrow) owing to the presence of a soft tissue mass (short arrow). This proved to be a caecal carcinoma. The acute clinical presentation resulted from perforation as evidenced by the presence of extraluminal air (long arrow). (b) An adjacent axial image in the same CT scan reveals extraluminal fluid (open arrow). Note the abnormal local lymphadenopathy (tiny arrows). (c) An axial CT image in another patient reveals a large mass (large arrow), with central fluid-appearing areas compatible with necrosis (small arrows) extending into the caecum (curved arrows) accounting for the palpable abnormality. (d) A scan cranial to (c) reveals that the mass (solid arrows) originates in the right kidney and is a pathologically proven renal cell carcinoma. The right renal vein (open arrows) is enlarged and contains solid material compatible with thrombus extending into the inferior vena cava (curved arrow). (e) Another axial image cranial to (d) demonstrates enhancing (tumour) thrombus (arrowhead) within the inferior vena cava (IVC) (long arrow). This information is vital for surgical planning. If thrombus is detected at the level of and distal to the intrahepatic IVC, a cardiac surgeon may be required during the surgery to excise tumour from the right atrium.





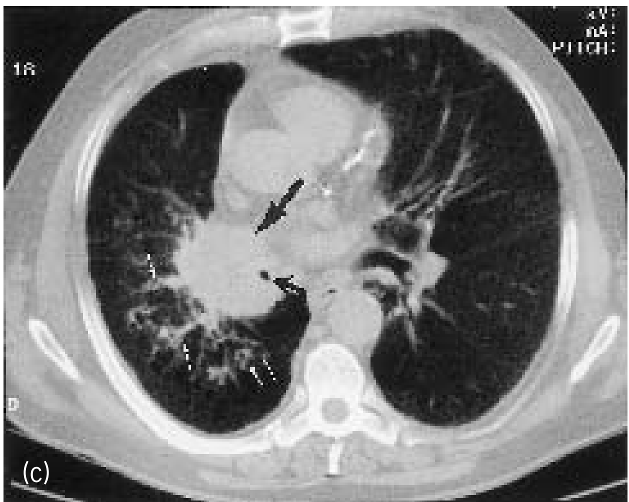
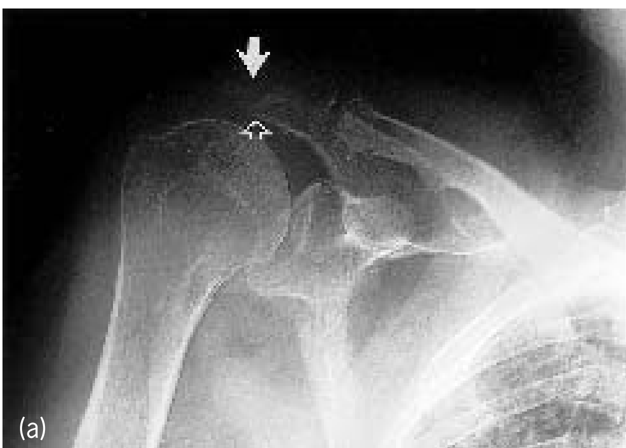


Figure 11 CT scan ordered for abnormal or suspicious radiographic findings. (a) A chest radiograph performed as part of the preoperative workup for coronary artery bypass graft surgery reveals a right apical (possibly cavitary) opacity (short black arrow), and an enlarged, dense right hilum (long black arrow) with adjacent thickened interstitial markings (white arrows). (b) A non-contrast CT image confirms the presence of the apical cavitary mass (black arrow). (c) Axial image at the level of the hila. A right hilar mass (long black arrow) encircles and narrows the bronchus intermedius (curved arrow). CT better demonstrates abnormally thickened perihilar interstitial markings compatible with lymphangitic tumour spread (white arrows). Pathology revealed a primary squamous cell lung carcinoma.



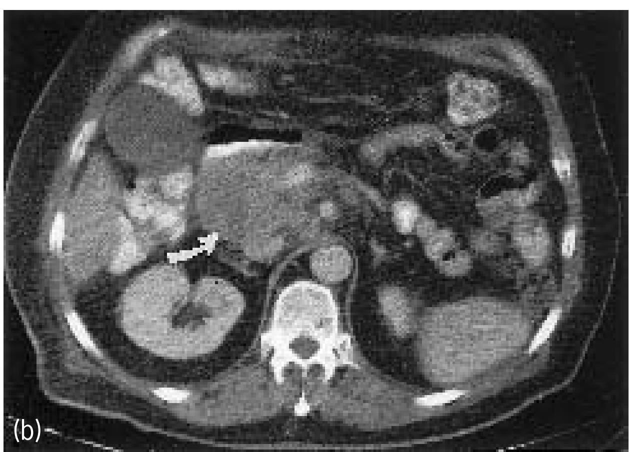


Figure 12 Clinical presentation of a metastatic lesion in the absence of a known primary. (a) A shoulder radiograph ordered because of pain in the absence of trauma demonstrates destruction of the distal acromion (open arrow) with bony fragments in the adjacent soft tissues (closed arrow). Biopsy revealed metastatic adenocarcinoma. (b) Abdominal CT identifies the primary neoplasm, a mass in the pancreatic head (curved arrow).

(NSCLC) (Patz *et al*., 1999). It provides information regarding the lung mass(es) including size and location (lobe, relationship to bronchi), the presence and location of suspicious lymphadenopathy and mediastinal, pleural, chest wall (**Figure 15a**) and local spinal involvement (**Figure 15b**).

Chest CT examinations routinely include the adrenal glands since up to 10% of NSCLC patients will have an adrenal mass as the only site of extrathoracic disease at initial presentation (Macari *et al*., 1998). Differentiation of an adrenal adenoma from a nonadenoma is necessary in order to prevent curative surgery in the presence of extrathoracic metastasis (**Table 4**) (Korobkin *et al*., 1996, 1998; Outwater *et al*., 1996; Boland *et al*., 1997, 1998; Macari *et al*., 1998). Scanning through the adrenal glands helps to ensure complete coverage of the lungs even in the presence of respiratory variability and/or hyperinflation secondary to emphysema. In the latter case, CT demonstration of extensive emphysematous lung destruction may suggest clinical inoperability secondary to an insufficient pulmonary reserve despite tumour resectability. As with pancreatic carcinoma, surgical resection provides the best chance for survival in patients with NSCLC and so extensive surgical procedures (rather than the standard lobectomy and mediastinal node dissection) may be considered. Preoperative CT scanning can not only suggest resectability, it can also influence the nature of surgical procedure itself. Pneumonectomy is indicated if the mass involves a main bronchus, chest wall involvement may require resection of multiple ribs and in some centres a solitary adrenal metastasis is resected at the time of the thoracic resection.

Preoperative identification of haematogenous or distant metastases is clinically useful even when resection of a primary tumour is indicated regardless of its stage. The role of CT in preoperative staging of colon cancer has been controversial with overall accuracies reported in the range 48–77% (Horton *et al*., 2000). These results are due in part, to the inability of CT to demonstrate microscopic tumour extension and to distinguish normal-sized lymph nodes with tumour from those without tumour (Zerhouni *et al*., 1996; Boiselle *et al*., 1998). These limitations are common to all CT oncological imaging and have contributed to the use of complementary imaging modalities such as magnetic resonance imaging (MRI) and positive emission tomography (PET) in addition to CT (Zerhouni *et al*., 1996; Abdel-Nabi *et al*., 1998; Boiselle *et al*., 1998). Indications for preoperative CT in colon cancer include suspected haematogenous or distal nodal metastases, suspected invasion into adjacent organs or abscess formation, atypical

Table 2 Bosniak classification of renal cystic masses

Category I
Simple, uncomplicated benign cyst
Category II
Minimally complicated cyst that is benign with
one or more of the following:
one or more thin septae
thin areas of calcification in the cyst wall or septation
fluid within the cyst has a higher attenuation than clear fluid
Category III
Complicated cyst with some findings seen in
malignant lesions including:
septal wall irregularity, thickness >1 mm, or associated solid elements at its attachment
multiple locules
more extensive areas of calcification, which are thick and/or irregular
inhomogeneous cyst fluid
Category IV
Cystic carcinoma
irregular margins
nonuniform wall thickening
solid component
enhancing nodularity

Table 3 CT characterization of a pulmonary nodule

	Benign	**Malignant**
Noncontrast		
Contour	Smooth, well defined	Irregular, spiculated
Composition		
Calcifications (if present)	Central, smooth margins	Peripheral, irregular margins
Fat	May be present	Absent
Periphery opacities	Adjacent lung parenchyma	Variable hazy
	Appears normal	Secondary to desmoplastic reaction
	Pleural tags absent	May be present
Postcontrast		
Enhancement	<15 HU	>20 HU
	(between 16 and 20 HU enhancement is indeterminate)	

Table 4 Phases of hepatic CT in malignancy

Phases	Time after onset of injection	Abnormalities best seen
Precontrast		Calcifications
Postcontrast		
Arterial	20–25 s	Hepatoma Hypervascular metastases
Late arterial	40 s	Hepatoma (some best seen here)
Portal venous	70 s	Most metastases
Equilibrium	8–12 min	Cholangiocarcinoma (delayed enhancement)
Delayed	4–6 h	Any mass (hepatic parenchyma accumulates iodine)

Table 5 CT characterization of adrenal masses

	Adenoma	Malignancy
Precontrast		
Size	<4 cm in diameter	Primary tumours usually >5 cm
Composition	<10 HU (MR is indicated when 10–20 HU)	>20 HU
Postcontrast	Rapid washout <37 HU at 30 min	Delayed washout >40 HU at 30 min

symptoms or unusual histology. The liver is the most common site for distant metastases. Their preoperative identification may be of importance as a solitary metastasis may be resected concurrently with the primary tumour. Dynamic contrast-enhanced CT and non-enhanced MRI both have 85% accuracy in diagnosing hepatic metastases (Zerhouni *et al*., 1996). Multidetector scanning (Weg *et al*., 1998) and CT arterial portography with or without CT hepatic arteriography have improved detection, with accuracies of 89–94% (Inaba *et al*., 2000).

A dedicated CT liver protocol is helpful in characterizing hepatic masses (Miller *et al*., 1988; Hollett *et al*., 1995; Baron *et al*., 1996; Chen *et al*., 1999; Murakami *et al*., 2001). Most hepatic metastases including those secondary to colon cancer are hypoattenuating masses best seen during the portal venous phase of enhancement as opposed to other liver lesions (**Figure 16**) (**Table 5**). Malignancies with a greater hepatic arterial blood supply and arterioportal shunting are hyperintense on the arterial phase (Chen *et al*., 1999). These include hepatomas (**Figure 17**) and hypervascular metastases secondary to carcinoid, pancreatic islet cell tumour, melanoma, thyroid carcinoma, phaeochromocytoma and breast and renal carcinoma (Miller *et al*., 1988; Hollett *et al*., 1995; Baron *et al*., 1996; Chen *et al*., 1999; Murakami *et al*., 2001).

Tissue Sampling

Histological confirmation of malignancy can often be made by a CT-guided transcutaneous needle biopsy (TCNB), obviating the need for an open biopsy. The initial diagnostic CT scan, by demonstrating suspicious primary or metastatic lesions, can assist in the selection of the most appropriate site and method for tissue sampling. In the case of a suspected lung cancer, which is central and communicates with a bronchus, sputum cytology or bronchoscopy with or without transbronchial biopsy may be appropriate. If the mass is peripheral, transthoracic needle biopsy (TTNA) may be performed under fluoroscopy if it is of sufficient size. If it is too small or surrounded by numerous bullae, CT guidance is preferable as it can reliably depict the mass and demonstrate the optimal trajectory, traversing the least amount of lung parenchyma and avoiding surrounding bullae if possible. Overall sensitivities of 82–100% and specificities of 94–100% have been reported for TTNA for carcinoma. The most common complication is a 5–57% risk of pneumothorax, 1.6–17% of which require a chest tube (Wescott *et al*., 1997; Dennie *et al*., 2001). CT fluoroscopy is currently available (Daly *et al*., 1999). Its utilization may reduce complications by providing continuous real-time monitoring, thus decreasing the total procedure time and the actual time during which a biopsy needle is in the lung. Aside from the histological confirmation of malignancy, differentiation between non-small-cell (NSCLC) and small-cell lung cancers is critical owing to their differences in treatment options. Small-cell cancers are generally effectively treated with chemotherapy alone, thus obviating surgical considerations.

CT-guided TCNB of a suspicious metastasis, even if the primary neoplasm is large and accessible, provides

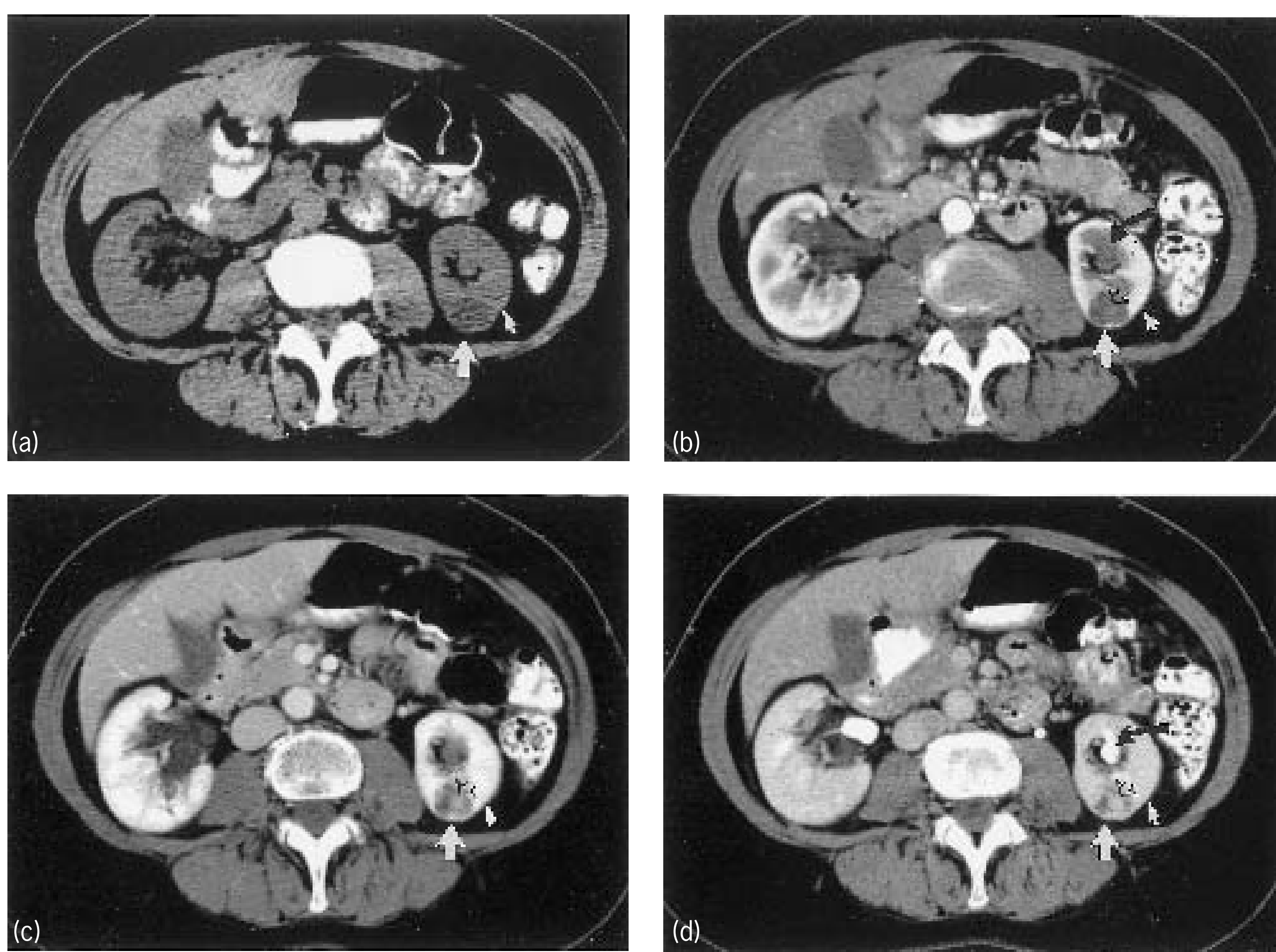


Figure 13 Multiphase imaging of a patient with left renal cell carcinoma. (a) Noncontrast CT though the kidneys demonstrates a mass in the posterior aspect of the left kidney (large arrow) which is hypodense relative to the normal renal parenchyma (small arrow). The noncontrast CT series is used to obtain baseline Hounsefield units (HU), density measurements, and to identify the presence of calcium, which may be obscured after injection of intravenous contrast medium. (b) Corticomedullary (vascular) phase imaging. The lesion is seen as a hypodense mass (large solid arrow) compared with the renal parenchyma (small solid arrow). Note the minimal enhancement within it (open arrow). The contrast has not reached the urine-collecting system which appears of lower attenuation (curved arrow). (c) Peak nephrographic phase imaging. Image at the same level performed just before the excretion of contrast medium shows loss of corticomedullary definition, which was present on the previous image and a more homogeneous parenchyma (small solid arrow). The lesion is well defined (large solid arrow), but exhibits greater enhancement (open arrow). (d) Excretory phase imaging. A delayed image done at the time of excretion of the contrast into the collecting system (curved arrow) shows a less intense nephrogram (small solid arrow) than that seen just prior to excretion but there is more intense tumoral enhancement (open arrow). Maximum tumoural enhancement can be seen at any phase, but is most typically seen on delayed images. Since maximum enhancement can be seen on any phase, multiphase imaging is a great aid in defining and enhancing borderline cases. In addition, different phases are needed to optimize the imaging of multiple organs. For example, the early phases (arterial and portal venous) are best to identify hepatic metastases, while the delayed (nephrographic and excretory) phases are superior for identification of renal lesions and of enhancement in them.

staging information. Determination of mediastinal nodal involvement is crucial for staging NSCLC which is required for appropriate prognostic assessment and therapeutic determination (Mountain, 1997; Mountain and Dresler, 1997). CT diagnosis of mediastinal disease is based upon nodal size. Utilizing 1 cm as the upper limit of normal for the short-axis dimension of a node, CT has been reported to have 62% sensitivity and

64% specificity in assessing lymph node metastases (McLoud *et al*., 1992). Its low sensitivity is secondary to the high frequency of microscopic metastases within normal-sized nodes, and its low specificity is due to the frequent occurrence of enlarged, hyperplastic nodes (Patz *et al*., 1999). Nevertheless, CT identification and localization of enlarged nodes can assist in the selection of the appropriate invasive procedure (mediastinoscopy, mediastinotomy, thoracoscopy and transbronchial biopsy) for staging.

Lymph nodes in the aortopulmonary window are generally not accessible with cervical mediastinoscopy and thus necessitate either parasternal or anterior mediastinotomy for biopsy (Boiselle *et al*., 1998). TTNA of an enlarged mediastinal node, if possible, expedites treatment initiation by eliminating the performance of one of the other biopsy procedures. Biopsy of a probable intra-abdominal or skeletal metastasis with CT-guided TCNB has the additional benefit of eliminating the risk of pneumothorax associated with TTNA. CT guidance has a role even in the biopsy of a superficial mass by demonstrating its internal architecture. The site for tissue sampling can be optimized away from areas of low yield such as necrosis, periosteal reaction or callus formation in a pathological fracture. Adjacent major vessels or nerves can also be avoided. CT-guided TCNB has a role in surveillance by confirming disease recurrence (**Figures 18** and **19**). This enables expeditious institution of appropriate therapy (Som *et al*., 1999).

Percutaneous CT sampling of other primary malignancies is equally safe and efficacious. Biopsy of abdominal primaries yields an approximately 95% positive result when malignancy is present with virtually

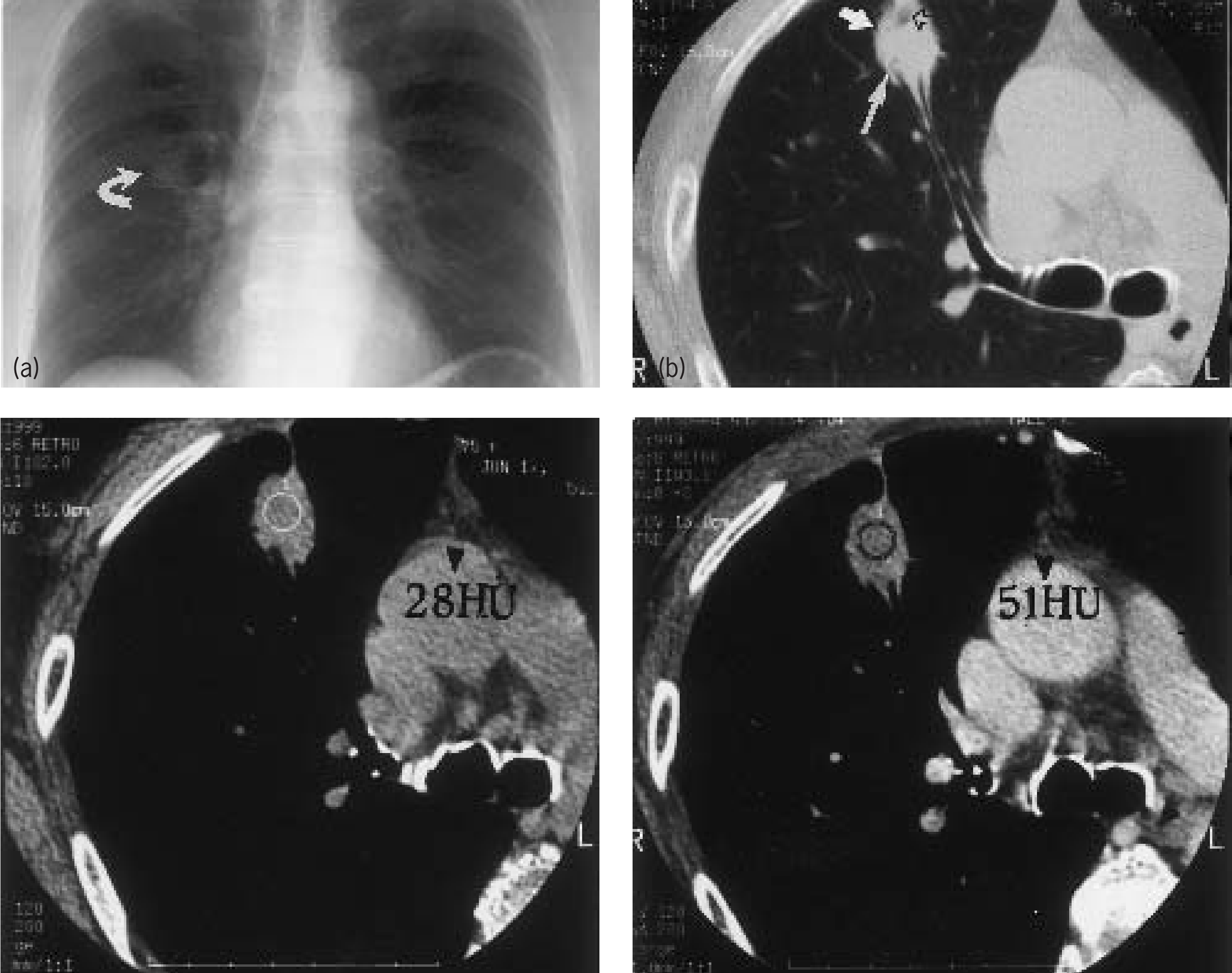
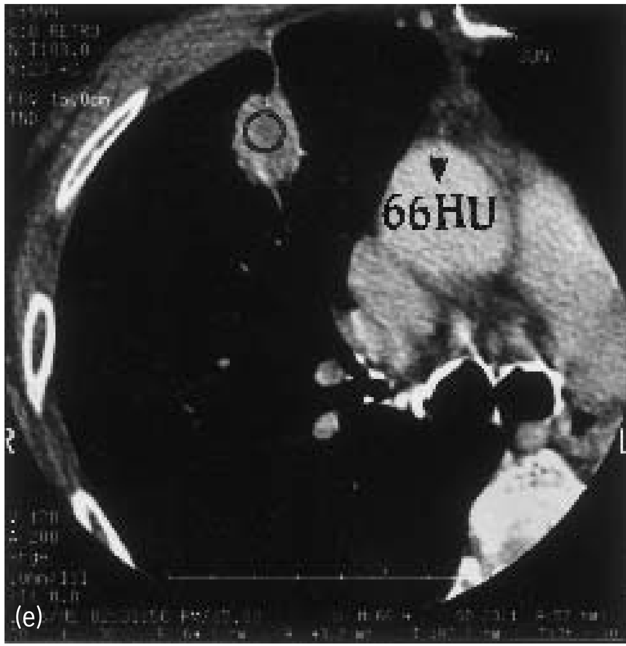
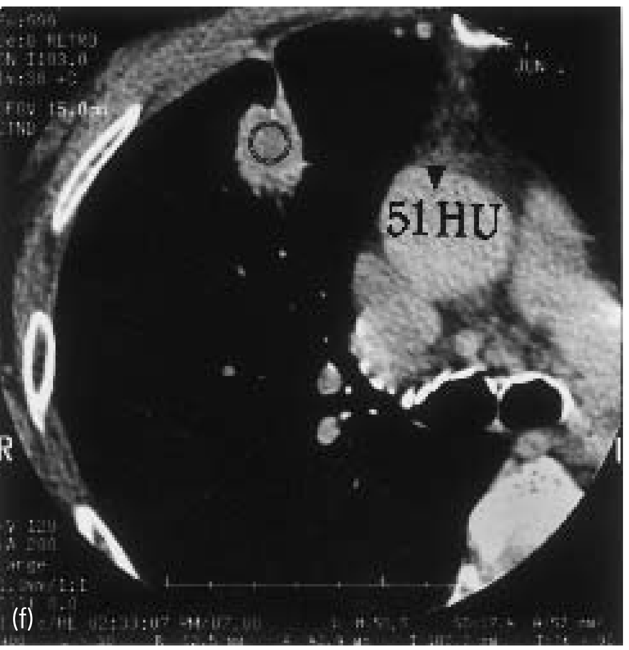


Figure 14 (Continued)





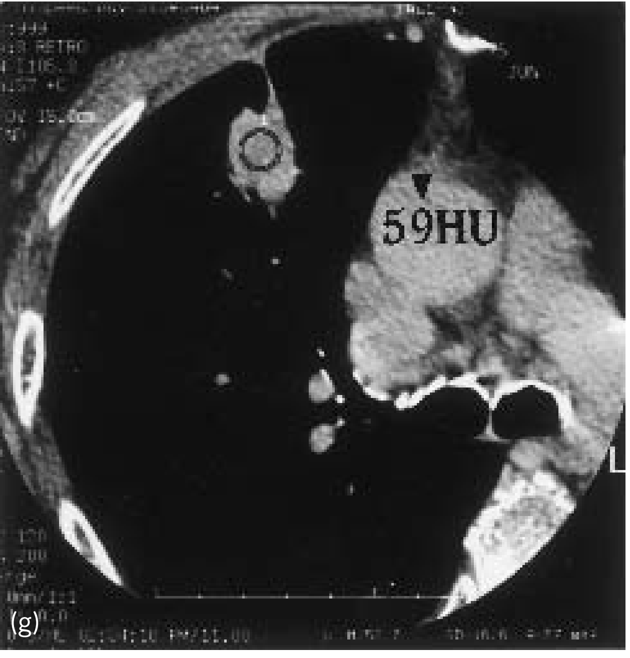
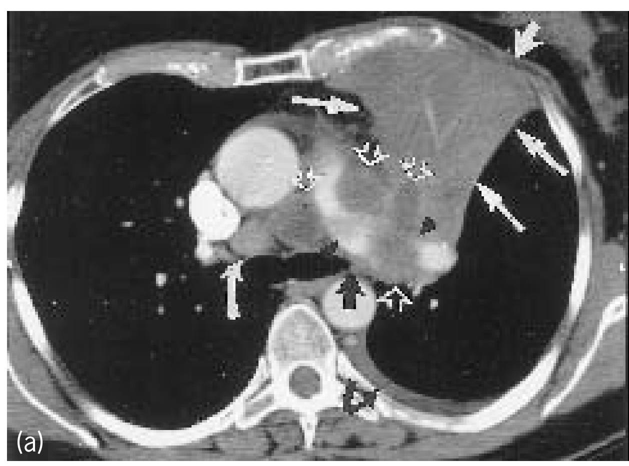


Figure 14 Imaging workup of a solitary pulmonary nodule, a biopsy-proven lung cancer. (a) A frontal chest radiograph reveals a solitary noncalcified nodule in the right lung (curved arrow). (b) A high-resolution CT scan provides further morphological detail. The nodule is fairly homogeneous, without calcifications, but with a focal luscency (open arrow) and an air bronchogram (long arrow). Its margins are irregular (short arrow). (c)–(g) CT evaluation of nodule enhancement. The white circles demarcate the region of interest used for measuring the nodule's density: at baseline (c) prior to the injection of intravenous contrast, medium and at 1 min (d), 2 min (e), 3 min (f) and 4 min (g) after the onset of the injection. The measurements in Hounsefield units (HU) are noted (arrowheads). Abnormal enhancement of 31 HU is present at 4 min, indicative of a malignant lesion. See colour plate section for parts (h) and (i).

no false-positive results and rare complications such as bleeding (Sundaram *et al.*, 1982). As with lung biopsies, the studies can be done on an outpatient basis. In addition to initial evaluation of neoplastic processes, percutaneous CT biopsies can be used to confirm recurrent disease, permitting optimal treatment.

Surgical Planning

CT plays a major role in surgical planning. In pancreatic carcinoma, which has a poor prognosis with a survival rate of 20% in the first and less than 5% in the fifth year, surgical resection offers the only chance for cure with a 20% 5-year survival. Only 20% of patients will have disease deemed surgically resectable at the time of diagnosis (O'Malley *et al.*, 1999). The ability to identify these patients accurately is essential, as there is a 20–30% surgical morbidity (Sheridan *et al.*, 1999). The role of CT scanning is for tumour detection and preoperative assessment of disease extent. Although its accuracy for tumour detection approaches 100%, determination of resectability is more problematic. Up to 25–50% of patients considered preoperatively to



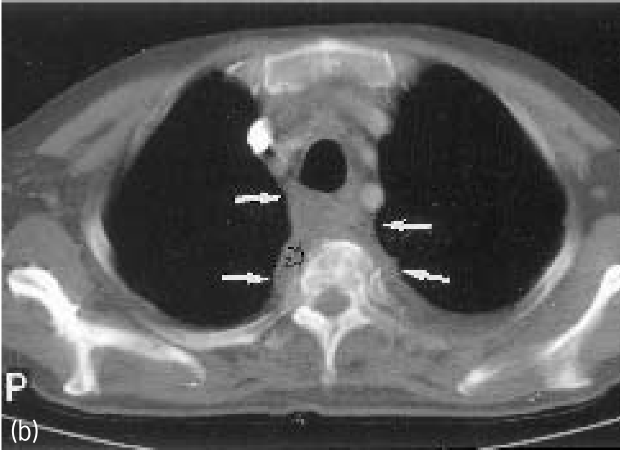
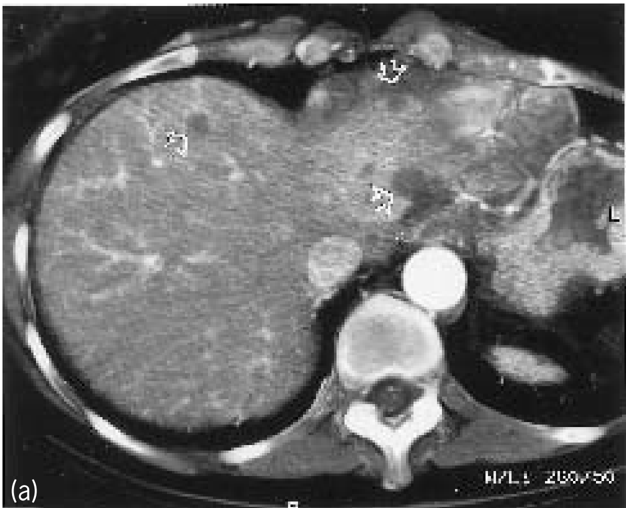


Figure 15 Evaluation of lung tumours for resectability. (a) Initial contrast-enhanced chest CT in a patient with newly diagnosed non-small cell carcinoma of the lung. The left main bronchus terminates abruptly at the level of the upper lobe bronchus (black arrow) by extensive necrotic (low attenuation) central (left hilar and mediastinal) masses (open arrows), with resultant left upper lobe collapse (long arrows). There is encasement of the left main pulmonary artery (arrowheads) as well as contralateral lymphadenopathy (curved white arrow). More peripherally, there is chest wall extension (broad arrow) and a small pleural effusion (curved black arrow). (b) Axial image from a thoracic CT of another patient with lung cancer reveals a mediastinal mass (solid arrows) with destruction of a contiguous vertebral body (open arrow).



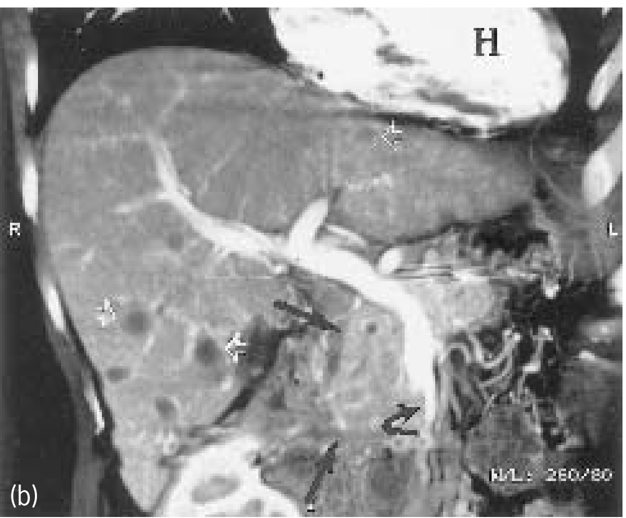


Figure 16 Hepatic metastases from pancreatic carcinoma. (a) Representative axial image through the liver demonstrates low-density hepatic metastases (open arrows). This is performed in the portal venous phase when most hepatic metastases are best appreciated as being less dense than normal parenchyma. (b) 2D coronal reconstruction image demonstrates multiple hepatic metastases (open arrows) as well as the pancreatic mass (straight black arrows). Note the involvement of the superior mesenteric vein (curved arrow) in the region of the mass. H = heart.

have resectable disease on CT were unresectable at surgery (O'Malley *et al*., 1999). Criteria include spread of disease outside the pancreas (usually the liver or peritoneum), contiguous invasion of adjacent organs including the stomach or colon and involvement of major peripancreatic vessels – arterial encasement or venous occlusion (**Figure 20**). A dedicated pancreatic mass CT is performed utilizing thin-section dynamic multiphasic

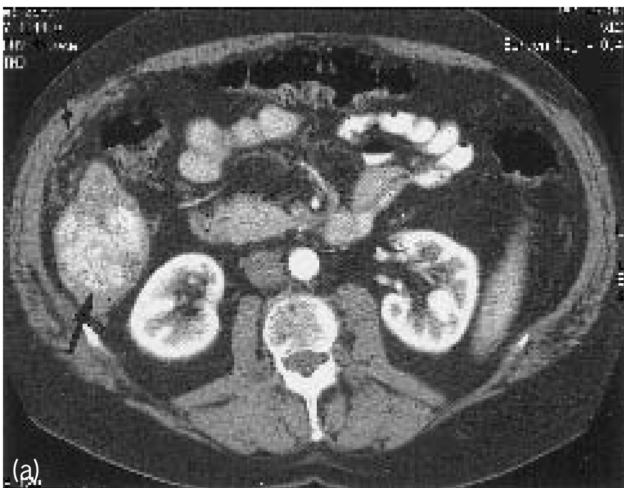
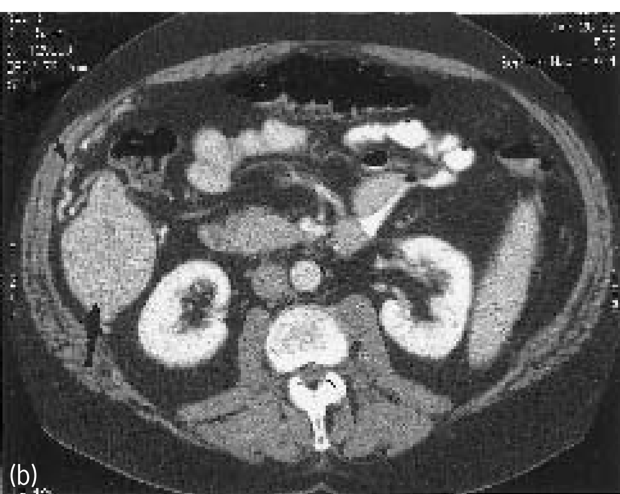
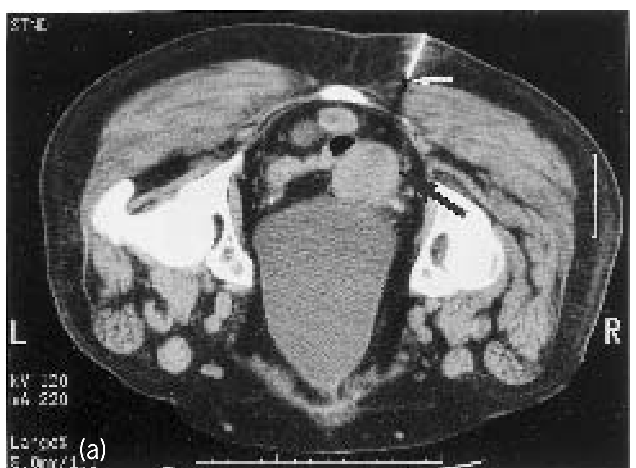



Figure 17 Multiphase imaging for the identification of a hepatoma. (a) Axial scan done through the right lobe of the liver during early imaging (arterial phase) demonstrates a hypervascular mass (large arrow) which enhances brightly on this phase. Maximum enhancement in the arterial phase is typical of hepatoma and certain other hypervascular tumours. In contrast, most hepatic tumours are best identified on the portal venous phase. Note the unenhanced collateral vessels in this cirrhotic patient (small arrow). (b) Same patient, scanning during the standard portal venous phase. The hypervascular tumour is not well appreciated (large arrow). Early scanning is necessary in patients with a suspicion of hypervascular tumours such as hepatoma. Most standard liver primary/secondary tumours are best seen on the portal venous phase. The collateral vessels have enhanced (small arrow).



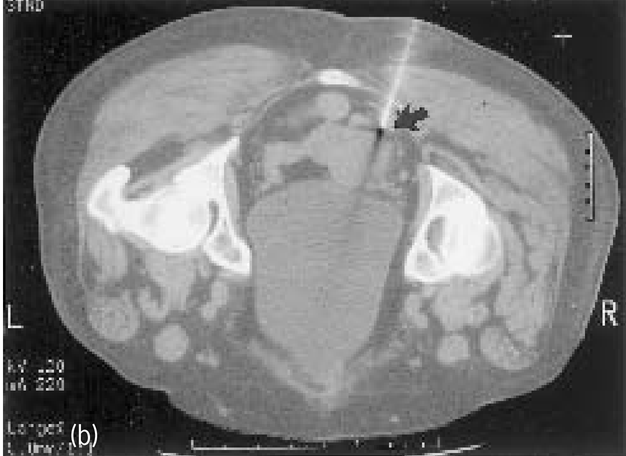
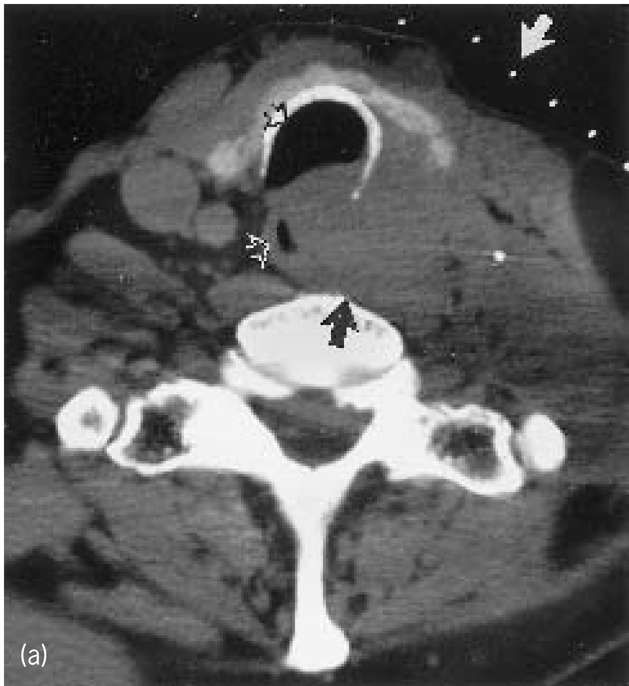
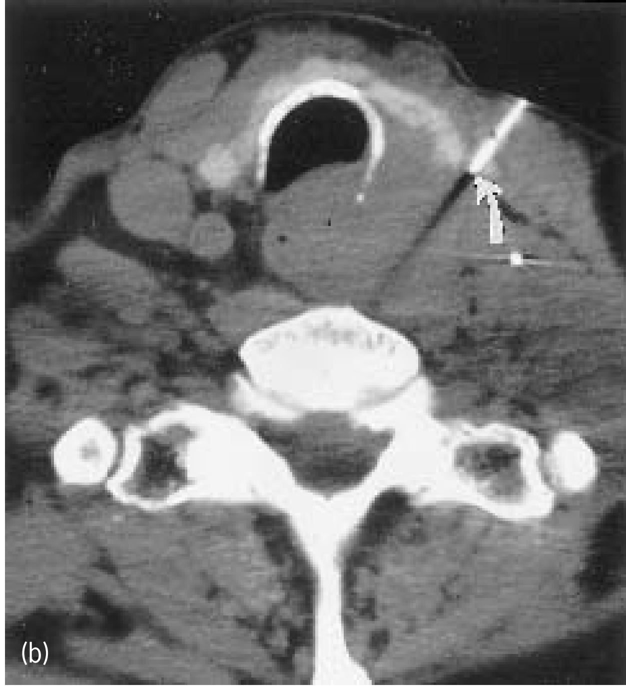


Figure 18 CT-guided biopsies. (a) CT-guided biopsy in the diagnosis of recurrent tumour. Patient with previous cervical carcinoma, who presented with a right pelvic mass (black arrow). On the initial image, the opaque needle is seen extending through the subcutaneous tissues. A black artifact extends from the needle tip (white arrow). (b) The needle has been advanced to the surface of the mass (arrow) permitting needle aspiration cytology or biopsy. This is a highly reliable technique with 95% sensitivity in patients having primary/secondary neoplastic disease. It is an out-patient procedure with extremely infrequent complications.

helical scanning (**Figure 21a**) (Nishiharu *et al.*, 1999; O'Malley *et al.*, 1999; Sheridan *et al.*, 1999). The acquired data set subsequently undergoes image analysis and interactive volume rendering. With multi-planar re-formation (MPR), reconstructed axial images are stacked to create a volume of imaging data from which two-dimensional images are generated (Ibukkuro *et al.*, 1995). A reference plane (coronal or sagittal) can then be selected for viewing (**Figure 21b and c**) and subsequent manipulation (tilted or rotated) (**Figure 21d and e**). CT angiography with MPR can display a vessel in a format similar to conventional angiography





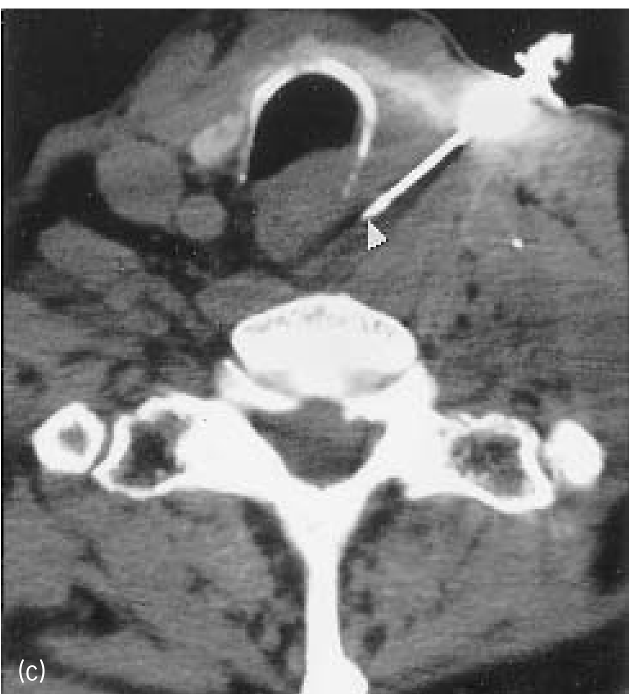
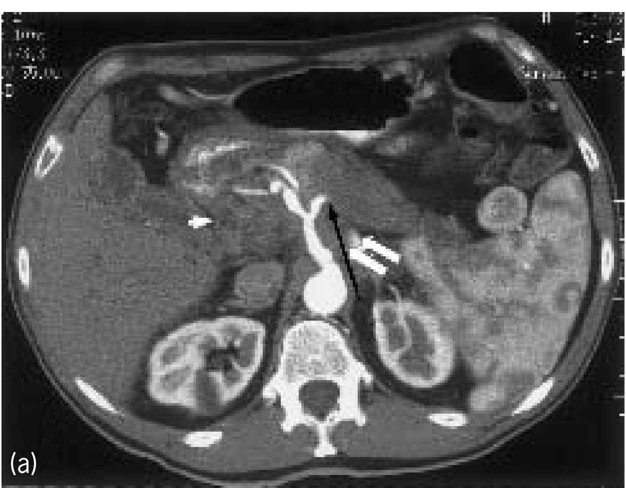


Figure 19 (a) CT-guided biopsy in the diagnosis of recurrent tumour in another patient. Patient with previous squamous cell carcinoma of the neck presented with a large soft tissue mass (black closed arrow) in the area of prior resection. It extends into the trachea (black open arrow) and abuts the oesophagus (open white arrow). Once a biopsy location is selected, a row of needles is placed on the skin. The appropriate entry site is selected relative to the corresponding needles (white closed arrow). (b) The biopsy needle is seen extending through the subcutaneous tissues. A black artifact extends from the needle tip (curved arrow). (c) The needle has been advanced into the mass with the tip successfully positioned (arrowhead) permitting needle aspiration cytology or biopsy. This is a highly reliable technique with 95% sensitivity in patients having primary/secondary neoplastic disease.

that is easier to understand than an axial CT image (**Figure 21f and g**). Furthermore, 2D and 3D (**Figure 21h–i**) vascular imaging techniques have the additional advantage over conventional angiography that they are capable of displaying a structure or mass surrounding a blood vessel in addition to a lesion inside the vessel. This is particularly helpful in pancreatic cancer imaging since many of the important anatomical landmarks lie in different planes (Novick and Fishman, 1998). Multiphase thin-section helical CT with 2D and 3D re-formation can assist in all stages of surgical planning for potentially resectable cases (**Figure 22**), or provide follow-up of a palliative procedure in others (**Figure 23**).



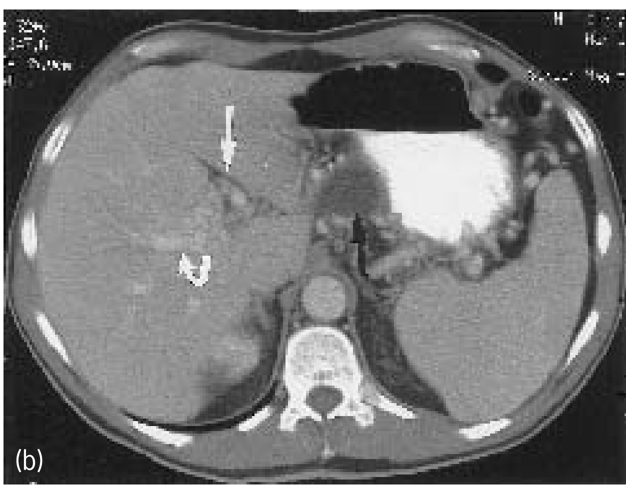
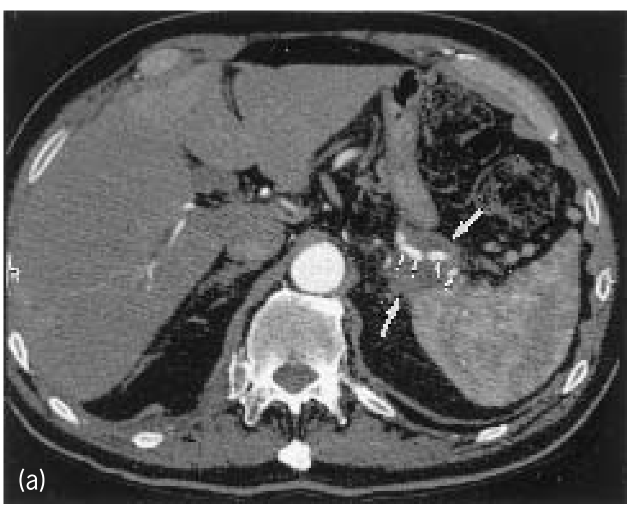
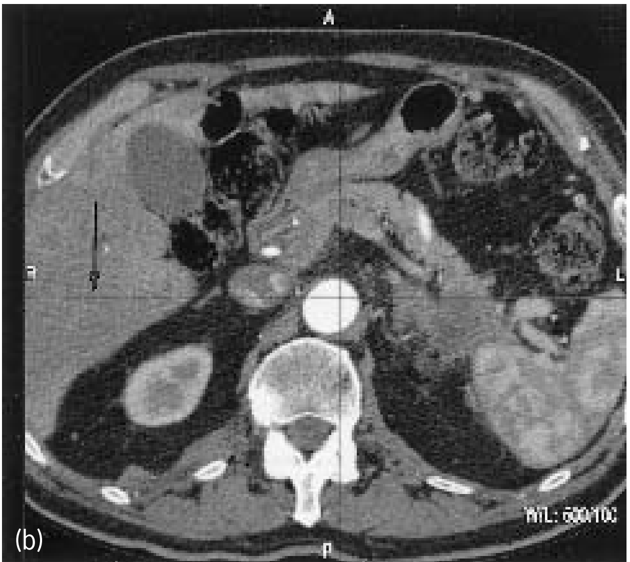


Figure 20 Local invasion. Pancreatic carcinoma with involvement of the hepatic artery, and obstruction of the portal vein with cavernous transformation. (a) Arterial phase imaging at the level of a pancreatic tumour (short white arrows) demonstrates the encasement of the splenic artery (long black arrow) and the failure to identify the portal vein. (b) A scan on a later phase through the liver demonstrates reconstitution of the right portal vein (curved arrow) and cavernous transformation of the main portal vein (white straight arrow). Note the necrotic adenopathy in the lesser sac (black arrow). CT produces exquisite detail of vessels and readily identifies local involvement of the structures.





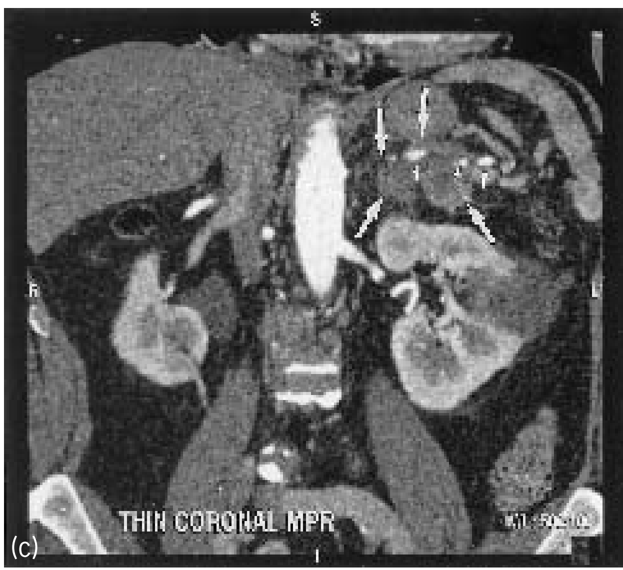


Figure 21 (*continued*)

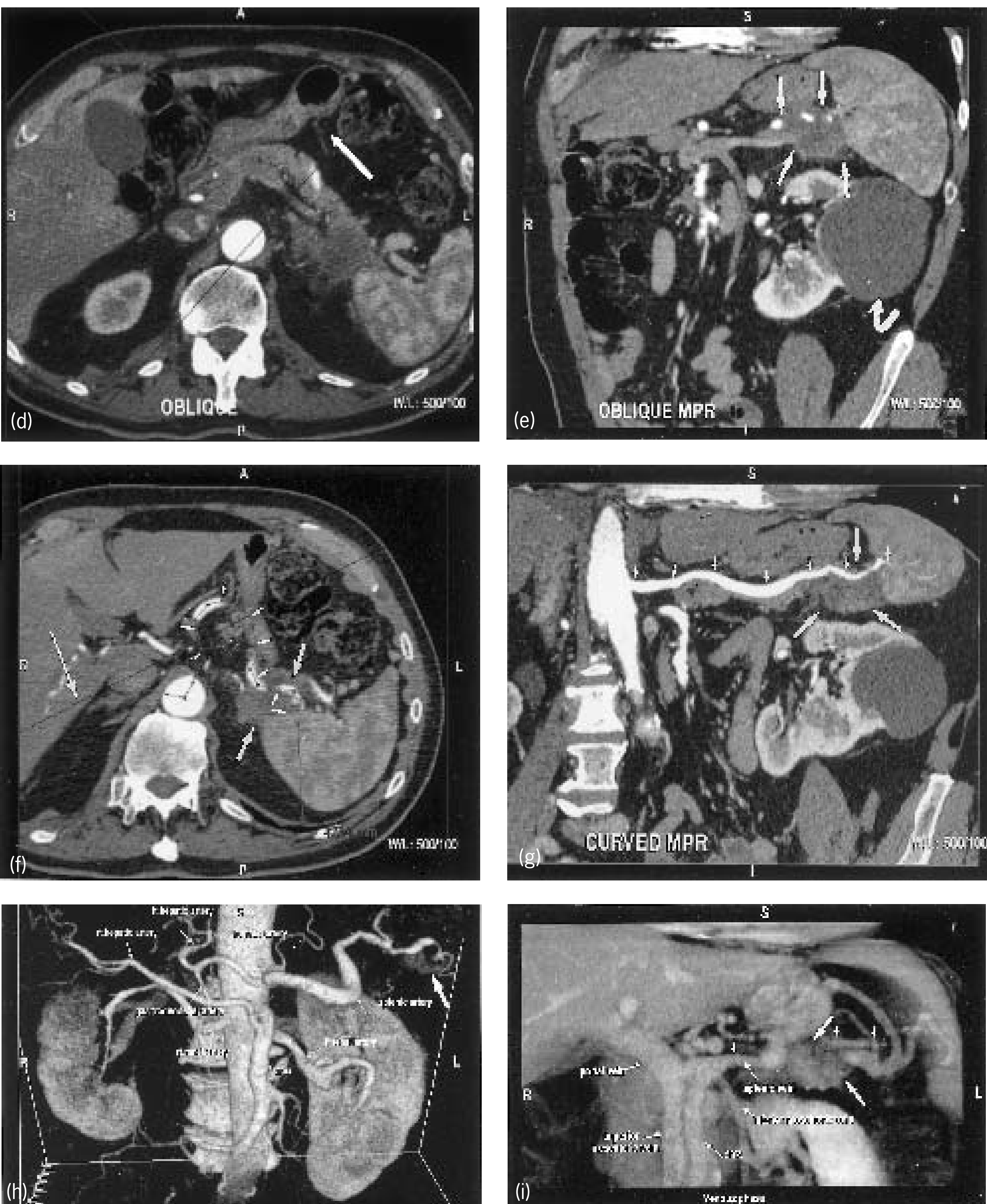


Figure 21 Utility of two-dimensional multiplanar re-formation (MPR) and three-dimensional helical CT angiography (CTA) in staging and surgical planning. (a) A thin section axial image in the arterial phase demonstrates a pancreatic tail mass (large arrows) involving the distal splenic artery (small arrows). (b),(c) A more caudal axial scan. Crosshairs have been placed for selection of the coronal reformation plane (long arrow in (b)). Coronal MPR image (c) demonstrates the distal splenic artery (small arrows) and its relationship to the mass (large arrows). (d),(e) Axial image (d) demonstrating selection of an oblique plane (long arrow) to produce an oblique MPR (e) with a more elongated presentation of the

(*continued*)

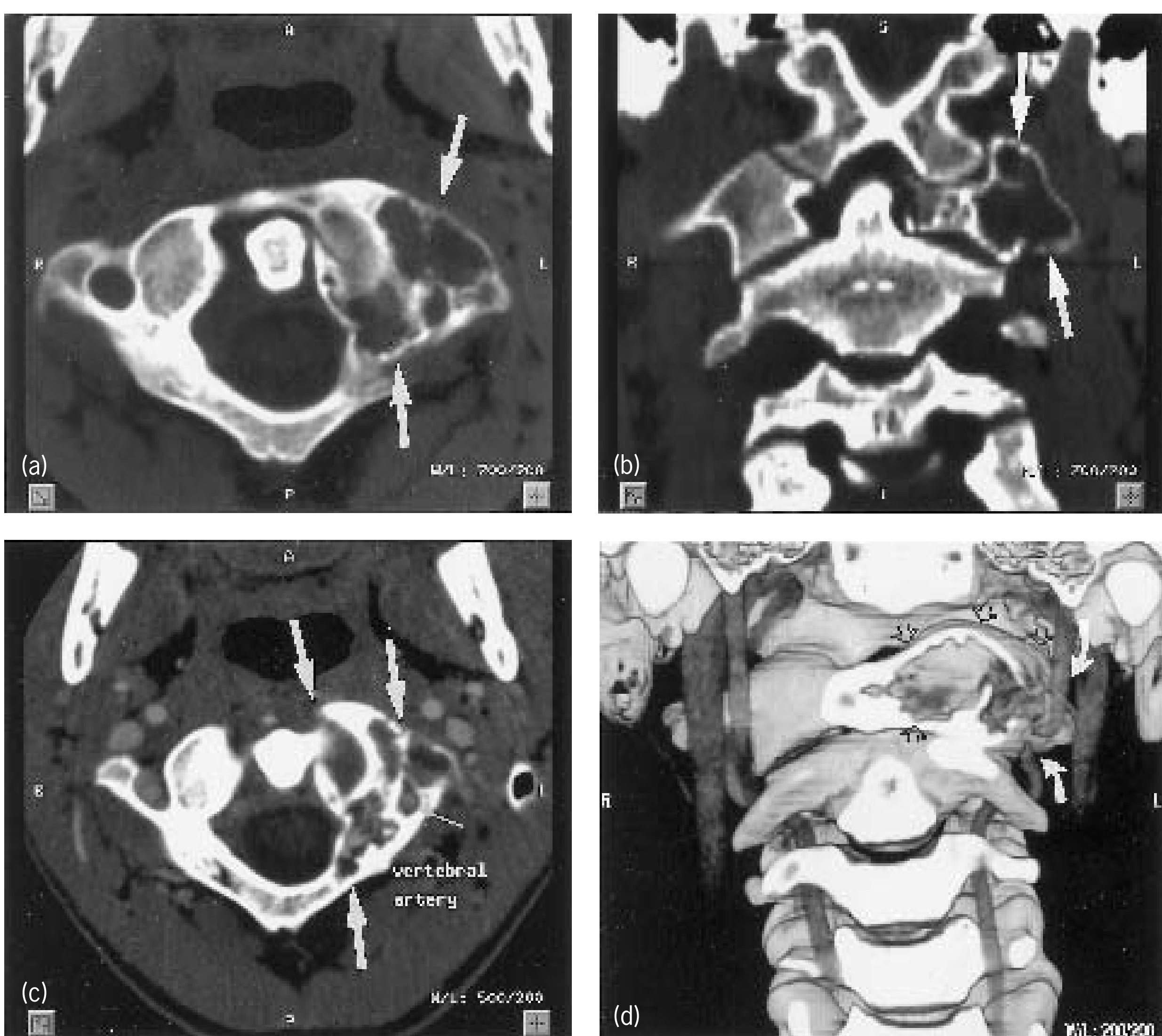
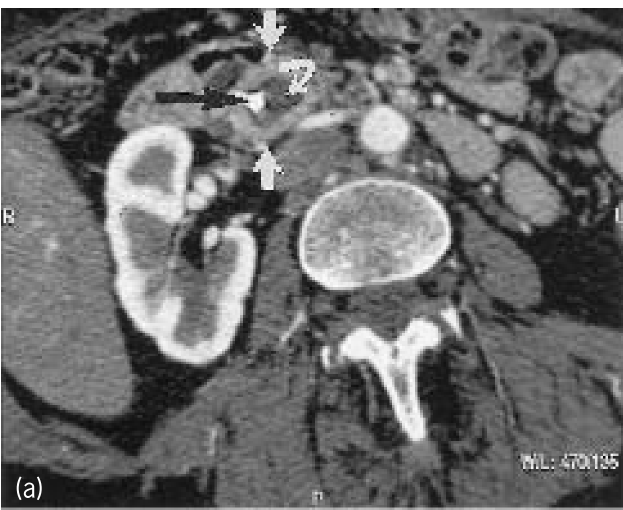


Figure 22 CT angiography with 3D reconstruction in a 17-year-old diagnosed with a vertebral osteosarcoma and referred for a CTA prior to a presurgical embolization of the left vertebral artery. (a),(b) Axial (a) and coronal (b) images demonstrate a lytic mass (arrows) in the C1 vertebral body representing the known osteosarcoma. (c) Axial image following injection of intravenous contrast reveals the location of the vertebral artery (yellow arrow) relative to the mass (white arrows). (d) Coronal 3D shaded surface area reconstruction exquisitely demonstrates the course of the vertebral artery (curved arrows) and its relationship relative to the tumour (open arrows).

Figure 21 (*caption continued*)
mass (straight arrows). Note the large left renal cyst (curved arrow). (f),(g) Curved re-formation. A line is drawn by the operator (long white arrow) on the axial scan (f) along the projected course of the splenic artery (small arrows) including its portion through the mass (large arrows identify mass). The result is a curved MPR (g) which displays the entire artery (small arrows) as well as the pancreatic tail mass (large arrows). (h) 3D volume-rendered coronal model of the arterial anatomy with shaded surface display (SSD). Pancreatic tail mass (large arrow). (i) Coronal 3D display image showing the venous anatomy. There is occlusion of the splenic vein in the area of the pancreatic mass (large arrows) with distal reconstitution (small arrows). These images provide better delineation of a neoplasm itself as well as the status of adjacent structures. Surgical planes can be virtually created to further assist in surgical planning. The arterial anatomy described in (h) is in actuality figure (i) and the venous anatomy described under (i) is actually (h).



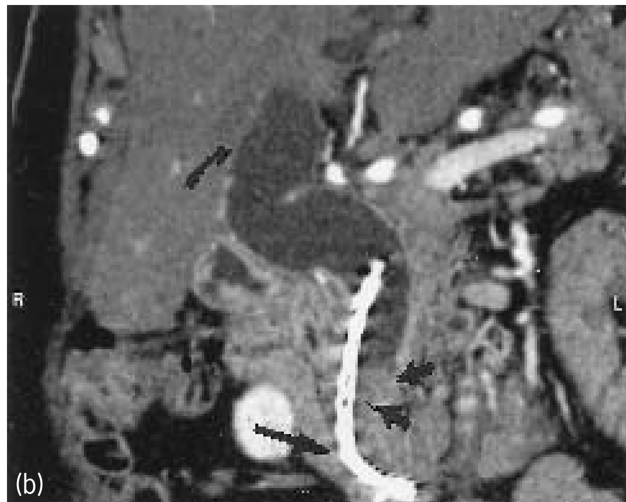


Figure 23 Cholangiocarcinoma in the distal common bile duct. (a) Axial scan through the thick-walled distal common bile duct (white arrows) in a patient with a stent (black arrow). The duct is distended with low-density bile (curved arrow). (b) Standard two-dimensional reconstruction shows the stent going through the dilated common bile duct (curved arrow) into the duodenum (long arrow). The tumour is well appreciated as an abrupt obstructing mass (short arrows).

Nephron-sparing surgery is the surgical treatment of choice for patients with renal cell carcinoma who have a solitary kidney, contralateral renal impairment or bilateral tumours. Its success rate is similar to that of radical nephrectomy in patients with a single tumour under 4 cm in size and with a normally functioning contralateral kidney. It is more technically challenging than radical nephrectomy and is benefitted by the diagnostic information and preoperative planning information provided by triphasic helical CT scanning and 3D volume rendering (Coll *et al*., 2000). Volume rendering techniques incorporate the entire CT data set into a 3D image (Calhoun *et al*., 1999). Preoperative information includes renal position relative to the lower rib cage, spine and iliac crest, tumour location and depth of extension into the kidney, relationship of the tumour to the collecting system and renal arterial and venous anatomy, including the segmental arterial supply to the tumour (**Figure 24**).

Therapy

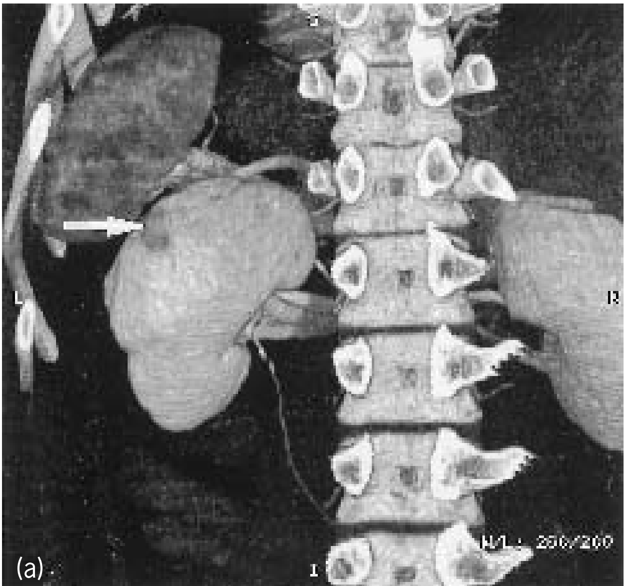
Local Nonsurgical Treatment of Disease

The art of radiation therapy entails the administration of a precise dose of ionizing radiation to a designated tissue volume. Treatments are often limited by the tolerance of surrounding normal tissue (Mohan *et al*., 1998). The role of CT includes depicting anatomical structures for better delineation of the tumour volume and healthy critical organs, allowing planning in a three-dimensional setting, permitting accurate placement of beams and field-shaping devices and allowing more accurate prediction of dose at various sites. CT-guided conformal radiation therapy permits accurate and consistent tumour localization immediately prior to treatment while sparing adjacent normal tissues. This permits the administration of higher radiation doses to tumours such as prostate, with reported improvements in local tumour control (55–94%) and reduction of complications in normal tissues from 10 to 20% (Zelefsky *et al*., 1988). Percutaneous ablation of both primary and metastatic tumours is currently being performed. Accurate preprocedural tumour localization, and immediate and long-term follow-up of its efficacy, are accomplished utilizing CT.

Surgical resection and orthotopic liver transplantation currently offer the best hope for survival for patients with hepatocellular carcinoma (HCC). The disease in unresectable in 80% of patients at the time of initial diagnosis (Okuda *et al*., 1985). Interventional procedures including local ethanol injection are currently being performed in patients who are unresectable or whose clinical status renders them inoperable (Livarhi *et al*., 1986). CT provides imaging guidance for the injection, immediate confirmation of the success of the procedure and long-term follow-up (**Figure 25**).

Palliation

CT may confirm the necessity and success of various palliative procedures. The placement of stents may be endobronchial when there is proximal bronchial obstruction sufficient to cause significant respiratory compromise, endovascular as in the case of superior vena cava obstruction or endoluminal in the presence of biliary obstruction (**Figure 23**).



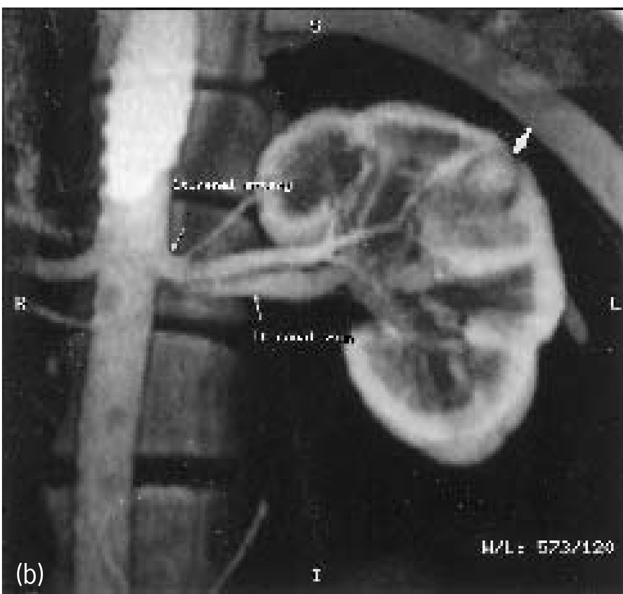


Figure 24 Renal-cell carcinoma. Role of 3D imaging in surgical planning for partial nephrectomy. (a) 3D reconstruction when viewed from behind demonstrates a 1-cm tumour on the surface of the left kidney (arrow). (b) Imaging during the arterial phase demonstrates the small tumour (arrow) and clearly defines the planes that can be used for a surgical approach. In addition, a single left renal artery and single left renal vein are identified and labelled. It is extremely important in doing partial nephrectomies to know the vascular anatomy. (c) 3D axial image viewed from above, which is a view commonly used in surgery, demonstrates the tumour (short arrow). The left renal artery (long arrow) is seen coming off the aorta with an anterior branch and the left renal vein (curved arrow) is readily seen anterior to it (the anterior is inferior and posterior is superior in the view). Identification of exact tumour location and vascular anatomy is crucial to the surgeon in planning a nephron-sparing, partial nephrectomy. The 'CT angiogram' obviates the need for a standard angiogram, which is a more invasive study.

Follow-up

Surveillance

Routine surveillance of oncological patients is frequently conducted utilizing CT. They are performed at appropriate intervals to monitor treatment efficacy (**Figure 26**). CT surveillance provides an opportunity for early detection and biopsy of a suspected lesion (**Figures 18** and **19**), which in turn can result in prompt institution of appropriate therapy. Local and potential distant sites of recurrence are scanned. CT findings indicative of unresponsiveness includes lack of tumour regression, local increase in size and/or number of lesions or adenopathy and the development of distant metastases (**Figure 27**).

The importance of CT in determining local recurrence is exemplified in patients who have undergone radiation for supraglottic or glottic carcinoma. Accurate interpretation of CT examinations in patients who undergo radiation requires that the expected changes due to treatment not be misinterpreted as residual or recurrent tumour. A baseline CT showing complete resolution at the primary site and symmetric–appearing laryngeal and hypopharyngeal tissues (i.e. expected radiation therapy-related changes) is predictive of permanent local control. However, if less than 50% estimated volume

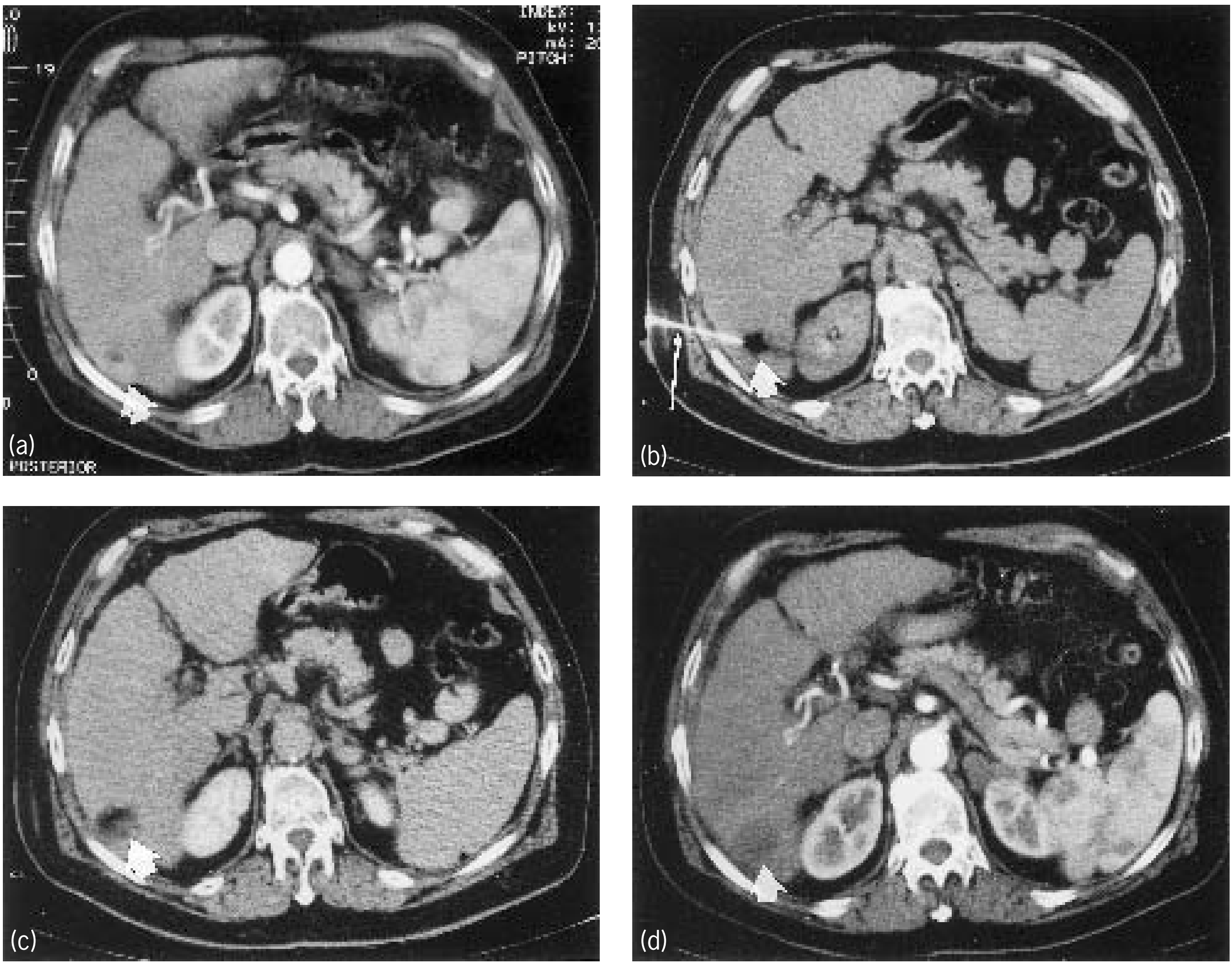


Figure 25 Ethanol ablation of a hepatoma utilizing CT. (a) Arterial phase CT reveals an enhancing mass in the posterior segment of the right hepatic lobe (arrow). Biopsy-confirmed hepatoma. (b) Axial scan performed during the CT-guided percutaneous ethanol injection procedure. Following confirmation of accurate needle placement (long arrow), ethanol is injected resulting in a defect (short arrow) in the mass. (c) Image obtained at completion of the procedure reveals an area of mixed attenuation in the mass (arrow). (d) Followup study. Arterial phase scan at the same level as the preprocedure CT (a) demonstrates the absence of early enhancement. Instead, only a relatively low attenuation area is seen (arrow). This disappearance of the hypervascularity denotes a successful ablation procedure.

reduction or a focal mass with a diameter $\geq$1 cm is found, immediate further investigation is warranted, as the likelihood of local failure is high (Hermans *et al*., 2000).

Even in situations where the earliest evidence of recurrence is local, the possibility of distant metastases must be considered and appropriate scans obtained. Recurrent cervical carcinoma typically occurs in the pelvis with masses or lymphadenopathy (**Figure 19**). However, with the introduction of intensive pelvic radiation therapy, there has been an increase in atypical manifestations of recurrence, including distant solid organ metastases (Fulcher *et al*., 1999). In addition, follow-up CT scans may provide early detection and diagnosis of a second primary cancer (Munden *et al*., 1997; Quint *et al*., 2000).

Disease progression may be clinically silent and detected on routine surveillance studies, or it may present with clinical findings requiring investigation. Dyspnea is a relatively common symptom in oncological patients. A high-resolution CT scan of the chest performed for dyspnea in a patient with breast cancer may reveal pulmonary lymphangitic carcinomatosis – haematogenous spread of tumour into the lungs, penetrating vessel walls and the surrounding interstitium and lymphatics (**Figure 28**) (Johkoh *et al*., 1992). Patients may present with an abnormal physical examination, such as a superficial mass (**Figure 29**), superior vena cava syndrome (**Figure 30**), or an acute change in mental status (**Figure 8**).

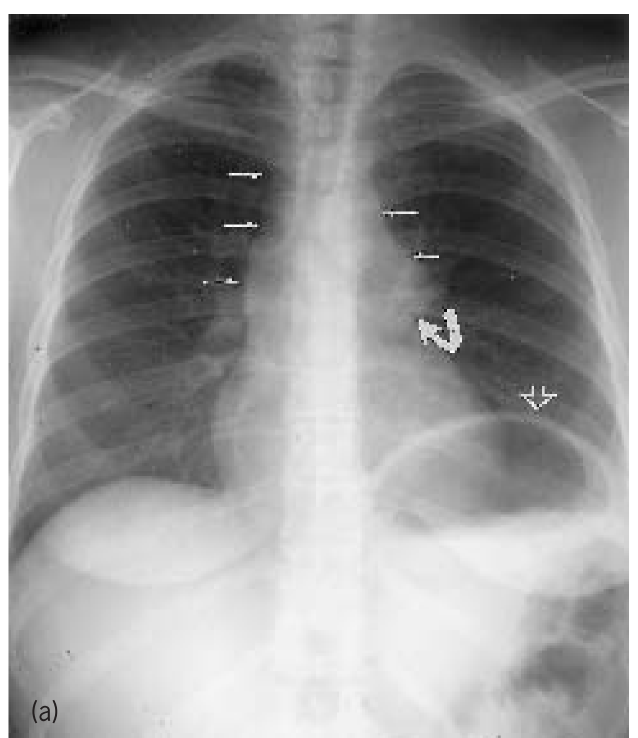
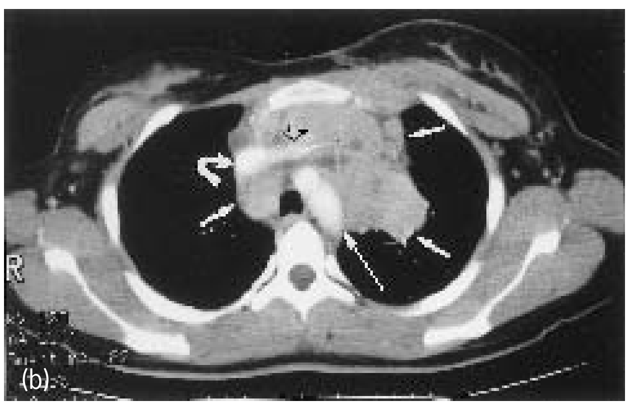
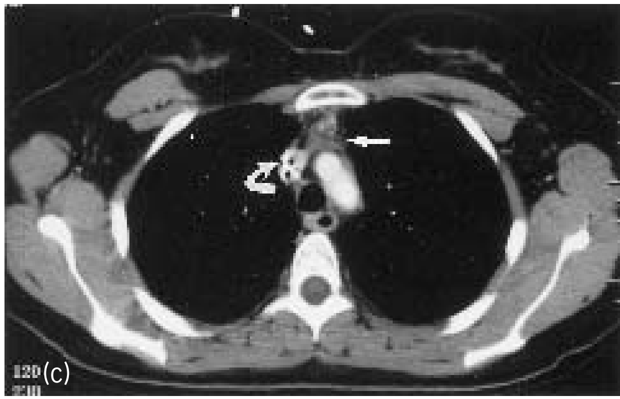
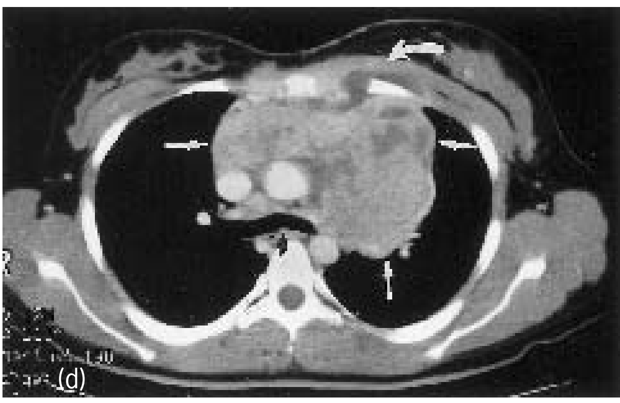
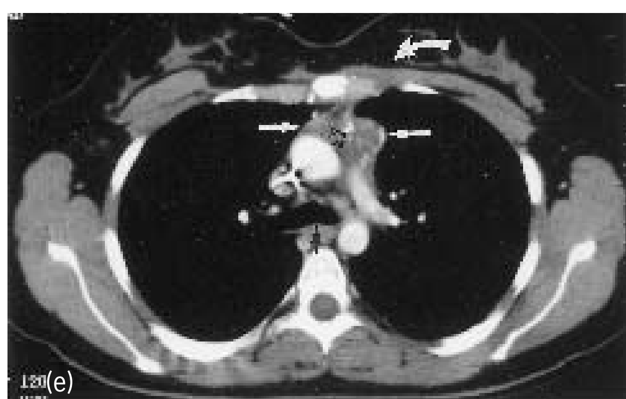


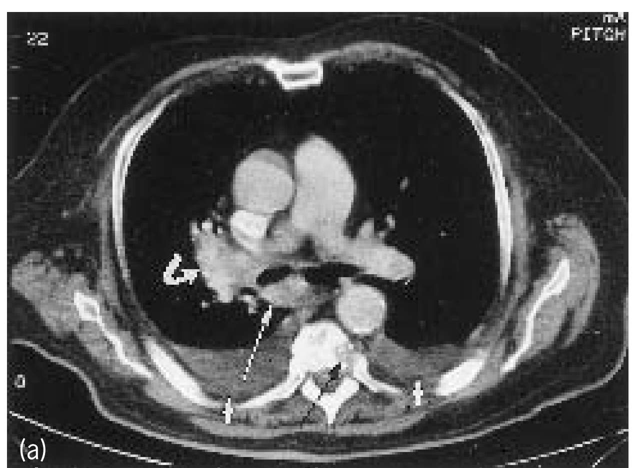
Figure 26 CT in staging non-Hodgkin lymphoma. (a) A chest radiograph in a patient with non-Hodgkin lymphoma demonstrates widening of the mediastinum (straight arrows), hilar enlargement (curved arrow) and elevation of the left diaphragm (open arrow). Pretreatment (b) and posttreatment (c) intravenous contrast-enhanced thoracic CT examinations. Initial scan at the level of the superior vena cava (SVC) (curved arrow) reveals extensive bilateral mediastinal lymphadenopathy (straight arrows). While the vessels including the aorta (long arrow) and brachiocephalic vein (open arrow) are encased they are not occluded, a feature which distinguishes lymphomas from most carcinomas of similar size. Posttreatment axial image (c) at a similar level reveals treatment response with marked diminution of the mediastinal mass (straight arrow). A central venous catheter is present in the SVC through which chemotherapy had been administered (curved arrow). (d), (e) Pretreatment (d) and posttreatment (e) axial images at the level of carina (black arrows). Prior to therapy (d) the extensive mediastinal lymphadenopathy (straight white arrows) encompasses the phrenic nerve accounting for the left diaphragmatic elevation. Chest wall extension (curved arrow) is present, a feature associated with non-Hodgkin rather than Hodgkin lymphoma. Following therapy, (e) calcifications (open arrow) are present in the residual adenopathy (white arrows). The chest wall involvement has regressed (curved arrow).

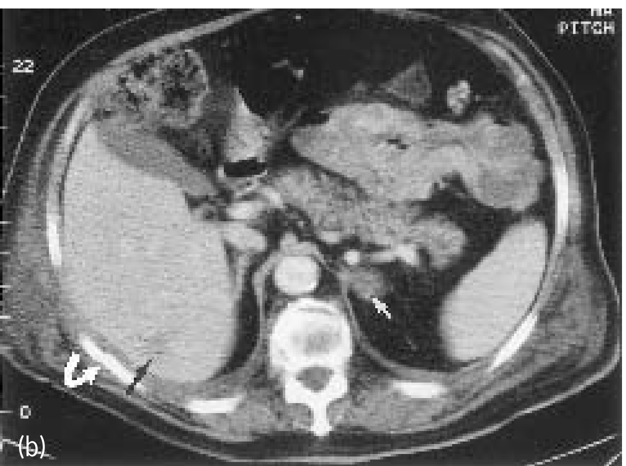












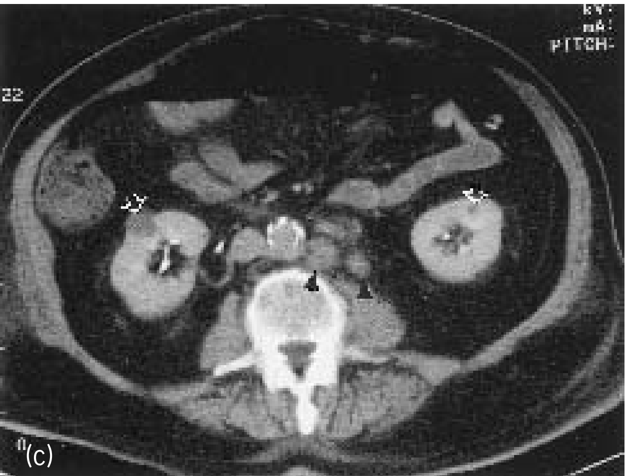


Figure 27 Unresponsive small-cell lung cancer. Multiple sites of neoplastic involvement are revealed during a contrast-enhanced thoracic and abdominal CT. (a) Axial image of the chest demonstrates adenopathy in the right hilum (curved arrow) and mediastinum (long white arrow). Bilateral pleural effusions (short arrows) and vertebral body destruction (long black arrow) are also present. (b) Axial image of the upper abdomen reveals a hepatic metastasis (black arrow), a left adrenal mass (white arrow) and a pathological right rib fracture (curved arrow). (c) Image caudal to (b) demonstrates bilateral renal masses (open arrows) and left para-aortic adenopathy (arrowheads).

Complications

The oncological treatments themselves have risks of complications. Postsurgical complications are often worked up with CT. Acute complications such as anastomotic leakage or abscess formation may appear as extraluminal collections on CT. Intestinal obstruction is usually a more chronic complication, especially in patients with a history of surgery for abdominal malignancy. CT is valuable for determining the presence or absence and the level and cause of bowel obstruction (Ha *et al*., 1998). The obstruction may be benign, secondary to adhesions or related to neoplastic disease. Recurrent tumour or peritoneal carcinomatosis causes malignant obstruction. CT findings not only confirm the diagnosis of intestinal obstruction but also are useful in management. The long-term survival rate in the presence of a malignant obstruction is poor, and the incidence of strangulated obstruction is rare, so early surgical intervention is not generally performed. On the other hand, up to one quarter of the patients have a benign cause, a resectable or easily bypassed local recurrence or a potentially curable new primary, making them candidates for early surgery should conservative management fail.

Complications secondary to nonsurgical treatment modalities are also amenable to CT scanning. CT is more sensitive and demonstrates radiation-induced lung parenchymal changes earlier than chest radiography (Park *et al*., 2000). When pulmonary areas of increased opacity are seen following radiation, the differential diagnosis includes radiation pneumonitis, local recurrence, lymphangitic tumour spread and infectious pneumonitis. The appearance of radiation pneumonitis is unique as it is generally confined to the radiation portals. CT is useful as it can depict not only the nature of the opacities (diffuse haziness and areas of consolidation in the acute, and fibrotic changes in the chronic phases), but also localize them. High-resolution chest CT is valuable in demonstrating findings suggestive of pulmonary toxicity secondary to chemotherapeutic agents in the appropriate clinical setting. In cases where the findings are nonspecific, HRCT is helpful in directing a biopsy to more abnormal areas thus increasing the yield of diagnostic material. CT may also assist in the workup of other more generalized complications such as infection and pulmonary emboli, the latter with and without deep venous thrombosis (Loud *et al*., 2000; Qanadli *et al*., 2000).

FUTURE DIRECTIONS

Advances in CT technology continue. Some involve the CT scanner itself such as the development of portable machines which can be placed in hospital elevators and moved to different patient care areas, or scanners with rotational capabilities to assist in conformal radiotherapy delivery. Faster scanners and improvements in software promise to further expand the imaging capabilities of CT. Multislice imaging utilizing a greater number of detectors will result in increased scanning speed, the creation of more scanning options and improved utilization of administered contrast.

Software advances are continuing to expand the imaging capabilities of CT. New virtual imaging techniques

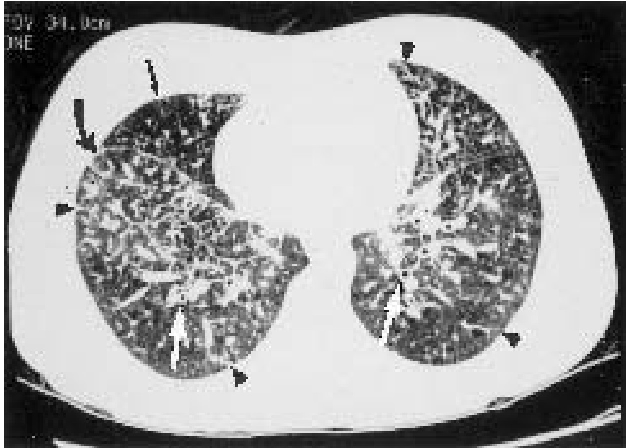


Figure 28 High-resolution CT (HRCT) of the lungs. A patient with known breast cancer developed dyspnea with no clinical evidence of infection. An axial image reveals bilateral patchy interstitial abnormalities including: thickening of peribronchial (white arrows), interlobular (arrowheads) and interlobar (curved arrow) interstitium. Note the normal appearing interlobular septa in the uninvolved right middle lobe (straight arrow). The HRCT technique employs increased milliamperage or kilovolt peak X-rays, thin collimation and an edge-enhanced reconstruction algorithm in order to display the pulmonary interstitium optimally.

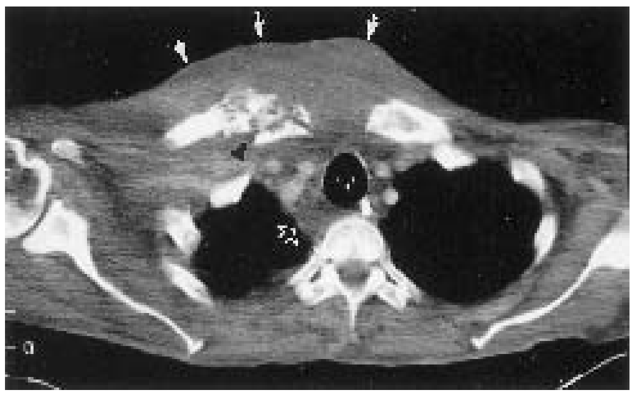
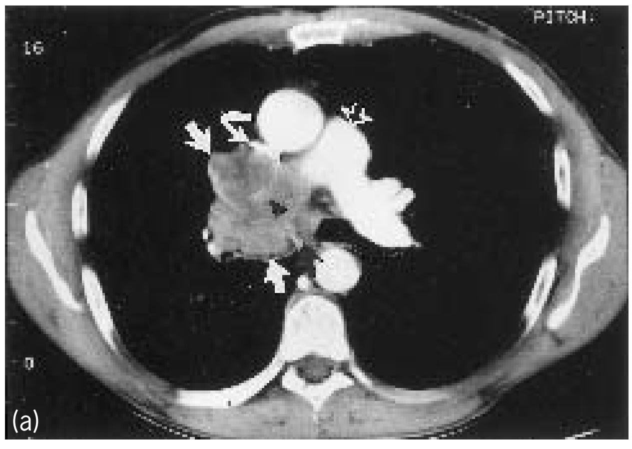


Figure 29 Thoracic CT image of a patient with a history of non-Hodgkin lymphoma who presented with a large visible and palpable mass which is shown to be dramatic chest wall extension (closed arrows) of the mediastinal adenopathy (open arrow), with destruction of the medial right clavicle (arrowhead).



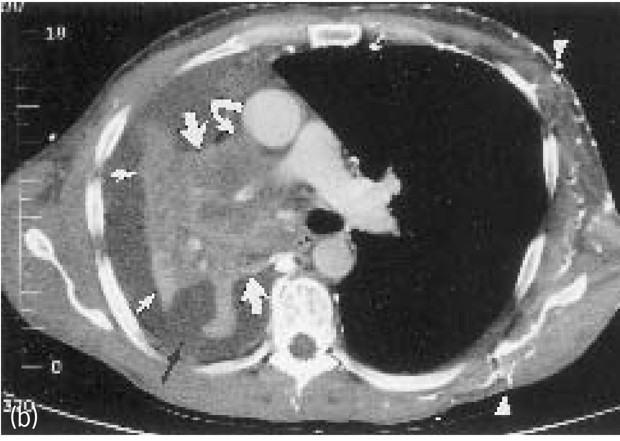


Figure 30 Progression of disease (lung cancer). (a) Axial CT scan of a patient with biopsy-proven adenosquamous carcinoma of the lung at the level of the main pulmonary artery (open arrow) reveals a large, heterogeneous mass (arrows). The barely visible right pulmonary artery (arrowhead) and absence of an air-filled right mainstem bronchus (black arrow) indicate their involvement. The SVC is compressed but patent (curved arrow). (b) Follow-up scan obtained 1 month later when the patient presented with SVC syndrome-engorgement of the veins of the upper chest, neck and face. Axial image at the same level reveals disease progression. The margins of the actual mass (large white arrows) are obscured owing to the development of right lung collapse (small white arrows). A contrast-filled SVC is no longer visualized owing to interval tumour growth (curved arrow). Numerous contrast-enhanced collateral vessels are now visualized throughout the left chest wall (arrowheads). Note the presence of a new pleural effusion (black arrow).

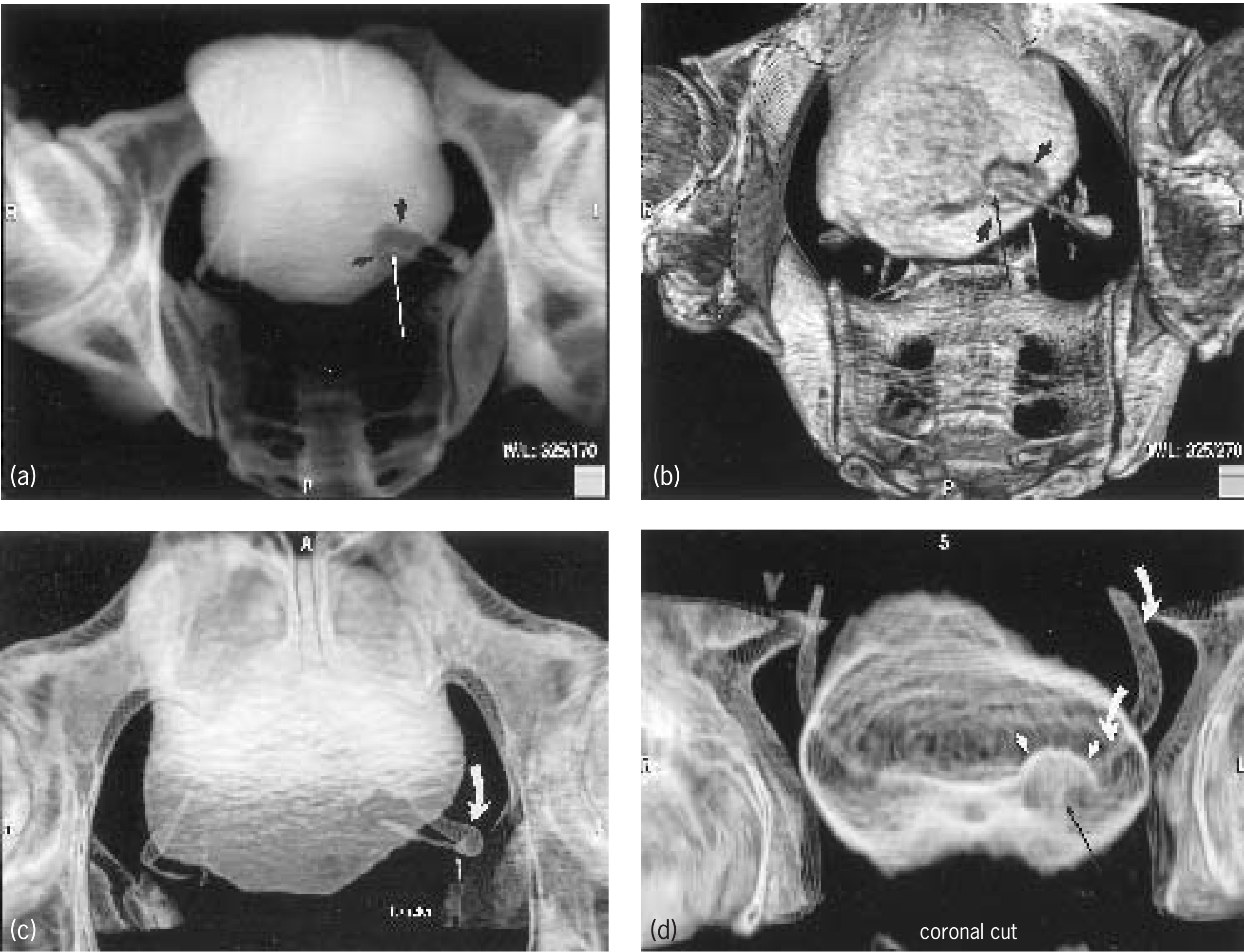


Figure 31 Bladder tumour with 'virtual' imaging. These volume rendered 3D models were reconstructed from axial CT images acquired as part of a CT/IVP examination in which a vesical abnormality was detected. Different techniques were employed to alter their transparencies. (a) 3D reconstruction in the axial plane as viewed from above demonstrates a tumour in the left hemitrigone (short arrows) where the ureter enters the bladder without obstruction to the ureter (long arrow). (b) The 3D reconstruction utilizing a shaded surface display further accentuates the tumour margins (short arrows). The distal ureter (long arrow) and its relationship to the tumour are again displayed. (c) The same 3D reconstruction rendered with a 'transparency technique' permits evaluation of the inner lumen of the left ureter (curved arrow). (d) 3D reconstruction in the coronal plane with further transparency application demonstrates not only the inner lumen of the distal left ureter (curved arrow), but also of the bladder. Note the mass in the hemitrigone (short arrows) and the distal ureter within it (long arrow). Postprocessing data manipulation can permit added information to be extracted from the image and can also be a great aid in guiding the surgeon to the correct approach.

are being evaluated. Three-dimensional helical CT of the tracheobronchial tree (Zeiberg *et al.*, 1996) and virtual bronchoscopy (Ferretti *et al.*, 2000) show promise in evaluating the central airway. Multiplanar CT pancreatography and distal cholangiography appear promising (Prassopoulos *et al.*, 1998; Raptopoulos *et al.*, 1998). Initial studies suggest that virtual cystoscopy maybe helpful in detection of bladder tumours larger than 5 mm (**Figure 31**) (Song *et al.*, 2001). Computer-assisted detection of lesions is also being developed (Summers *et al.*, 1998).

CT scans are being further integrated with complementary imaging modalities such as ultrasound (US) and magnetic resonance imaging (MRI). The correlation of a tumour's appearance and location on CT with its physiology (such as increased glucose uptake in a metabolically active lesion) as demonstrated by positron emission tomography (PET) has tremendous implications for oncological imaging.

REFERENCES

Abdel-Nabi, H., *et al*. (1998). Staging of primary colorectal carcinomas with fluorine-18 fluorodeoxyglucose whole body PET: correlation with histopathologic and CT findings. *Radiology*, **206**, 755–760.

Aronson, S., *et al*. (1991). Cystic renal masses: usefulness of the Bosniak classification. *Urologic Radiology*, **13**, 83–90.

Barnes, G. T., *et al*. (1979). Optimizing computed tomography (CT) scanner geometry. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, **173**, 225–237.

Baron, R. J., *et al*. (1996). Hepatocellular carcinoma: evaluation with biphasic contrast-enhanced, helical CT. *Radiology*, **199**, 505–511.

Birnbaum, B. A., *et al*. (1996). Multiphasic renal CT: comparison of renal mass enhancement during the corticomedullary and nephrographic phases. *Radiology*, **200**, 753–758.

Boiselle, P. M., *et al*. (1998). Imaging of mediastinal lymph nodes: CT, MR, and FDG PET. *RadioGraphics*, **18**, 1061–1969.

Boland, G. W. L., *et al*. (1997). Adrenal masses: characterization with delayed contrast-enhanced CT. *Radiology*, **202**, 693–696.

Boland, G. W. L., *et al*. (1998). Characterization of adrenal masses using unenhanced CT: an analysis of the CT literature. *American Journal of Roentgenology*, **171**, 201–204.

Bosniak, M. A. (1986). The current radiological approach to renal cysts. *Radiology*, **158**, 1–10.

Bosniak, M. A. (1997). Diagnosis and management of patients with complicated cystic lesions of the kidney. *American Journal of Roentgenology*, **169**, 819–821.

Brink, J. A., *et al*. (1994). Helical CT: principles and technical considerations. *RadioGraphics*, **14**, 887–893.

Calhoun, P. S., *et al*. (1999). Three-dimensional volume rendering of spiral CT data: theory and method. *RadioGraphics*, **19**, 745–764.

Chen, J., *et al*. (1999). Proximal arterioportal shunting associated with hepatocellular carcinoma: features revealed by dynamic helical CT. *American Journal of Roentgenology*, **172**, 403–407.

Coll, D. M., *et al*. (2000). Preoperative use of 3D volume rendering to demonstrate renal tumors and renal anatomy. *RadioGraphics*, **20**, 431–438.

Daly, B., *et al*. (1999). Percutaneous abdominal and pelvic interventional procedures using fluoroscopy guidance. *American Journal of Roentgenology*, **173**, 637–644.

Dennie, C. J., *et al*. (2001). Transthoracic needle biopsy of the lung: results of early discharge in 506 outpatients. *Radiology*, **219**, 247–251.

Fenlon, H. M., *et al*. (1999). Occlusive colon carcinoma: virtual colonoscopy in the preoperative evaluation of the proximal colon. *Radiology*, **210**, 423–428.

Ferretti, G. R., *et al*. (2000). Benign abnormalities and carcinoid tumors of the central airways: diagnostic impact of CT bronchography. *American Journal of Roentgenology*, **174**, 1307–1313.

Fox, S. H., *et al*. (1998). Future directions in CT technology. *Neuroimaging Clinics of North America*, **8**, 497–513.

Fulcher, A. S., *et al*. (1999). Recurrent cervical carcinoma: typical and atypical manifestations. *RadioGraphics*, **19**, S103–S116.

Ha, H. K., *et al*. (1998). Usefulness of CT in patients with intestinal obstruction who have undergone abdominal surgery for malignancy. *American Journal of Roentgenology*, **171**, 1587–1593.

Henschke, C. I., *et al*. (1999). Early Lung Detection Action Project: overall design and findings from baseline screening. *Lancet*, **354**, 99–105.

Hermans, R., *et al*. (2000). Laryngeal or hypopharyngeal squamous cell carcinoma: can follow-up CT after definitive radiation therapy be used to detect local failure earlier than clinical examination alone? *Radiology*, **214**, 683–687.

Hollett, M. D., *et al*. (1995). Dual-phase helical CT of the liver: the value of arterial scans in the detection of small (<1.5 cm) malignant hepatic neoplasms. *American Journal of Roentgenology*, **164**, 879–884.

Horton, K. M., *et al*. (2000). Spiral CT of colon cancer: imaging features and role in management. *RadioGraphics*, **20**, 419–430.

Hounsefield, G. N. (1973). Computerized transverse axial scanning (tomography): Part I. Description of system. *British Journal of Radiology*, **46**, 1016–1022.

Hu, H., *et al*. (2000). Four Multidetector-row helical CT: image quality and volume coverage speed. *Radiology*, **215**, 55–62.

Ibukkuro, K., *et al*. (1995). Helical CT angiography with multiplanar reformation: techniques and clinical applications. *RadioGraphics*, **15**, 671–682.

Inaba, Y., *et al*. (2000). Revealing hepatic metastases from colorectal cancer: value of combined helical CT during arterial portography and CT hepatic arteriography with a unified CT and angiography system. *American Journal of Roentgenology*, **174**, 955–961.

Itoh, S., *et al*. (2000). Lung cancer screening: minimum tube current required for helical CT. *Radiology*, **215**, 175–183.

Johkoh, T., *et al*. (1992). CT findings in lymphangitic carcinomatosis of the lung: correlation with histologic findings and pulmonary function tests. *American Journal of Roentgenology*, **158**, 1217–1222.

Kalendar, W. A., *et al*. (1990). Spiral volumetric CT with single-breath-hold technique, continuous transport, and continuous scanner rotation. *Radiology*, **176**, 181–183.

Koike, T., *et al*. (1999). The influence of lung cancer mass screening on surgical results. *Lung Cancer*, **24**, 75–80.

Korobkin, M., *et al*. (1996). Delayed enhanced CT for differentiation of benign from malignant adrenal masses. *Radiology*, **200**, 737–742.

Korobkin, M., *et al*. (1998). CT time-attenuation washout curves of adrenal adenomas and nonadenomas. *American Journal of Roentgenology*, **170**, 747–752.

Landis, S. H., *et al*. (1998). Cancer statistics 1998. *CA Cancer Journal for Clinicians*, **48**, 6–29.

Livarhi, T., *et al*. (1986). US-guided percutaneous alcohol injection of small hepatic and abdominal tumors. *Radiology*, **161**, 309–312.

Loud, P. A., *et al*. (2000). Combined CT venography and pulmonary angiography in suspected thromboembolic disease: diagnostic accuracy for deep venous evaluation. *American Journal of Roentgenology*, **174**, 61–65.

Macari, M., *et al*. (1998). Non-small cell lung carcinoma: usefulness of unenhanced helical CT of the adrenal glands in an unmonitored environment. *Radiology*, **209**, 807–812.

Macari, M., *et al*. (1999). Usefulness of CT colonography in patients with incomplete colonoscopy. *American Journal Roentgenology*, **173**, 561–564.

McLoud, T. C., *et al*. (1992). Bronchogenic carcinoma: analysis of staging in the mediastinum with CT by correlative lymph node mapping and sampling. *Radiology*, **182**, 319–323.

Miller, F. H., *et al*. (1988). Using triphasic helical CT to detect focal hepatic lesions in patients with neoplasms. *American Journal of Roentgenology*, **171**, 643–649.

Mohan, R., *et al*. (1998). A comprehensive three-dimensional radiation treatment planning system. *International Journal of Radiation Oncology Biological Physics*, **15**, 481–495.

Morrin, M. M., *et al*. (1999). Endoluminal CT colonography after an incomplete endoscopic colonoscopy. *American Journal Roentgenology*, **172**, 913–918.

Mountain, C. F. (1997). Revisions in the international system for staging lung cancer. *Chest*, **111**, 1710–1117.

Mountain, C. F. and Dresler, C. M. (1997). Regional lymph node classification for lung cancer staging. *Chest*, **111**, 1718–1723.

Munden, R. F., *et al*. (1997). Small pulmonary lesions detected at CT: clinical importance. *Radiology*, **202**, 105–110.

Murakami, T., *et al*. (2001). Hypervascular hepatocellular carcinoma: detection with double arterial phase multi-detector row helical CT. *Radiology*, **218**, 763–767.

Nishiharu, T., *et al*. (1999). Local extension of pancreatic carcinoma: assessment with thin-section helical CT versus with breath-hold fast MR imaging-ROC analysis. *Radiology*, **212**, 445–452.

Novick, S. L. and Fishman, E. K. (1998). Three-dimensional CT angiography of pancreatic carcinoma: role in staging extent of disease. *American Journal of Roentgenology*, **170**, 139–143.

Okuda, K., *et al*. (1985). Natural history of hepatocellular carcinoma and prognosis in relation to treatment: study of 850 patients. *Cancer*, **56**, 918–928.

O'Malley, M. E., *et al*. (1999). Adenocarcinoma of the head of the pancreas: determination of surgical unresectability with thin-section pancreatic-phase helical CT. *American Journal of Roentgenology*, **173**, 1513–1518.

Outwater, E. K., *et al*. (1996). Adrenal masses: correlation between CT attenuation value and chemical-shift MR imaging. *Radiology*, **200**, 749–752.

Park, J. P., *et al*. (2000). Radiation-induced lung disease and the impact of radiation methods on imaging features. *Radio-Graphics*, **20**, 83–98.

Patz, E. F. Jr, *et al*. (1999). Lung cancer staging and management: comparison of contrast-enhanced and nonenhanced helical CT of the thorax. *Radiology*, **212**, 56–60.

Prassopoulos, P., *et al*. (1998). Development of virtual CT cholangiopancreatoscopy. *Radiology*, **209**, 570–574.

Qanadli, S. D., *et al*. (2000). Pulmonary embolism detection: prospective evaluation of dual-section helical CT versus selective pulmonary arteriography in 157 patients. *Radiology*, **217**, 447–455.

Quint, L. E., *et al*. (2000). Solitary pulmonary nodules in patients with extrapulmonary neoplasms. *Radiology*, **217**, 257–261.

Raptopoulos, V., *et al*. (1998). Multiplanar CT pancreatography and distal cholangiography with minimum intensity projections. *Radiology*, **207**, 317–324.

Read, T. E., *et al*. (1997). Importance of adenomas 5 mm or less in diameter that are detected by sigmoidoscopy. *New England Journal of Medicine*, **336**, 8–12.

Sheridan, M. B., *et al*. (1999). Dynamic contrast-enhanced MR imaging and dual-phase helical CT in the preoperative assessment of suspected pancreatic cancer: a comparative study with receiver operating characteristic analysis. *American Journal of Roentgenology*, **173**, 583–590.

Som, P. M., *et al*. (1999). Surveillance CT and the prompt use of CT-guided fine-needle aspiration in patients with head and neck cancer who have undergone surgery. *American Journal of Roentgenology*, **173**, 1505–1508.

Song, J. H., *et al*. (2001). Bladder tumor detection at virtual cystoscopy. *Radiology*, **218**, 95–100.

Summers, R. M., *et al*. (1998). Polypoid lesions of airways: early experience with computer-assisted detection by using virtual bronchoscopy and surface curvature. *Radiology*, **208**, 331–337.

Sundaram, M., *et al*. (1982). Utility of CT-guided abdominal aspiration procedures. *American Journal of Roentgenology*, **139**, 1111–1115.

Swensen, S. J., *et al*. (2000). Lung nodule enhancement at CT: multicenter study. *Radiology*, **214**, 73–80.

Szolar, D. H., *et al*. (1997). Multiphasic helical CT of the kidney: increased conspicuity for detection and characterization of small (<3 cm) renal masses. *Radiology*, **202**, 211–217.

Ter-Pogossian, M. M. (1977). Computerized cranial tomography: equipment and physics. *Seminars in Roentgenology*, **12**, 13–25.

Warshauer, D. M., *et al*. (1988). Detection of renal masses: sensitivities and specificities of excretory urography/liner tomography, US, and CT. *Radiology*, **169**, 363–365.

Weg, N., *et al*. (1998). Liver lesions: improved detection with dual-detector-array CT and routine 2.6-mm thin collimation. *Radiology*, **209**, 417–426.

Wescott, J. L., *et al*. (1997). Transthoracic needle biopsy of small pulmonary nodules. *Radiology*, **202**, 97–103.

Winawer, S. J., *et al*. (1993). Prevention of colorectal cancer by colonscopic polypectomy. *New England Journal of Medicine*, **329**, 1977–1981.

Yamashita, K., *et al*. (1995). Solitary pulmonary nodule: preliminary study of evaluation with incremental dynamic CT. *Radiology*, **194**, 399–405.

Yankelevitz, D. F., *et al*. (2000). Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology*, **217**, 251–256.

Yuh, B. I. and Cohan, R. H. (1999). Different phases of renal enhancement: role in detecting and characterizing renal masses during helical CT. *American Journal of Roentgenology*, **173**, 747–755.

Zeiberg, A. S., *et al*. (1996). Helical (spiral) CT of the upper airway with three-dimensional imaging: technique and clinical assessment. *American Journal of Roentgenology*, **166**, 293–299.

Zelefsky, M. J., *et al*. (1988). Locally advanced prostatic cancer: long-term toxicity outcome after three-dimensional conformal radiation therapy – a dose-escalation study. *Radiology*, **209**, 169–174.

Zerhouni, E. A., *et al*. (1996). CT and MR imaging in the staging of colorectal carcinoma: report of the Radiology Diagnostic Oncology Group II. *Radiology*, **200**, 443–451.

Zhang, M. and Kono, M. (1997). Solitary pulmonary nodules: evaluation of blood flow patterns with dynamic CT. *Radiology*, **205**, 471–478.

FURTHER READING

Bluemke, D. A., *et al*. (1995). Potentially resectable pancreatic adenocarcinoma: spiral CT assessment with surgical and pathologic correlation. *Radiology*, **197**, 381–385.

Brink, J. A. (1995). Technical aspects of helical (spiral) CT. *Radiologic Clinics of North America*, **33**, 825–841.

Buetow, P. C., *et al*. (1995). Colorectal adenocarcinoma: radiologic–pathologic correlation. *RadioGraphics*, **15**, 127–146.

Cirillo, R. L., Jr, *et al*. (1998). Pathology of the adrenal glands: imaging features. *American Journal of Roentgenology*, **170**, 429–435.

Curry, T. S., III, *et al*. (1990). *Christensen's Physics of Diagnostic Radiology*, 4th edn (Lea and Febiger, Philadelphia).

Daly, B. and Templeton, P. A. (1999). Real-time CT fluoroscopy: evolution of an interventional tool. *Radiology*, **211**, 309–315.

Erasmus, J. J., *et al*. (2000). Solitary pulmonary nodules: part I. Morphologic evaluation for differentiation of benign and malignant lesions. *RadioGraphics*, **20**, 43–58.

Erasmus, J. J., *et al*. (2000). Solitary pulmonary nodules: part II. Evaluation of the indeterminate nodule. *RadioGraphics*, **20**, 59–66.

Eubank, W. B., *et al*. (1998). Imaging of oncologic patients: benefits of combined CT and FDG PET in the diagnosis of malignancy. *American Journal of Roentgenology*, **171**, 1103–1110.

Foley, W. D. and Oneson, S. R. (1994). Helical CT: clinical performance and imaging strategies. *RadioGraphics*, **14**, 894–904.

Frush, D. P. and Donnelly, L. F. Helical CT in children: technical considerations and body applications.

Hara, A. K., *et al*. (1997). Colorectal lesions: evaluation with CT colography. *RadioGraphics*, **17**, 1157–1167.

Heiken, J. P., *et al*. (1993). Spiral (helical) CT. *Radiology*, **189**, 647–653.

Hendee, W. R. (1983). *The Physical Principles of Computed Tomography* (Little Brown, Boston).

Hopper, K. D., *et al*. (2000). Body CT and oncologic imaging. *Radiology*, **215**, 27–40.

Ko, J. P., *et al*. (2000). CT depiction of regional nodal stations for lung cancer staging. *American Journal of Roentgenology*, **174**, 775–782.

Marshall, C. (ed.) (1982). *The Physical Basis of Computed Tomography* (Warren H. Green, St. Louis, MO).

Rubin, G. D. (1994). Three-dimensional helical CT angiography. *RadioGraphics*, **14**, 905–912.

Sprawls, P. (1992). CT image detail and noise. *RadioGraphics*, **12**, 1041–1046.

Valls, C., *et al*. (1999). Hyperenhancing focal liver lesions: differential diagnosis with helical CT. *American Journal of Roentgenology*, **173**, 605–611.

Wong, K., *et al*. (2001). Breath-hold three-dimensional CT of the liver with multi-detector row helical CT. *Radiology*, **219**, 75–79.

Zeman, R. K., *et al*. (1998). Helical body CT: evolution of scanning protocols. *American Journal of Roentgenology*, **170**, 1427–1438.

Chapter 77

Ultrasound

Lucy E. Hann
Memorial Sloan–Kettering Cancer Center, New York, NY, USA

CONTENTS

- Introduction
- General Principles of Ultrasound as Applied to Cancer Imaging
- Doppler Ultrasound
- Ultrasound Contrast Agents
- Specific Clinical Applications of Ultrasound for Cancer Diagnosis
- New Advances in Ultrasound Technology
- Conclusion

INTRODUCTION

Ultrasound is a widely available cross-sectional imaging technique that is used for both diagnosis of cancer and image-directed therapy. This chapter will discuss general principles of ultrasound imaging, including Doppler blood flow analysis and how these techniques are applied for cancer detection. Potential applications of intraoperative ultrasound and new ultrasound techniques such as laparoscopic ultrasound and contrast imaging will also be addressed. Ultrasound-directed biopsy and therapy are discussed elsewhere.

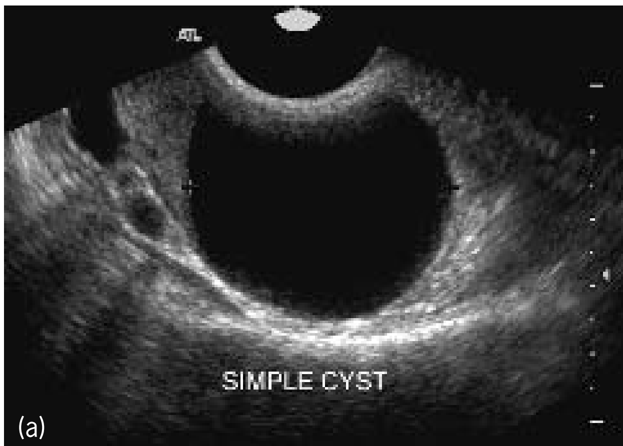
GENERAL PRINCIPLES OF ULTRASOUND AS APPLIED TO CANCER IMAGING

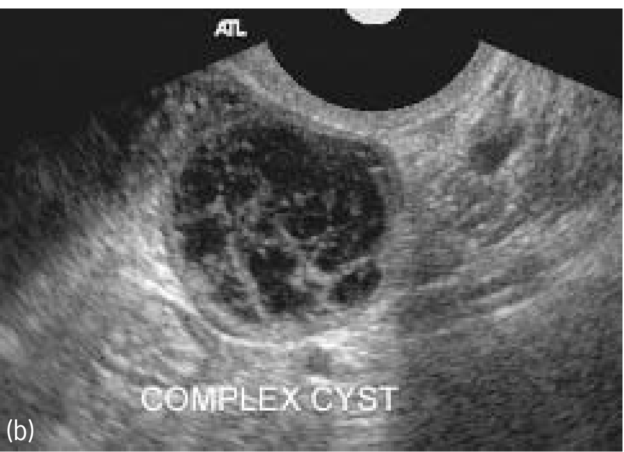
Ultrasound imaging is based on the transmission of high-frequency sound through tissue by using a hand-held transducer that both transmits and receives sound. The varying attenuation of sound by different tissue interfaces is used to generate an image. As sound travels through tissue, echoes are generated when some sound is reflected back to the transducer at each tissue interface. The distance of the interface or object is determined by the elapsed time between transmitted and received echo signals. The terms used to describe findings on ultrasound images relate to acoustic properties such as echogenicity, reflection and acoustic transmission.

In order to detect malignancies it is necessary to have sufficient soft tissue contrast to differentiate normal tissue from malignant masses and adequate spatial resolution to detect small masses or normal anatomical structures. Ultrasound has excellent soft tissue contrast and is the best imaging modality for differentiation of cystic, fluid containing masses from solid lesions (**Figure 1**). Fluid-containing masses have no internal echoes, and since sound is not attenuated by fluid, there is acoustic enhancement posterior to fluid-containing structures. Solid masses have internal echoes on ultrasound and are readily differentiated from fluid-containing cystic lesions. Ultrasound soft tissue contrast is used to characterize the internal architecture of lesions as a means of differentiating benign from malignant disease.

Soft tissue contrast provided by ultrasound has been improved further by the new technique of harmonic imaging. Harmonics are weak ultrasound signals that are produced when the incoming signal wavelength is altered by varying speed of sound transmission within tissue. The harmonic frequencies produced within tissue are multiples of the original signal frequency. Since harmonics are generated within the tissues there is less image background noise from scatter. Harmonic imaging, now routinely available, improves the depiction of fluid-containing structures, and better defines internal architecture of masses. In cancer patients, the information from harmonic ultrasound is particularly useful for the detection of small hepatic masses or for the identification of hepatocellular carcinomas that have similar echogenicity to normal liver (Hann *et al.*, 1999). Harmonic ultrasound is reported to provide additional information for liver imaging compared with standard ultrasound in approximately 30% of cancer patients because improved soft tissue contrast allows identification of smaller and more subtle lesions.

Ultrasound has excellent spatial resolution, in the range of 1 mm with standard transducers. New high-frequency transducers have sub-millimetre resolution. Normal anatomy such as extrahepatic bile ducts or small structures such as mural nodules within masses are readily resolved by sonography. The spatial resolution of ultrasound is





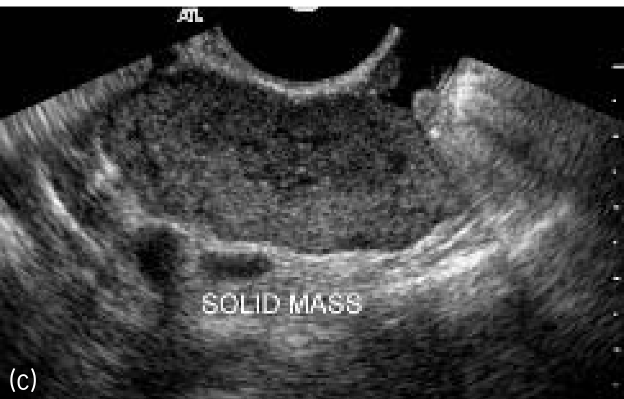
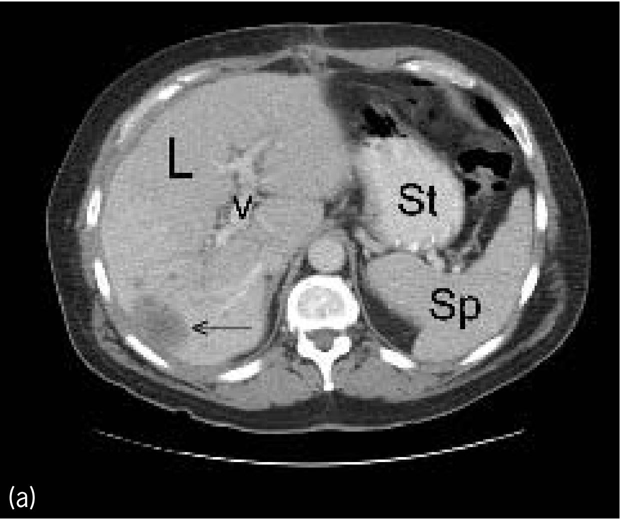


Figure 1 Ovarian masses in three different patients illustrate the use of ultrasound to differentiate between fluid and solid masses. (a) Simple benign cyst has no internal echoes (black area) and has a thin, smooth, white echogenic wall. (b) Complex ovarian mass has fluid (black areas) and solid components (white areas) due to multiple septations within the mass. With increasing complexity, there is increased likelihood of malignancy. (c) Solid ovarian carcinoma has internal echoes throughout the lesion.

directly related to transducer frequency. Higher frequency transducers have better resolution, but tissue penetration is limited because high-frequency signals are attenuated faster. In general, the highest frequency transducer that provides adequate image depth is chosen for examinations. For examination of superficial structures such as breast, thyroid and scrotum, high-frequency transducers in the range 7.5–15 MHz are used, but for abdominal imaging lower frequency transducers in the range 3–6 MHz are required for greater sound penetration to image deep organs.

Compared with computed tomography, ultrasound has comparable spatial resolution and better contrast resolution. There are other important differences in the two imaging techniques. Ultrasound images are generated manually and the quality of the study is directly related to the experience of the ultrasonographer. Ultrasound is therefore more operator dependent than computed tomography or magnetic resonance imaging. Another difference between ultrasound and computed tomography relates to the image field of view. Computed tomography provides a cross-sectional view of an entire anatomical region, but ultrasound images show cross-sectional anatomy in limited projections determined by acoustic windows that avoid artifact-producing bone and air. Because the acoustic properties of bone and air are so markedly different from soft tissue, at bone and air interfaces most of the transmitted sound is reflected back to the transducer and structures deep to bone and air are obscured by an acoustic shadow. For this reason, ultrasound images are generated in such a way as to avoid regions of bone- or gas-containing structures such as lung or gastrointestinal tract (**Figure 2**). Superficial organs such as thyroid, scrotum and breast are readily accessed by ultrasound. In the abdomen, liver, gallbladder, bile ducts, spleen and kidneys are well visualized sonographically, but the retroperitoneum may be obscured by bowel gas. To overcome potential artifacts from bowel gas, pelvic sonography is performed by displacing bowel from the pelvis with a filled urinary bladder or by direct transvaginal ultrasound that places the transducer close to the ovaries and uterus without interposed bowel.

Although ultrasound may give a limited field of view, it does provide multiplanar imaging similar to computerized tomography and magnetic resonance imaging. Ultrasound has an advantage in that the images are obtained in real time so there is no motion degradation and subtle changes during respiration or the cardiac cycle may be used to aid diagnosis. For example, ultrasound real-time imaging can identify changes in blood vessel calibre during respiration to differentiate between compression by tumour or direct tumour invasion of a vessel.



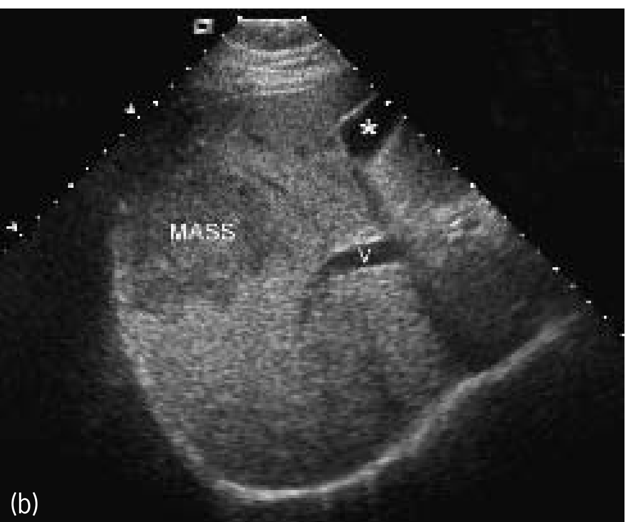


Figure 2 Comparison of computed tomography and ultrasound images in a patient with hepatic metastasis. (a) Computed tomographic axial image reveals the entire liver including the posterior metastasis (arrow). The stomach and spleen are also shown. (b) Ultrasound transverse image reveals the internal architecture of the metastasis (MASS, arrow heads). A portion of the portal vein (V) and gallbladder (asterisk) are included in the image, but since ultrasound is limited by air, the field of view is smaller than for computed tomography.

An encased blood vessel remains constricted, but a compressed vessel lumen changes during quiet respiration. Also, observations may be made regarding tissue compliance. Soft masses may be deformed by variations in applied pressure by the examiner at the skin surface but firm tumours remain rigid.

DOPPLER ULTRASOUND

Doppler imaging is unique to ultrasound and provides information regarding changes in blood flow during the cardiac cycle, direction of blood flow and patency of vessels. Assessment of blood flow characteristics is useful since tumours require angiogenesis for growth and since malignant tumours locally encase or obstruct blood vessels. The Doppler information can be displayed in three ways. Duplex Doppler samples flow within a vessel and displays velocity of flow over time (both mean and peak) as well as direction of flow relative to the position of the transducer. Some tumours such as hepatocellular carcinomas are characterized by high peak systolic velocity flow, a useful sign for diagnosis. This technique also readily identifies flow reversal within a vessel, as seen in the portal vein when portal venous hypertension is present. Power Doppler reveals amplitude of blood flow, but not flow direction. It is sensitive for slower flow and is the best technique for displaying the distribution of blood vessels within a mass. These Doppler techniques may be obtained individually or displayed simultaneously and often a combination of Doppler techniques is used. (**Figure 3; see colour plate section**).

Blood vessels in malignant tumours often lack a normal branching pattern and have a tortuous angulated course with areas of displacement by mass effect from tumour nodules (Tschammler *et al*., 1998). Tumour vessels lack muscular layers, so the Doppler waveform is often altered with low-resistance, high-diastolic flow that may be detected on duplex Doppler imaging. Doppler information is particularly useful for determining tumour involvement of vessels and can detect both intraluminal thrombi and occlusion. The pattern of vascular involvement by tumours may determine therapy. For example, liver tumours, regardless of size, that involve hepatic veins and portal veins may not be amenable to curative surgical resection, whereas large hepatic tumours without vascular involvement may be successfully resected.

ULTRASOUND CONTRAST AGENTS

Ultrasound intravenous contrast agents have recently been developed to improve the detection and characterization of tumour perfusion. These agents are highly reflective microbubbles that are encapsulated within a shell for stability. The microbubbles are less than 10 μm in diameter so that they can clear the pulmonary capillaries and remain in circulation. Contrast enhancement is produced by the interaction of the microbubbles and the ultrasound frequency. When micro bubbles are struck by an ultrasound signal, they alternately expand and contract, emitting harmonic frequencies that are multiples of the original transmitted ultrasound frequency. The ultrasound machine can selectively receive these harmonic

signals to show the distribution of arteries and veins. In addition, as the contrast agent recirculates through tissue capillaries, parenchymal enhancement also occurs (**Figure 4; see colour plate section**).

Ultrasound contrast agents have several potential applications for cancer diagnosis and treatment. For example, malignant liver tumours that are primarily supplied by the hepatic artery, such as hepatocellular carcinoma and vascular metastases, show hypervascularity with rapid arterial enhancement (Kim *et al.*, 1998). This pattern of enhancement differs from that of benign haemangiomas that exhibit gradual enhancement, filling from peripheral to central regions. During the parenchymal phase of contrast enhancement, hepatic tumours appear more conspicuous as dark masses against the contrast enhanced normal liver (Wilson *et al.*, 2000) (**Figure 5**). Improved detection of renal, breast and prostate tumours by contrast-enhanced ultrasound has also been reported. Some ultrasound contrast agents have a different mode of action: as reticuloendothelial agents, they are picked up by normal liver cells, but not by tumours.

Ultrasound contrast agents may also be used to monitor the response of tumours to treatment. Hepatic tumours are often treated with radiofrequency probes or alcohol injections that are placed directly into the mass. Both methods coagulate the tumour and destroy tumour vascularity. Ultrasound contrast agents are reported to be highly sensitive for the detection of any residual vascularity and are thus useful to assess adequacy of treatment (Koito *et al.*, 2000).

SPECIFIC CLINICAL APPLICATIONS OF ULTRASOUND FOR CANCER DIAGNOSIS

Ultrasound is generally the initial imaging study for cancer detection since it is readily available and relatively inexpensive. Ultrasound may detect abnormalities, but differentiation of benign from malignant disease often is not possible without biopsy. However, masses in the abdomen or pelvis as well as superficial structures are easily accessed by ultrasound-guided biopsy that can provide a definitive tissue diagnosis. Because ultrasound field of view is limited by bone and air, ultrasound is not suited for cancer staging that requires accurate visualization of all anatomical regions. Computed tomography is the preferred imaging technique for tumour staging.

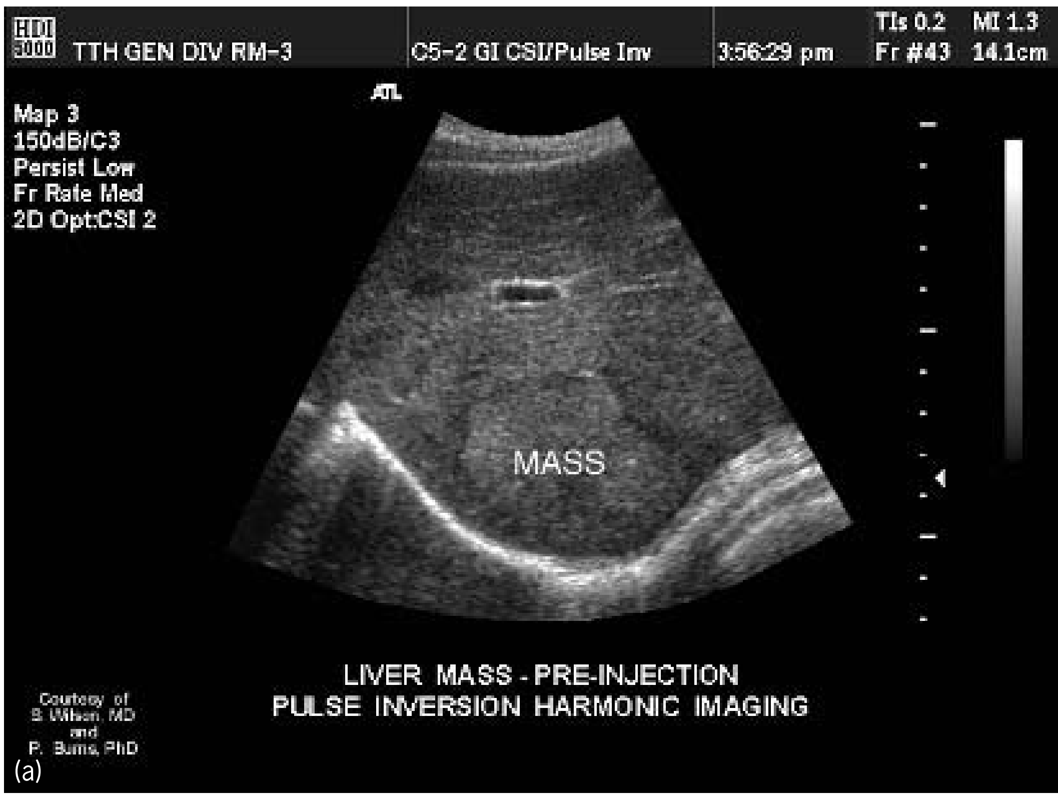
Breast

Breast cancer screening is performed by mammography but when a mass is evident by mammography or physical examination, ultrasound evaluation may provide additional useful information by differentiating benign cysts from solid masses. It is essential that the ultrasound examination be interpreted with mammographic correlation to ensure that ultrasound findings anatomically correspond in location to mammographic abnormalities. Benign simple cysts have anechoic interiors, sharp margins, a thin wall and posterior enhancement. Breast cancers are solid hypoechoic masses that appear dark compared with normal breast tissue on ultrasound images. Malignant infiltrating ductal carcinomas are often irregular in contour, taller than wide, and may have posterior acoustic shadowing because of the fibrous response associated with breast cancer growth (Stavros *et al.*, 1995). Microcalcifications, commonly present in ductal carcinomas, are best evaluated mammographically, but also may be seen on ultrasound images performed with very high-frequency transducers. Infiltrating lobular carcinoma characteristically has irregular ill-defined margins on ultrasound with distortion of adjacent breast parenchyma. In contrast to the pattern seen in carcinomas, benign fibroadenomas are smooth in contour, have an oblong shape, and may have a thin echogenic halo. The shape and border of a mass are the most useful predictors of malignancy for solid masses, but there is considerable overlap in ultrasound appearance. Noninvasive identification of vascular changes and estimates of blood vessel density by Doppler ultrasound are also being explored for breast cancer detection since blood vessels comprise a significant portion of tumour volume and tumours have both altered flow states and vascular distribution.

For breast masses that are indeterminate for malignancy by mammography or ultrasound, biopsy is required for definitive diagnosis. Percutaneous image-guided core biopsy of the breast has become standard as a minimally invasive, cost-effective alternative to surgical biopsy for histological diagnosis. Ultrasound guidance for breast core biopsy using a 14-gauge needle has been shown to have 100% correlation with surgical biopsy in a study of 181 lesions (Parker *et al.*, 1993). Ultrasound guidance for core biopsy of breast masses is advantageous because the procedure is done in real time and precise location of the needle can be determined throughout the procedure.

Thyroid

Ultrasound is superior to physical examination for the detection of thyroid nodules and localization of palpable neck masses. The majority of thyroid nodules are benign and there is considerable overlap of appearances between benign and malignant thyroid nodules, but nodules with irregular margins, microcalcifications and thick irregular peripheral haloes on ultrasound images are suspicious for malignancy (Tessler and Tublin, 1999) (**Figure 6**). Papillary, medullary and Hurthle cell thyroid cancers



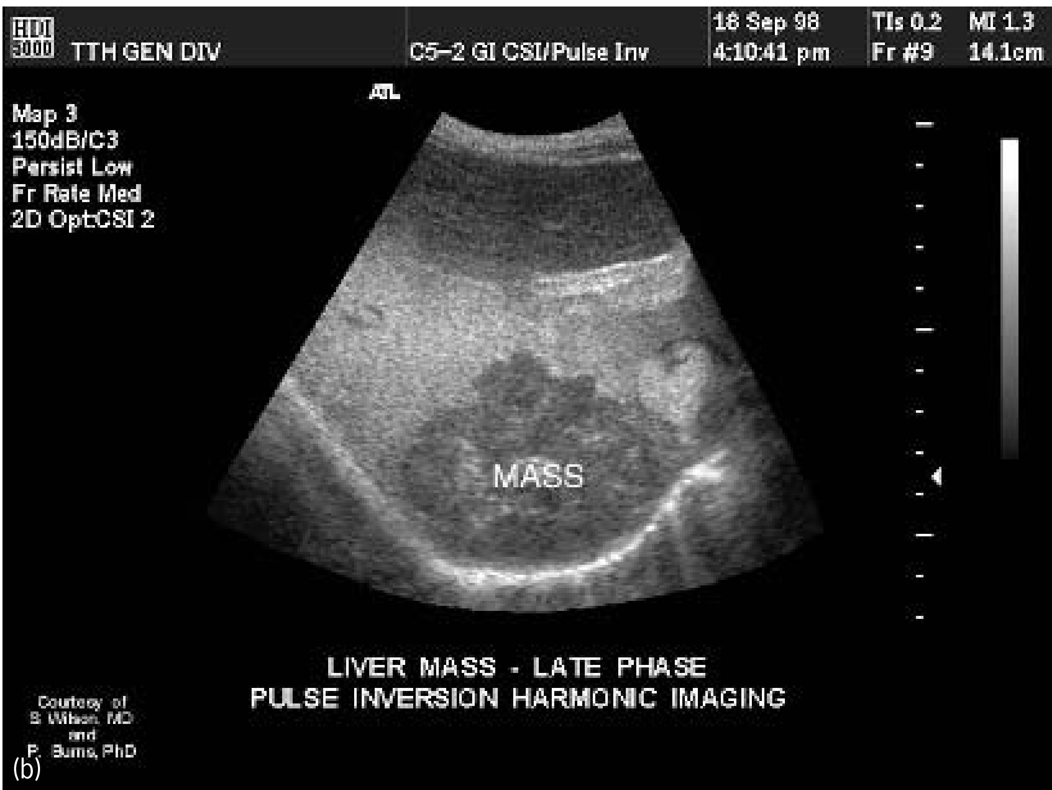


Figure 5 Images of the liver pre- and postadministration of ultrasound contrast show improved visualization of liver mass after contrast administration. Images courtesy of ATL Ultrasound, a Philips Medical Systems Company (Advanced Technology Ultrasound, Bothell, WA, USA). (a) Transverse view of the liver shows a mass posteriorly. (b) Following administration of intravenous microbubble contrast, the contour of the mass is better delineated.

frequently metastasize to cervical nodes. Metastatic nodes are typically enlarged, round in shape and have inhomogeneous texture, calcifications and distorted internal vascularity.

Thyroid nodules and cervical lymph nodes are easily biopsied under sonographic guidance for cytological diagnosis. Fine-needle biopsy, rather than large-bore core biopsy, is used because the gland is small, hypervascular and adjacent to major blood vessels. Ultrasound guidance for thyroid biopsy improves accuracy of diagnosis and reduces insufficient samples compared with standard fine-needle aspiration biopsy.

In patients with thyroid carcinoma previously treated by total thyroidectomy and radioablation, radioiodine uptake and serum thyroglobulin assays have been routinely used to monitor for recurrence. Ultrasound is being used increasingly in this setting since 20% of tumours are not iodine avid and antibodies to thyroglobulin may give spurious results. In the post-thyroidectomy cancer patient, ultrasound may detect residual or recurrent disease in the neck, and masses in the thyroid bed or abnormal cervical lymph nodes may be biopsied under ultrasound guidance for cytopathological diagnosis (**Figure 7; see colour plate section**). For the subgroup of thyroid cancer patients who have been treated by subtotal thyroidectomy, radioiodine and thyroglobulin cannot be used since there is residual thyroid tissue. For these patients, ultrasound is superior to physical examination, the only other means of monitoring for recurrent disease.

Scrotum

Ultrasound has sensitivity approaching 100% for detection of intratesticular masses and it is the gold standard for the diagnosis of testicular cancer (Benson *et al*., 1989). Ultrasound has 98% accuracy for distinguishing between intra- and extratesticular scrotal masses. This distinction is essential since any intratesticular solid mass is suspicious for testicular carcinoma whereas masses arising in the epididymis are commonly benign. All solid intratesticular masses are considered potentially malignant and are treated surgically by inguinal orchiectomy. Percutaneous biopsy is not performed for intratesticular masses because it increases the risk of tumour spread.

The majority of primary testicular cancers appear as hypoechoic masses on ultrasound. Germ-cell tumours are most common, representing approximately 90–95% of all testicular neoplasms, and the majority of tumours have mixed cell type. There are a few ultrasound features that may be used to distinguish between histological subtypes, but correlation with age at presentation and serum tumour markers is also helpful. Seminomas are the most common single cell type and usually are hypoechoic, homogeneous

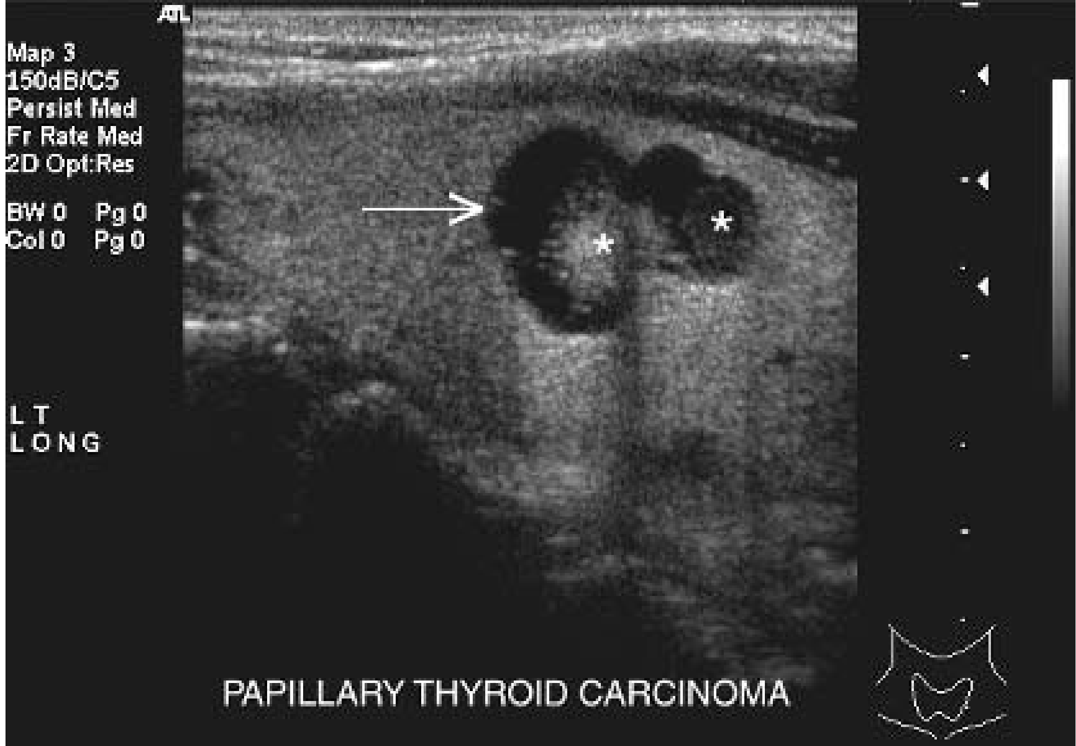


Figure 6 Longitudinal ultrasound image of a papillary thyroid carcinoma (arrow) shows that it contains fluid (dark areas) and solid papillary projections (asterisks).

and smoothly marginated on ultrasound images. Embryonal cell tumours, teratomas and choriocarcinomas are more often inhomogeneous, and may contain cystic areas or calcification evident on ultrasound. Doppler ultrasound provides little additional information for characterization of testicular tumours because tumour hypervascularity is proportional to increased tumour size rather than histopathology.

Multiple masses or bilateral intratesticular masses on ultrasound most frequently represent metastases from leukaemia, lymphoma, lung, genitourinary tumours or melanoma. Leukaemia and lymphoma are infiltrative and appear as an enlarged hypoechoic heterogeneous testicle on ultrasound. The testicle is a sanctuary site for leukaemic cells during chemotherapy and relapse in the testicle detected by ultrasound is relatively common.

Scrotal ultrasound is also valuable in men who present with mediastinal, retroperitoneal or supraclavicular lymphadenopathy to identify a primary occult testicular tumour. Testicular tumours may metastasize widely despite their small size, and the primary tumour may regress with only a residual echogenic shadowing focus seen on ultrasound to indicate the site of the 'burned-out' primary that becomes calcified (**Figure 8**).

Uterus

Ultrasound imaging of the female pelvis may be performed transabdominally with a filled urinary bladder or transvaginally by a high-frequency ultrasound probe placed in the vagina. The transvaginal technique provides optimum resolution because the probe is close to the uterus and ovaries, and a higher frequency transducer can be used.

Uterine cancer is usually diagnosed by endometrial biopsy performed for the clinical symptom of postmenopausal bleeding, but it may also be detected by demonstration of thickened endometrium exceeding 5 mm on transvaginal ultrasound of a postmenopausal woman (Dubinsky *et al.*, 1997). Endometrial cancers may be heterogeneous in echogenicity and hypervascular on Doppler imaging, but these signs are less reliable than endometrial thickness measurements for diagnosis of endometrial cancer. Depth of myometrial invasion in endometrial cancer is poorly assessed by ultrasound and magnetic resonance imaging is considered superior for that determination.

Saline infusion sonohysterography may be used also to evaluate endometrial abnormalities (Laifer-Narin *et al.*, 1999). Saline infusion hysterography is performed by instillation of sterile saline into the uterine cavity followed immediately by imaging with transvaginal ultrasound. The fluid outlines the contour of the endometrium and aids diagnosis of diffuse endometrial thickening and focal masses. Such information may guide subsequent intervention. Diffuse endometrial thickening may be biopsied by blind endometrial biopsy in the gynaecologist's office. Focal endometrial masses require dilation, curettage and

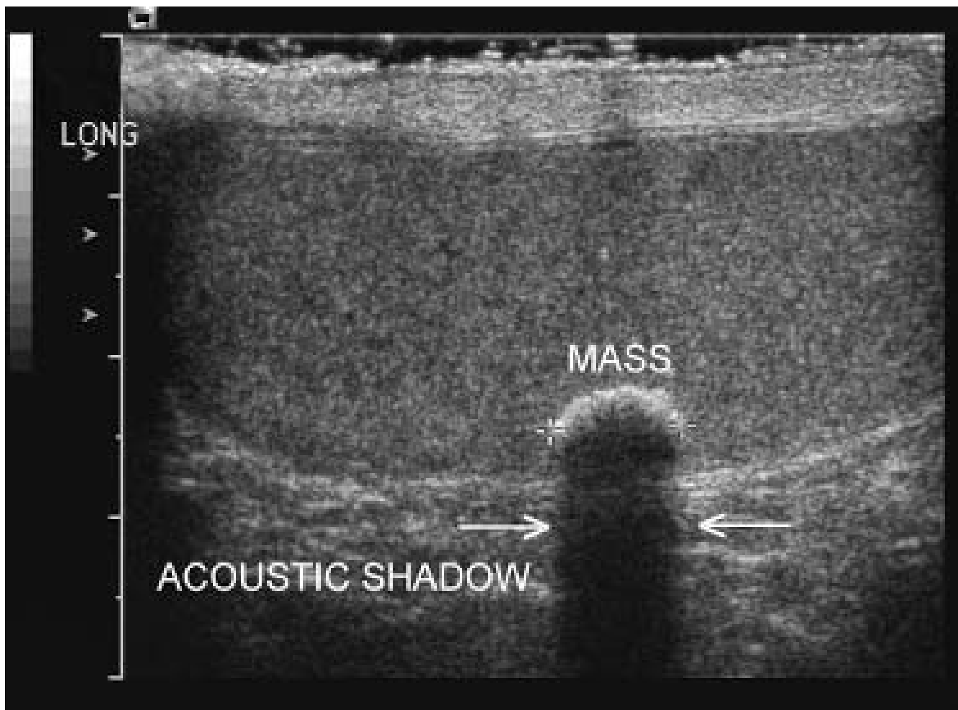


Figure 8 Longitudinal scrotal ultrasound image reveals a primary intratesticular germ-cell tumour that is calcified with resultant posterior acoustic shadowing.

hysteroscopy for removal. Saline infusion sonohysterography is particularly useful for the assessment of endometrial abnormalities in breast cancer patients treated with tamoxifen (**Figure 9**) Tamoxifen has a weak oestrogenic effect that results in endometrial abnormalities including endometrial carcinoma, hyperplasia and polyps. On saline infusion sonohysterography, polyps are shown as masses that often contain cystic spaces and project into the endometrial cavity. In contrast, endometrial carcinoma and hyperplasia cause focal or diffuse endometrial thickening.

Ovary

Ultrasound diagnosis of ovarian cancer is based on ovarian volume, morphology and blood flow. Ultrasound provides excellent visualization of the ovary in all three dimensions and is the best imaging method for evaluating ovarian internal architecture. Normal ovarian volume in premenopausal women ranges up to 18–20 cm^3. After menopause ovarian volume ranges up to 8 cm^3, but the ovaries continue to decrease in size with advancing age. An enlarged ovary for age and increasing ovarian volume on serial ultrasound examinations are findings suspicious for ovarian carcinoma.

Ovarian carcinomas and benign ovarian masses are often cystic on ultrasound, but the morphology of the ovary may differentiate between benign and malignant disease. Benign ovarian cysts have smooth, thin walls, no internal echoes and enhancement of the ultrasound signal posterior to the cyst. Cysts with thin septations or uniform low level echoes throughout are also probably benign. Solid ovarian masses and cysts with thick irregular walls, nodular septations and solid components are suspicious for malignancy by ultrasound criteria (**Figure 10**). Metastases to ovaries are often bilateral and may be either cystic or solid depending on the primary carcinoma. Breast, pancreas and gastric metastases to the ovary are typically solid whereas secondary tumours from mucin-secreting primaries from colon or stomach are more likely cystic. In premenopausal women physiological cysts that occur as part of the menstrual cycle may mimic the ultrasound appearance of ovarian neoplasms, but physiological cysts are transient. Premenopausal women with suspicious ovarian masses should always have follow-up sonograms in 6 weeks or after the next menses, to exclude physiological cysts and decrease false-positive ultrasound diagnoses.

Assessment of tumour angiogenesis by Doppler flow is another method for diagnosis of ovarian cancer by ultrasound. Ovarian tumours may have an increased number of blood vessels that can be shown by power Doppler imaging and the distribution of vascularity in ovarian cancers may be more central in location than in benign ovarian masses, which often show peripheral distribution of blood flow. Tumour vessels lack a muscular layer and are characterized by low-impedence, low-resistance blood flow that may be measured with the duplex Doppler technique. Vascular impedence is measured by ratios such as resistive index and pulsatility index that compare flow in diastole relative to systole. When diastolic flow is increased as with tumour vascularity, the ratios are low. Unfortunately, many benign conditions, including the physiological corpus luteum of the normal

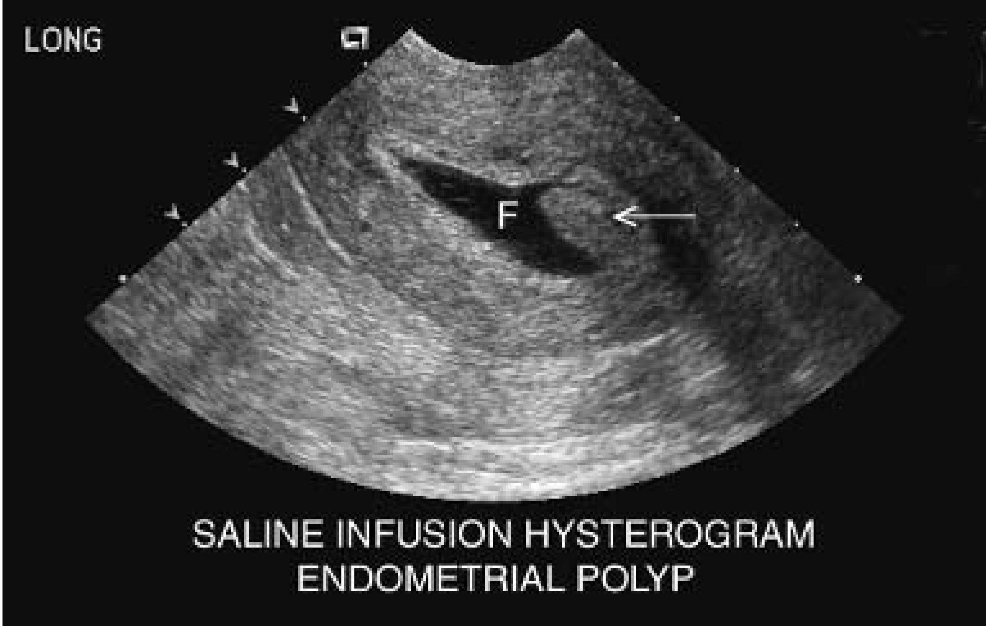


Figure 9 Longitudinal image of the uterus from a saline infusion sonohysterogram in a tamoxifen-treated breast cancer patient shows fluid (F) filling the endometrial cavity and an endometrial polyp (arrow) in the fundus.

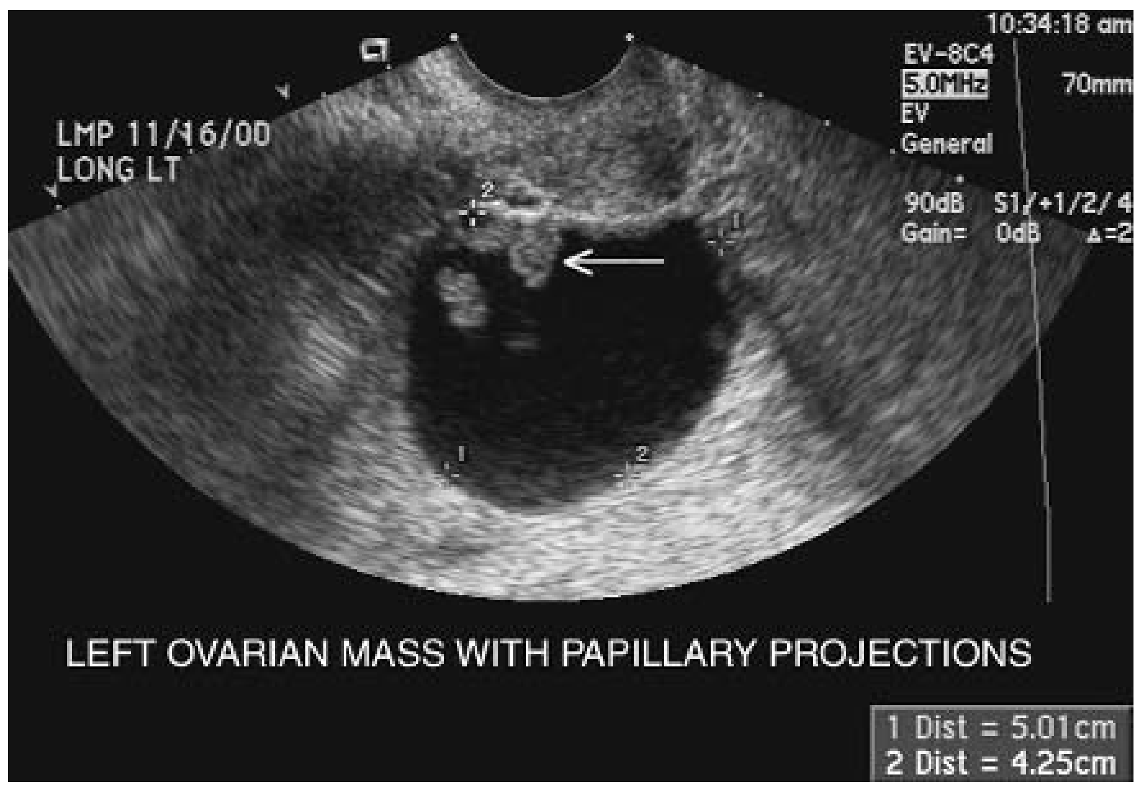


Figure 10 Transverse ultrasound image of an ovarian carcinoma reveals suspicious solid mural nodules (arrow) projecting from the wall of the cyst.

menstrual cycle, have similar flow patterns. Despite initial optimism, Doppler indices lack sufficient specificity to serve as independent predictors of malignancy, particularly in premenopausal women.

The best results for ultrasound diagnosis of ovarian cancer are obtained by the combination of Doppler and ovarian morphology as shown in a meta-analysis of 5159 patients studied with current ultrasound technique (Kinkel *et al*., 2000). Using this approach, ultrasound has approximately 93% sensitivity and 93% specificity for diagnosis of ovarian cancer. The best predictors of malignancy are considered the presence of solid components within an ovarian mass, central flow on Doppler, ascites and septations within a mass (Brown *et al*., 1998). Results of ultrasound are improved further by correlation with serum tumour markers, such as CA-125 that is elevated in women with advanced ovarian cancer and in 50% of women with stage 1 ovarian cancer.

There has been interest in using ultrasound to screen women for ovarian cancer since ultrasound has such high diagnostic accuracy for detection and characterization of ovarian masses. However, there is a difference between using a test for diagnosis in the clinical setting versus screening of asymptomatic women. For diagnostic ultrasound women are symptomatic or have suspected pelvic pathology and the prevalence of disease is relatively high. A positive test result is therefore more likely to identify ovarian cancer correctly since it is more common within the group. In contrast, the majority of asymptomatic women in a screened population are normal and the prevalence of ovarian cancer is low even in women with a family history of ovarian cancer. Since ovarian cancer is infrequent in a screened group, a positive test result is less reliable owing to more false-positive diagnoses because the proportion of women with ovarian cancer is so low. Current recommendations are that available evidence does not support ultrasound screening for ovarian cancer in pre- or postmenopausal women with or without a family history of ovarian cancer.

Liver and Bile Ducts

Ultrasound is used to detect primary and metastatic hepatobiliary tumours and to provide guidance for diagnostic biopsy. Ultrasound also has a role for preoperative and intraoperative planning to determine if tumours are amenable to surgical resection. Often management of patients with hepatobiliary tumours requires the use of multiple imaging modalities. The choice of cross-sectional imaging technique may vary by availability, cost and experience. In Europe and the UK ultrasound is more widely used for imaging of hepatic malignancies; in the United States computed tomography is preferred.

Hepatic tumours appear as focal masses that have different echogenicity than adjacent normal liver. Lesion

detection is determined by the relative difference in echogenicity rather than size of the tumour. Small tumours, approximately 1 cm in diameter, are detectable by ultrasound if there is sufficient contrast between the tumour and normal liver, but if the tumour has echogenicity similar to liver, it may be difficult to visualize even if it is relatively large. Conspicuity of a liver mass may be different on ultrasound as compared with computed tomography or magnetic resonance imaging because ultrasound is based on the acoustic properties of tissue rather than tissue attenuation in a radiation beam or magnetic field. Harmonic ultrasound has been shown to improve the ultrasound detection of hepatic masses through better soft tissue contrast. In a study comparing harmonic imaging of liver tumours with standard ultrasound, eight of 48 patients had lesions evident only on harmonic ultrasound (Hann *et al*., 1999). Harmonic ultrasound is now used routinely as the primary ultrasound technique for detection of hepatic lesions.

Hepatocellular carcinoma is the most common primary hepatic tumour and has an especially high incidence in Asia and Africa. These tumours may be focal, infiltrative or multiple and the ultrasound appearance is variable. Small hepatocellular carcinomas <5 cm in diameter are usually less echogenic than the adjacent liver and may have a thin peripheral halo. Larger tumours are ill-defined and heterogeneous owing to areas of fibrosis and necrosis. Doppler ultrasound may detect increased vascularity within these hypervascular tumours that have arteriovenous shunts and high velocity arterial flow with frequency shifts exceeding 4.5 kHz (**Figure 11; see colour plate section**). One study found that 32 of 46 hepatocellular carcinomas, four of 86 metastases and none of 66 benign hepatic lesions had frequency shifts of 4.5 kHz or greater (Reinhold *et al*., 1995). Doppler evaluation may also suggest the diagnosis of hepatocellular carcinoma by identification of portal vein thrombosis that occurs in 40% of hepatocellular carcinomas and hepatic vein thrombosis, present in approximately 15% of hepatocellular carcinomas.

Metastases to liver are multiple and echogenicity is variable. Echogenic metastases are often of gastrointestinal origin or are from vascular tumours such as carcinoid, pancreatic islet cell, choriocarcinoma or renal cell carcinoma. A hypoechoic halo at the peripheral margin of a liver mass is highly suggestive of colorectal metastasis to liver, but may also be seen with hepatocellular carcinomas. This target or bull's eye appearance is due to distended sinusoidal spaces and new tumour vessel formation (Kruskal *et al*., 2000) (**Figure 12**). Appearance of metastases may provide clues for site of origin. For example, calcified metastases are seen from mucin-secreting tumours of the colon or medullary carcinoma of the thyroid, and cystic, fluid-containing metastases often arise from ovary or colon carcinoma, sarcoma or melanoma. Lymphoma, lung, melanoma and pancreas metastases are frequently hypoechoic on ultrasound. Chemotherapy may alter the

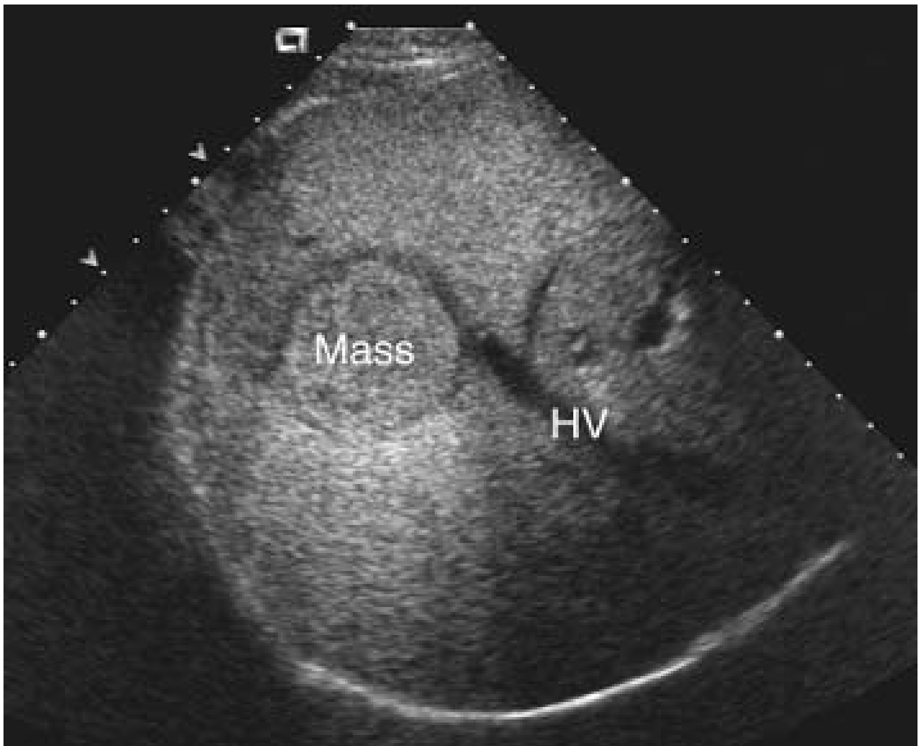


Figure 12 Transverse ultrasound image of the liver reveals a mass with surrounding hypoechoic (dark) halo that is characteristic of colorectal metastasis. HV = hepatic vein.

appearance of metastases and the background echogenicity of the liver. Breast carcinoma metastases to liver are often mixed in echogenicity prior to treatment, but may appear echogenic postchemotherapy. Chemotherapy-associated fatty change in the liver produces heterogeneous texture with patchy areas of increased echogenicity. This ultrasound appearance due to metabolic change in the liver may make metastases less evident or less well defined on ultrasound images. For this reason, computed tomography is preferred to monitor the response of hepatic metastases to chemotherapy.

Ultrasound is useful for the preoperative assessment of patients with primary and secondary hepatic malignancies that are being considered for curative resection. For the 25% of patients with colorectal metastases limited to the liver, surgical resection offers improved survival in the range 20–30% compared with chemotherapy. Hepatic resections are performed along planes between anatomical hepatic segments each of which has portal venous inflow, hepatic venous outflow and bilary drainage. Using the segmental approach to hepatic resection it is possible to remove up to approximately 70% of the liver because the remaining disease-free liver has intact vascular supply and biliary drainage and the liver hypertrophies in volume postoperatively. For preoperative planning, computed tomography or ultrasound is used to determine the distribution of tumour within the liver and to ensure that the remaining liver has normal vascularity. Tumours confined to a localized area of the liver are potentially resectable but scattered tumours throughout multiple segments are not amenable to surgical resection. Surgery is also contraindicated if tumours involve one side of the liver and the portal or hepatic vein on the opposite side, since blood flow must be preserved in the remaining liver. Doppler and grey scale ultrasound are used in combination to evaluate the hepatic vasculature preoperatively. Encased veins have irregular contour and narrowed calibre. When the portal or hepatic veins are thrombosed, the vessel may appear expanded or have an intraluminal filling defect. Demonstration of arterial flow within an occluded expanded portal vein or hepatic vein is diagnostic of tumour growth into the vessel as may be seen with hepatocellular carcinoma.

Biliary tumours are managed by curative surgical resection or by placement of stents or biliary drainage catheters to relieve bile duct obstruction. Ultrasound is reliable for determining the level of bile duct obstruction and to image the tumour mass directly. As with hepatic tumours, the distribution of bile duct involvement within the hepatic segments determines whether surgical resection is feasible. Cholangiocarcinomas appear as nodular bile duct masses, irregular scirrhous tumours that stricture the bile ducts or intraluminal papillary masses that expand the duct. The scirrhous and nodular tumours are locally invasive and often encase or occlude adjacent hepatic blood vessels. It is preferable to image bile duct tumours prior to placement of biliary drainage catheters because the catheter allows reflux of intestinal air into the biliary tree and the pneumobilia causes artifacts that may obscure the bile duct tumour.

Intraoperative ultrasound provides higher resolution for imaging hepatic and biliary tumours. Intraoperative ultrasound is performed with the transducer placed directly on the hepatic surface. Because the liver is closer to the transducer, it is possible to scan with higher frequency, higher resolution probes. Intraoperative ultrasound is reported to detect 25–35% additional liver lesions in patients who have been evaluated preoperatively by computed tomography, ultrasound or magnetic resonance imaging (Kane *et al*., 1994). In a study of patients with hepatic cancer, Kane *et al*. reported that intraoperative ultrasound altered surgical management in 51% of patients. The relative contribution of intraoperative ultrasound for the detection of additional liver masses may decrease as thin-slice, high-resolution, multidetector computed tomography scanners are more widely available.

Kidney

It is now recommended that computed tomography be used for the initial evaluation of patients with microhaematuria in whom renal or bladder tumours are suspected clinically. This change in the imaging algorithm for exclusion of renal masses has altered the role of ultrasound. Renal ultrasound in the cancer patient is primarily used to characterize renal masses detected by computed tomography and to exclude hydronephrosis in patients with renal failure or abdominal masses.

The excellent cyst–solid differentiation of ultrasound makes it ideal for the evaluation of renal masses that are indeterminate by computed tomography. Even small renal cysts less than 1 cm in diameter may be resolved by ultrasound. Renal cysts with thin walls and no internal septations are benign; any solid renal mass is considered suspicious for renal cell carcinoma. Cystic masses with thick nodular septations or calcifications also may be malignant. If internal vascularity is present within a renal mass on Doppler imaging, the likelihood of malignancy is increased. It is easier to identify a renal cyst than a solid renal mass on ultrasound because there is a great deal of contrast between the fluid-containing cyst and the adjacent normal renal parenchyma. Small solid renal masses are more difficult to visualize by ultrasound but variations in echogenicity, Doppler blood flow patterns and abnormalities in contour of the kidney are all used to improve detection (Forman *et al*., 1993; Jinzaki *et al*., 1998).

Transitional cell tumours are evident on ultrasound as solid masses that are less echogenic than the fatty tissue in the renal hilum. These tumours may cause hydronephrosis or locally obstruct renal calyces. The dilated fluid-filled urinary collecting system aids identification of masses in the renal pelvis or bladder. For tumours of the prostate ultrasound is no longer used for screening. Endorectal

prostate ultrasound is performed routinely, however, for guiding systematic biopsies from various quadrants of the gland in order to establish a definitive diagnosis in men with elevated serum prostate-specific antigen.

NEW ADVANCES IN ULTRASOUND TECHNOLOGY

With advances in computer technology, manufacturers are now able to produce smaller ultrasound units and miniaturized transducers. These changes allow for the potential expansion of the role of ultrasound for imaging diagnosis. Smaller, less expensive ultrasound machines may be used by general physicians as an extension of physical examination. However, since the quality of the ultrasound examination is so operator dependent, practitioners or oncologists using ultrasound would need sufficient ultrasound training in order to standardize quality of care. Portable ultrasound units could also be used in the operating room or at the bedside for the evaluation of critically ill patients.

Miniature transducers allow for intraluminal imaging of blood vessels or the gastrointestional tract. Small transducers may be mounted on tips of catheters for angiographic imaging or attached to endoscopes for staging of gastrointestinal and biliary tumours. Similarly, miniature ultrasound probes may be used for ultrasound imaging during minimally invasive surgical procedures such as laparoscopy for tumour staging (Liu *et al*., 1995).

Three-dimensional ultrasound is also being evaluated in the clinical setting for the estimation of tumour volume, for assessment of tumour vascularity and for lesion characterization. Ultrasound three-dimensional techniques are not as developed as computed tomography and magnetic resonance imaging multiplanar imaging, but they do show potential for readily accessible volumetric data regarding size of tumours. Preliminary studies have shown that three-dimensional techniques for transvaginal ultrasound of the ovaries may improve differentiation of benign from malignant ovarian tumours (Kurjak *et al*., 2000).

Ultrasound contrast agents have been shown to improve the visualization of blood vessels and tumour angiogenesis. Currently ultrasound contrast agents are used to assess tumour viability postradiofrequency ablation (Koito *et al*., 2000). Ultrasound contrast agents may be useful also to monitor response to treatment of antiangiogenesis agents as these become more available. Another theoretical application for ultrasound contrast microbubbles is for targeted delivery of chemotherapeutic agents or gene therapy. Intravenously administered microbubbles burst when exposed to an ultrasound pulse. If the ultrasound signal was given in a selected region and the microbubbles contained therapeutic agents, this could serve as a method for selective drug delivery. This method would be particularly useful for tumours that have local regional spread, rather than systemic metastases.

CONCLUSION

Ultrasound is used for initial diagnosis of malignancy and image-guided biopsy, but not for tumour staging. It is an inexpensive, readily available imaging technique, but ultrasound requires an experienced operator for optimal results. Doppler imaging is unique to ultrasound and provides information regarding tumour vascular distribution and flow dynamics. The excellent spatial resolution and soft tissue contrast provided by ultrasound are the major advantages of this imaging technique.

REFERENCES

Benson, C. B., *et al*. (1989). Sonography of the male genital tract. *American Journal of Roentgenology*, **153**, 705–713.

Brown, D. L., *et al*. (1998). Benign and malignant ovarian masses: selection of the most discriminating gray-scale and Doppler sonographic features. *Radiology*, **208**, 103–110.

Dubinsky, T. J., *et al*. (1997). The role of transvaginal sonography and endometrial biopsy in the evaluation of peri- and postmenopausal bleeding. *American Journal of Roentgenology*, **169**, 145–149.

Forman, H. P., *et al*. (1993). Hyperechoic renal cell carcinomas: increase in detection at US. *Radiology*, **188**, 431–434.

Hann, L. E., *et al*. (1999). Hepatic sonography: comparison of tissue harmonic and standard sonography techniques. *American Journal of Roentgenology*, **173**, 201–206.

Jinzaki, M., *et al*. (1998). Small solid renal lesions: usefulness of power doppler US. *Radiology*, **209**, 543–550.

Kane, R. A., *et al*. (1994). Impact of intraoperative ultrasonography on surgery for liver neoplasms. *Journal of Ultrasound Medicine*, **13**, 1–6.

Kim, A. Y., *et al*. (1998). Hepatocellular carcinoma: power Doppler US in a contrast agent – preliminary results. *Radiology*, **209**, 135–140.

Koito, K., *et al*. (2000). Power Doppler sonography: evaluation of hepatocellular carcinoma after treatment with transarterial embolization or percutaneous ethanol injection therapy. *American Journal of Roentgenology*, **174**, 337–341.

Kinkel, K., *et al*. (2000). US characterization of ovarian masses: a meta-analysis. *Radiology*, **217**, 803–811.

Kruskal, J. B., *et al*. (2000). Hepatic colon cancer metastases in mice: dynamic *in vivo* correlation with hypoechoic rims visible at US. *Radiology*, **215**, 852–857.

Kurjak, A., *et al*. (2000). Three-dimensional ultrasound and power Doppler improve the diagnosis of ovarian lesions. *Gynecologic Oncology*, **76**, 28–32.

Laifer-Narin, S. L., *et al*. (1999). Transvaginal saline hysterosonography: characteristics distinguishing malignant and

various benign conditions. *American Journal of Roentgenology*, **172**, 1513–1520.

Liu, J. B., *et al*. (1995). Laparoscopic gray-scale and color Doppler US: preliminary animal and clinical studies. *Radiology*, **194**, 851–857.

Parker, S. H., *et al*. (1993) US-guided automated large-core needle biopsy. *Radiology*, **187**, 507–511.

Reinhold, C., *et al*. (1995). Characterization of focal hepatic lesions with duplex sonography: findings in 198 patients. *American Journal of Roentgenology*, **164**, 1131–1135.

Stavros, A. T., *et al*. (1995). Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology*, **196**, 123–134.

Tessler, F. N. and Tublin, M. E. (1999). Thyroid sonography: current applications and future directions. *American Journal of Roentgenology*, **173**, 437–443.

Tschammler, A., *et al*. (1998). Lymphadenopathy: differentiation of benign from malignant disease – color doppler US assessment of intranodal angioarchitecture. *Radiology*, **208**, 117–123.

Wilson, S. R., *et al*. (2000). Harmonic hepatic US with microbubble contrast agent: initial experience showing improved characterization of hemangioma, hepatocellular carcinoma, and metastasis. *Radiology*, **215**, 153–161.

FURTHER READING

Goldberg, B. B., *et al*. (eds) (1997). *An Atlas of Ultrasound Color Flow Imaging* (Mosby, New York).

Hann, L. E., *et al*. (2000). Diagnostic imaging approaches and relationship to hepatobiliary cancer staging and therapy. *Seminars in Surgical Oncology*, **19**, 94–115.

Kane, R. A. (ed.) (1999). *Intraoperative Laparoscopic and Endoluminal Ultrasound* (Churchill Livingstone, London).

Rumack, C. M., *et al*. (eds) (1998). *Diagnostic Ultrasound*, 2nd edn (Mosby Year Book, New York).

Chapter 78

Magnetic Resonance Imaging

Dushyant Sahani, Michael Lev, Samantha Kubaska, Sanjeeva Kalva and Sanjay Saini
Massachusetts General Hospital, Boston, MA, USA

CONTENTS

- Introduction
- Brain and Spine
- Head and Neck
- Chest
- Liver
- Pancreas and Biliary Tree
- Renal–Adrenal
- Colon–Rectum
- Cervix
- Uterus
- Ovaries and Adnexa
- Breast
- Urinary Bladder
- Prostate
- Testis
- Musculoskeletal
- Lymph Node Imaging

INTRODUCTION

Magnetic resonance imaging (MRI) plays an important role in the diagnostic evaluation of patients with known or suspected cancer. Compared with other imaging studies, MRI provides intrinsically higher soft-tissue contrast resolution and is therefore an attractive alternative to computed tomography (CT), especially when the use of iodinated contrast material is contraindicated. In addition, because MRI is devoid of ionizing radiation, it can be safely used in children, young adults and pregnant women. Recent developments have extended its role to functional imaging, which holds the promise of providing biochemical information at a molecular level. The purpose of this chapter is to provide an overview of the role of MRI in diagnostic imaging of the oncological patient.

Clinical MRI is based on the distribution of hydrogen atoms in the body. T_1 and T_2 relaxation times reflect intrinsic magnetic properties of protons in hydrogen atoms. These processes can be measured to obtain high-resolution T_1-weighted and T_2-weighted images of the body. In general, on T_1-weighted images, both water and tumours appear darker (lower signal intensity) than surrounding normal tissue (**Figures 1** and **2**) and on T_2-weighted images, both water and tumours appear brighter (higher signal intensity) than surrounding normal tissue. Common contraindications to the use of MRI include the presence of pacemakers, cochlear implants and metallic aneurysm clips of uncertain manufacture. Claustrophobia, and also the inability to lie still for 30–45 min, are also not uncommon problems. Additionally, medically unstable patients may not be able to enter the magnet safely,

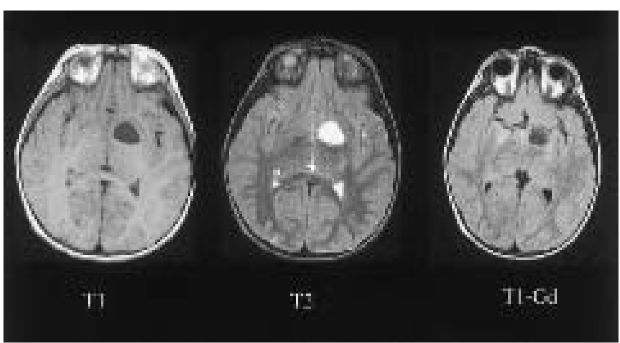


Figure 1 Low-grade glioma: well-defined T_1 low signal intensity and T_2 high signal intensity lesion located in the left basal ganglia without surrounding oedema. Minimal peripheral enhancement is seen on the postcontrast T_1 image.

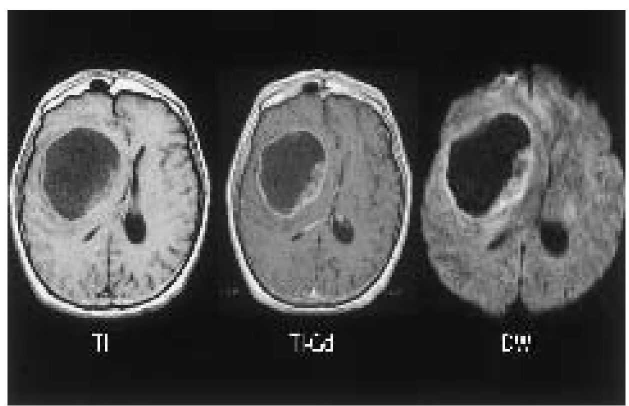


Figure 2 High-grade glioma: large necrotic mass in the right cerebral hemisphere with associated surrounding white matter oedema and mass effect on the ipsilaterlal ventricle. Heterogeneous peripheral nodular enhancement is present on post-Gd T_1-weighted image (middle). Diffusion-weighted MR (DWI) confirms a fluid necrotic centre.

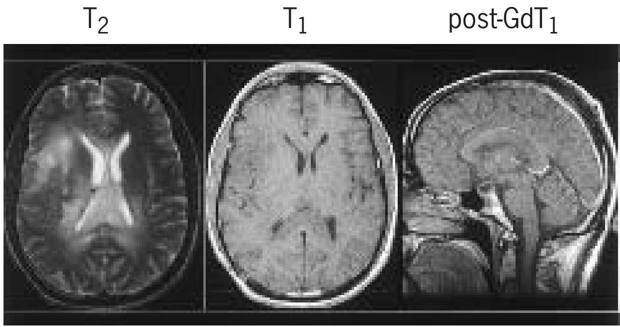


Figure 3 Glioma versus lymphoma versus demylinating disease: diffuse swelling on the distal body and splenium of corpus callosum, and right cerebral hemisphere, T_2 hyperintensity signal from white matter oedema. Lack of enhancement was observed on postcontrast T_1 image, more suggestive of demyelination.

because some monitoring and treatment devices are incompatible with the magnetic field produced by MRI systems.

BRAIN AND SPINE

MRI, because of its higher sensitivity than CT scanning in tumour detection, is typically the initial imaging modality of choice for CNS evaluation. When the presence of calcification or acute intra parenchymal haemorrhage is in question, however, CT scanning may still be required. Contrast-enhanced MRI is superior to contrast-enhanced CT in the detection of metastases. MRI is also not subject to beam hardening artifacts, which can limit CT's sensitivity in detecting leptomeningeal carcinomatosis, or small parenchymal metastases in the posterior fossa or near the calvarium (Hansberger and Dillon, 1988).

The primary role of the radiologist in oncological imaging of the central nervous system (CNS) is in the detection of tumours (**Figure 3**), MRI is well suited for this role, as it has intrinsically high sensitivity in the detection of primary or metastatic disease in brain, meninges or spinal canal. Intravenous administration of Gd–DTPA further improves the detection of intraparenchymal, leptomeningeal and dural-based metastases, and may also sometimes help in differentiating viable from necrotic tumour (Yuh *et al.*, 1992). Some authors advocate the use of double-or even triple-dose gadolinium to improve further the sensitivity of MR in lesion detection. Screening for drop metastases in the spinal canal, in cases of intraventricular or posterior fossa tumours, is best performed with MRI.

The secondary role of the radiologist in oncological imaging of the CNS is in the accurate anatomical localization of tumours, as the differential diagnosis, presurgical mapping and further work-up are highly dependent on precise lesion localization. Such localization is best achieved with MRI (Brant-Zawadzki *et al.*, 1984). Currently, since the introduction of blood oxygen level dependent (BOLD) functional MRI, the relationship between tumour location and the location of other critical structures, such as the motor and language centres of the brain, can now be accurately mapped (Buchbinder and Cosgrove, 1998).

It is difficult to predict brain tumour grade accurately using conventional MRI techniques alone. The presence or absence of enhancement, signal heterogeneity or frank necrosis only weakly correlate with neoplasm grade. Although the presence of calcifications and/or well-defined margins suggests low tumour grade, these findings are unreliable; similarly, the presence of haemorrhage or oedema and mass effect does not reliably correlate with high tumour grade (Arrington *et al.*, 1998). Moreover, the distinction between recurrent tumour and radiation-induced necrosis, using conventional MRI sequences only, is similarly problematic. Newer, dynamic contrast enhanced and spin-labelled functional MR perfusion imaging techniques, however, which assess the microvascularity of tumours (blood volume and blood flow) as markers of tumour grade, have shown great promise in the more sensitive determination of both brain tumour grade and recurrent tumour detection (Lev and Rosen, 1999). The sensitivity and specificity of perfusion MRI in this regard clearly exceeds that of nuclear medicine techniques such as FDG-PET (positron emission tomography) scanning. Functional MR spectroscopy has also been shown to play a beneficial role in the management of suspected brain tumours (Adamson *et al.*, 1998) (**Figure 4**). Image-guided neurosurgery and intraoperative MRI are additional recent developments that can assist neurosurgeons in both planning their surgical approach, and in the complete removal of brain tumours.

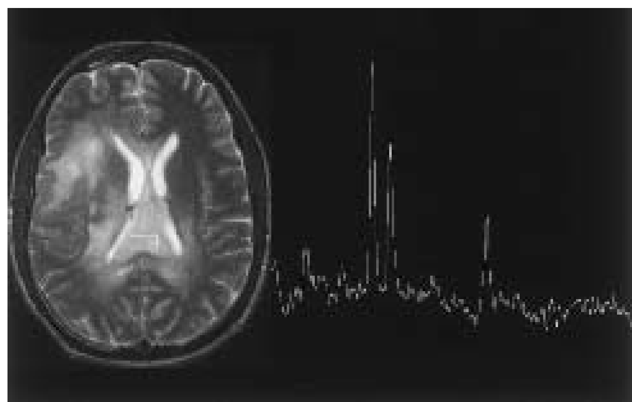


Figure 4 MR spectroscopy: spectral waveform of the signal abnormality in the distal body/splenium of corpus callosum shows reduced NAA peak favouring a diagnosis of tumour rather than demyelination. The lack of enhancement suggests gliomatosus cerebri.

Follow-up

Tumour progression is evaluated by noting the overall size, signal characteristics, pattern of contrast enhancement and associated evidence of spread of disease. The development of a new contrast-enhancing lesion or an area of T_2 hyperintense signal within post-treatment, irradiated tumour beds may represent dedifferentiation within the original tumour signifying a higher grade, tumour recurrence, radiation necrosis or post-surgical change (Dillon and Nelson, 1998). The differentiation of tumour recurrence from benign gliosis is facilitated with the use of MR spectroscopy by noting the ratios of the spectral peaks of metabolites such as creatine, choline and (*N*-acetylaspartate) (NAA) within the imaging voxel. As noted above, perfusion MRI techniques are probably more effective than PET in assessing brain tumour responses to novel therapies, and offer high resolution and convenient coregistration with conventional MRI, as well as time- and cost-effectiveness.

MRI remains the modality of choice for evaluating the spinal cord and the thecal sac (**Figures 5** and **6**).

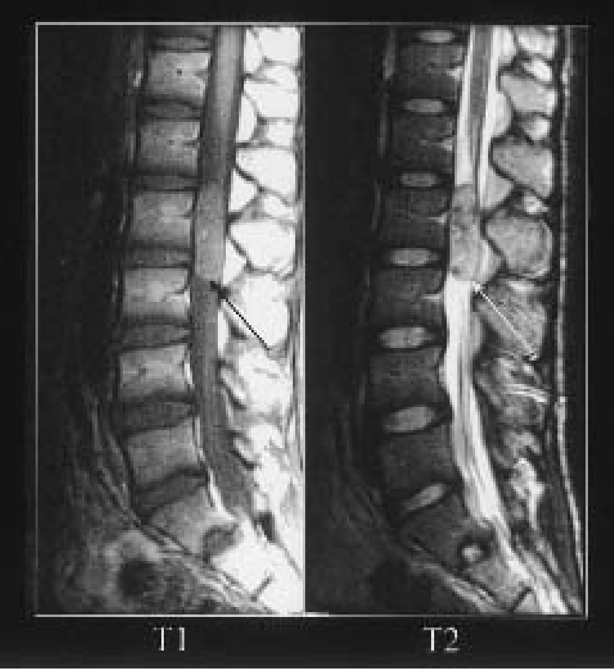


Figure 5 Ependymoma of the filum terminale: T_1 and T_2-weighted images show a well-defined intraspinal mass (arrow) arising from the filum terminale.

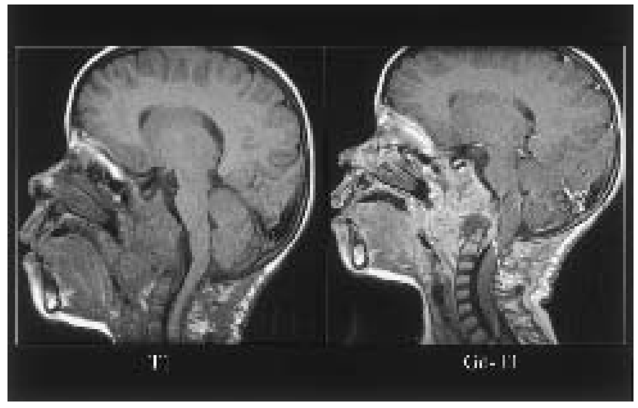


Figure 6 Clivus chordoma: pre- and postcontrast T_1-weighted image showing a large, enhancing and locally aggressive clival tumour, extending into the nasopharynx as well as intracranially (arrow).

HEAD AND NECK

Most mucosal tumours of the head and neck are detected clinically, by direct endoscopic visualization. The role of the radiologist in head and neck oncological imaging is therefore, not primarily, to detect or diagnose mucosal disease (predominantly squamous cell carcinoma), but, rather, to grade and stage tumours in anticipation of surgery and radiotherapy, and to determine submucosal or deep spread of primary or recurrent disease in 'blind spots' (such as the subglottic larynx) not amenable to direct visualization (Lev *et al*., 1998). The multiplanar capability of MRI allows complex anatomy and pathology of the skull base and neck to be well demonstrated, e.g. demonstration of perineural tumour extension into the skull basal foramina and fossa. The superior soft tissue resolution and tissue contrast of MRI allow tumour and involved lymph nodes to be differentiated from inflamed mucosa, retained secretions, fibrosis and normal structures, particularly skeletal muscle and blood vessels (Tabor and Curtin, 1989). Advantages of CT imaging of the head and neck include increased sensitivity for the detection of calcification, bony destruction and subtle bony erosion or skull base foraminal expansion.

Head and neck tumours typically demonstrate intermediate signal intensity, equal to or slightly greater than that of skeletal muscle on T_1-weighted MR imaging, and intermediate to high signal intensity on T_2-weighted MR imaging (the so-called 'evil grey' appearance). A very hyperintense signal on T_2-weighted imaging is more commonly due to soft tissue or mucosal inflammation or oedema, and is less suspicious of the presence of tumour. Gadolinium-enhanced T_1- or STIR- (short tau inversion recovery) weighted images are helpful in assessing intracranial extension and meningeal involvement of skull base tumours (**Figure 6**). Although in some series MRI has demonstrated a sensitivity as high as 100%, a specificity as high as 80% and an accuracy of almost 90% in the evaluation of head and neck tumours (Kabala *et al.*, 1992), it is critical to note again that direct visual inspection of the mucosal surfaces of the oro-, naso- and hypopharynx is essential in the work-up of head and neck neoplasms; potentially large mucosal tumours can easily be overlooked using conventional CT or MR imaging alone (**Figures 7 and 8**).

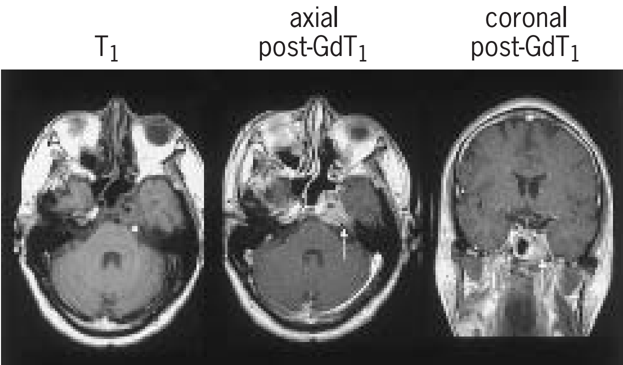


Figure 7 Skull base metastases: pre- and postcontrast T_1-weighted MR images in a patient with primary lung cancer, showing an ill-defined enhancing soft tissue mass centred in the left cavernous sinus and encasing the ipsilateral carotid artery (arrows).

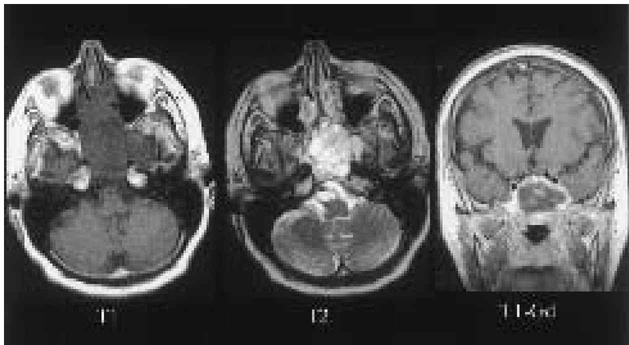


Figure 8 Sphenoid sinus cancer: a large soft tissue mass causing expansion of the sphenoid sinus with contiguous intra cranial extension.

The detection of metastatic lymph node involvement is based on size, shape and internal signal criteria (**Figure 9**). The criteria currently in use are imperfect, and represent a trade-off between sensitivity and specificity for neoplasm detection. According to one set of commonly used criteria, lymph nodes with a clinically relevant probability of metastatic involvement are considered to be those with (1) a maximum diameter of >1 cm, in the lower neck, (2) a maximum diameter of >1 cm and spherical, in the upper neck if the patient is in a high risk group, (3) a maximum diameter of >1.5 cm, in the upper neck if the patient is not in high risk group, or (4) any evidence of central necrosis (regions of increased signal intensity on T_2-weighted MR images, regions with relative lack of gadolinium enhancement or regions of focal low attenuation on CT imaging; central necrosis must always be distinguished from a 'fatty' hilum – a normal finding) (Som, 1987).

As in the central nervous system, the distinction between recurrent head and neck tumour and post-treatment change can be difficult. Early postradiation therapy changes (less than 4 months) may appear similar to inflammation, i.e. hypointense on T_1-weighted and hyperintense on T_2-weighted MR images. Beyond 6 months, with progressive fibrosis, the signal intensity changes tend to be hypointense on T_1-weighted and hypo- or isointense on T_2-weighted imaging. Criteria that suggest tumour recurrence include increased signal intensity on T_2-weighted sequences more than 6 months after radiation treatment, bulky mass beyond the original primary site and obliteration of tissue planes. These findings can often be better detected on postcontrast, fat-suppressed STIR-weighted MR images, although precise anatomical localization is typically best delineated on conventional nonenhanced T_1-weighted MR images. Preliminary results with both MR spectroscopy and MR perfusion imaging in the evaluation of native and post-treatment head and neck tumours have been encouraging.

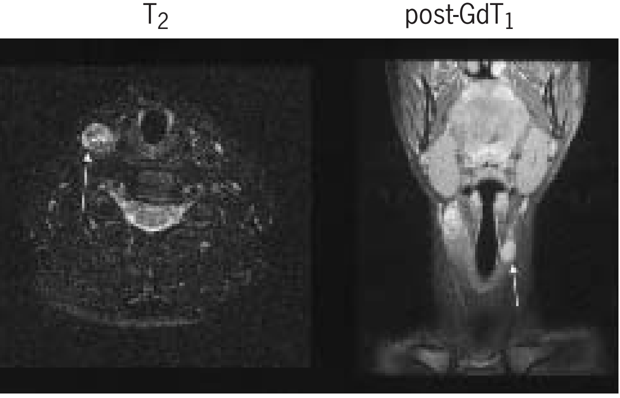


Figure 9 Recurrent thyroid cancer and lymph node metastases: T_2-weighted bright signal mass in the right neck consistent with recurrent tumour. Postcontrast T_1-weighted coronal image shows enhancing enlarged contralateral lymph node (arrow), consistent with metastasis.

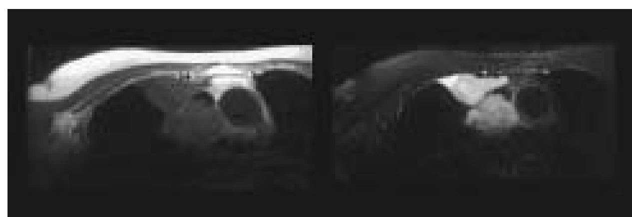


Figure 10 Primary lung carcinoma, T_1-weighted (left) and T_2-weighted (right) image of the chest are shown above which demonstrates tumour invasion into the superior vena cava.

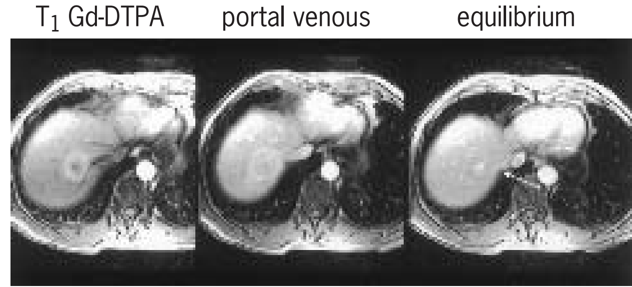


Figure 11 Liver metastases: sequential postcontrast T_1-weighted MR images show characteristic peripheral rim enhancement in the early phase images with subsequent peripheral washout of contrast on the delayed phase image (arrow).

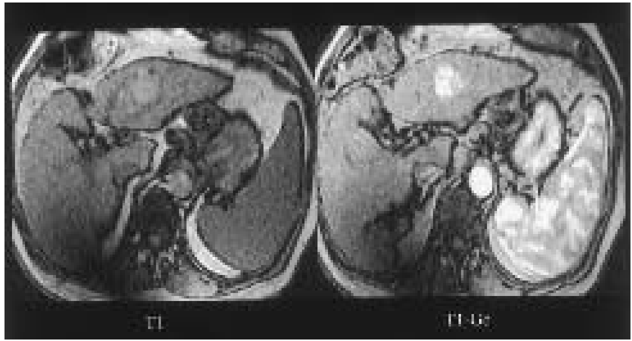


Figure 12 Hepatocellular carcinoma: pre- and post-contrast T_1-weighted axial images demonstrate cirrhotic, nodular liver with presence of a well-defined mass in the left lobe which enhances strongly in arterial phase.

CHEST

CT is the primary cross-sectional imaging modality for the evaluation of the lungs because of its superior spatial resolution, which permits better visualization of lung parenchyma and bronchial structures than does MRI. CT also helps detect calcification, which can be a useful finding in differentiating benign from malignant tumours.

MRI is a problem-solving modality in staging of the lung cancer. The multiplanar imaging capability of MRI is very useful in evaluating the extent of chest neoplasm and its relationship to the organs. Anatomical areas with a more vertical orientation – such as lung apices, thoracocervical junction, aortopulmonic window, subcarinal region and peridiaphragmatic areas – are better imaged with MRI. In patients with contraindication to iodinated CT contrast media, MRI can noninvasively assess the presence of vascular invasion. In Pancoast tumours, MRI can be utilized to determine brachial plexus involvement, chest wall invasion, spinal canal invasion and direct mediastinal invasion (White *et al*., 1993). MRI is useful in determining the involvement of pericardium, cardiac chambers and great vessels (**Figure 10**). Extension into the bronchus and trachea can be seen on MRI but CT is a better modality for such delineation.

LIVER

The liver is a common site for primary and secondary tumours. Owing to the high prevalence of benign primary hepatic tumours, the goals of liver imaging in patients with cancer are tumour detection and lesion characterization.

Liver evaluation is commonly undertaken in patients in whom, because of tumour type, symptoms of serum markers, liver metastases are suspected (**Figure 11**). The cornerstone of liver imaging is contrast-enhanced CT because it permits the rapid evaluation of the entire abdomen. Thus, live MRI is typically used when iodinated contrast media cannot be administered (e.g. prior contrast reaction; renal insufficiency). In addition, since liver lesion detection with CT or ultrasonography is compromised in patients with fatty infiltration, MRI serves an effective supplementary test. An important indication for liver MRI is the preoperative work-up of patients with resectable liver tumours (Ferrucci, 1994). More recently, MRI with liver specific contrast agents has been found to be highly effective in delineating and in staging tumour in the liver as well as in localizing tumours in various hepatic segments. MRI is also useful in patients with cirrhosis for detecting hepatomas and differentiating them from regenerating nodules and dysplastic nodules (**Figure 12**). However, owing to the high cost of MRI and limited access, ultrasonography is preferred as a screening test, leaving MRI for problem-solving equivocal sonographic studies. The role of MRI with liver-specific contrast is increasing in screening for malignant hepatic lesions, both primary and secondary.

Liver MRI is most commonly used for classifying indeterminate hepatic masses seen at CT or ultrasonography. For example, malignant liver masses can be readily differentiated from haemangiomas, cysts, focal nodular hyperplasia, dysplastic nodules and focal fatty infiltration. Signal intensity of the spleen is taken as an internal reference standard to characterize liver tumours. Typically most of the benign liver tumours are brighter

than spleen on T_2-weighted images, whereas malignant tumours have signal intensities similar to or less than those of the spleen (Wittenberg *et al*., 1988). Gadolinium–DTPA, essentially a blood pool agent, is commonly utilized to characterize tumours further by studying their perfusion pattern (Van Beers *et al*., 1997).

PANCREAS AND BILIARY TREE

Although MRI is highly effective in imaging the pancreas and biliary tree, the use of MRI in imaging these organs in the oncological population is limited. Usually, pancreatic MRI is used when administration of contrast media is contraindicated for CT. MRI may be equivalent to CT for the evaluation of the pancreas to rule out or confirm possible masses and/or stage pancreatic tumours with respect to vascular encasement and hepatic metastases (Fernando-del Castillo and Warshaw, 1990) (**Figure 13**). However in patients with suspected small functioning islet cell tumours, pancreatic MRI is superior to CT and should be the preferred initial test. Similarly, in patients with cystic pancreatic masses and a clinical suspicion of intrapapillary mucinous tumours (IPMT), MRI can be used to differentiate a focal from a diffuse process. This is undertaken with MR cholangiopancreatography (MRCP), which can also be used to noninvasively evaluate biliary obstruction (stone vs tumour) (Guibaud *et al*., 1995) (**Figure 14**). The tumour appears as a low-signal intensity mass lesion against the high signal intensity of the normal pancreas on fat-suppressed T_1-weighted spin-echo images. The tumour exhibits variable signal intensity on T_2-weighted images. Pancreatic tumours enhance less than normal pancreatic tissue and are therefore better appreciated on Gd-enhanced T_1-weighted images (Mergo *et al*., 1997). Preliminary results with MnDPDP (Teslascan, Nycomed Amersham, NJ) as a hepatobiliary contrast agent have reported a better contrast-to-noise ratio between the tumour and normal pancreas on T_1-weighted spin-echo and gradient-echo images, thereby increasing the lesion conspicuity.

Figure 13 Pancreatic cancer: pancreatic head mass best demonstrated on post-MnDPDP T_1-fat saturated image (arrow).

RENAL–ADRENAL

CT is the preferred test for evaluating the kidneys and adrenal glands. The indication for MRI is limited to patients unable to undergo a contrast-enhanced CT examination owing to prior contrast reaction or renal insufficiency. In addition, young adults with Von Hippel–Lindau syndrome who need serial scanning for monitoring the kidneys for development of malignancies are evaluated with MRI because MRI does not use ionizing radiation. Renal cell cancers appear hypointense on both T_1-weighted and T_2-weighted images. The latter is due to the high water content of the kidneys, making them hyperintense on T_2-weighted images (Semelka *et al*., 1993). The enhancement characteristics of renal masses are analagous to that of contrast-enhanced CT, with increased enhancement due to increased vascularity. The tumour thrombus enhances on Gd-enhanced images. MRI is useful to show the extent of tumour thrombus in the IVC which has implications for the surgical approach (**Figure 15**).

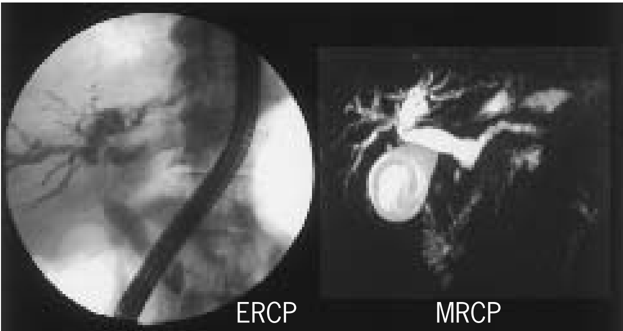


Figure 14 Cholangiocarcinoma: excellent ERCP and MRCP correlation of the level of CBD obstruction by a malignant CBD structure.

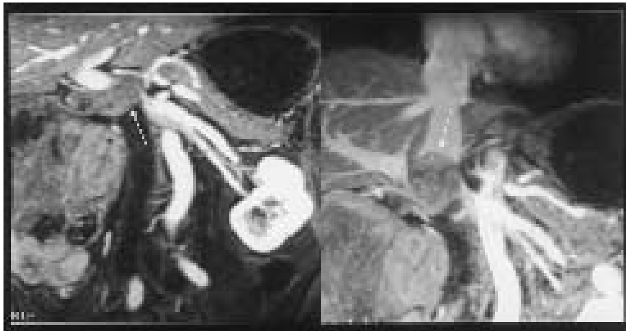


Figure 15 Renal cell carcinoma: MRI with Gd-enhanced MR angiography showing right renal tumour with contiguous invasion into the right renal vein (arrow) and the IVC (arrow).

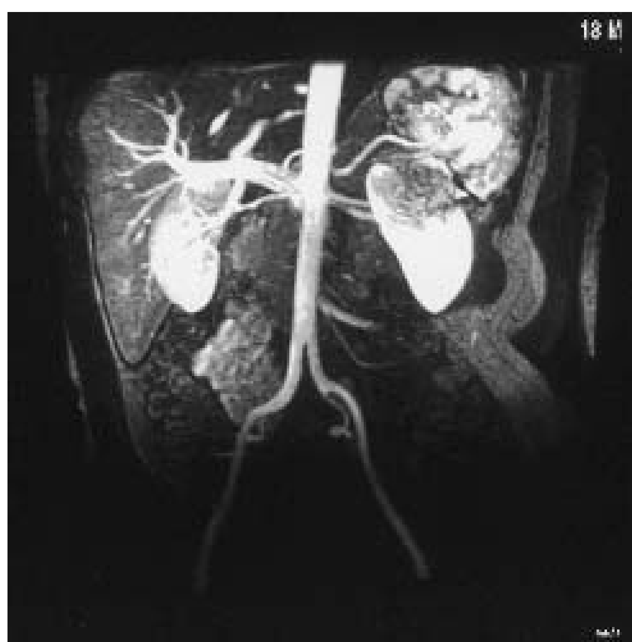


Figure 16 Renal cell carcinoma: Gd-enhanced MR angiography showing renal vessels free of the large, exophytic left upper pole renal cell carcinoma.

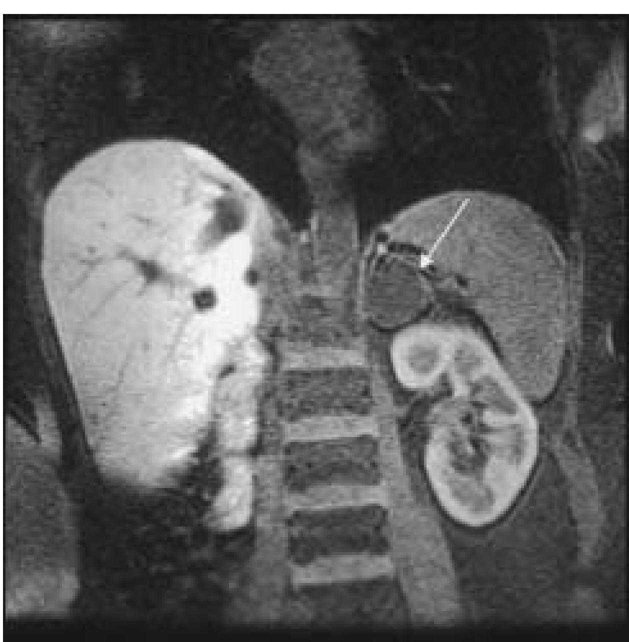


Figure 18 Adrenal metastases: post-MnDPDP enhanced T_1-weighted coronal MR image shows (arrow) a left suprarenal tumour with lack of any contrast uptake. Note the bright signal of the liver due to enhancement from MnDPDP uptake.

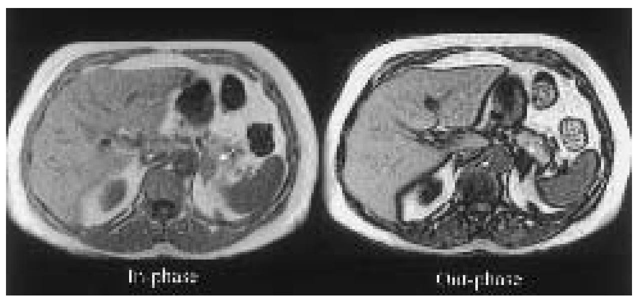


Figure 17 Adrenal adenoma: in-phase and out-of-phase T_1-weighted images. Note the drop in the signal intensity of adenoma on out-phase image.

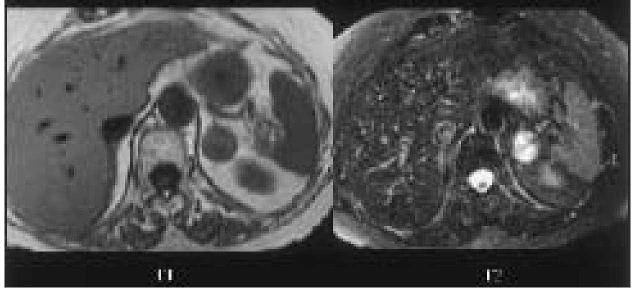


Figure 19 Adrenal phaeochromocytoma: left suprarenal tumour seen with low signal intensity on T_1-weighted image and very bright signal on T_2-weighted image.

MR imaging may be used to define whether a suprarenal mass is adrenal or renal, and to characterize a non-functioning adrenal adenoma (**Figures 16** and **17**). MRI is felt to be superior to CT for characterizing phaeochromocytomas. Adrenal adenomas resemble metastases in their signal intensities, with low signal on T_1-weighted and high signal on T_2-weighted images, but unlike metastases, adenomas may lose signal on the out-of-phase MR images in comparison with their signal on the in-phase images (Outwater *et al.*, 1996) (**Figures 17** and **18**). Phaeochromocytomas are characterized by their extreme hyperintensity on T_2-weighted images (**Figure 19**). Adrenal cortical carcinomas have no unique features that can be used to distinguish them from metastases.

COLON–RECTUM

Advances in the treatment of rectal carcinoma have increased the importance of accurate preoperative staging. Knowledge of the depth of tumour spread through and beyond the bowel wall influences the selection of patients who will benefit from preoperative adjuvant therapy (Drew *et al.*, 1999). CT is the primary modality to assess the lymph node and distant metastases. Endorectal coil MRI can be used as a focused study to evaluate the tumour extent within the muscle layers to differentiate the T_2 and T3 lesions and detect involvement of anal sphincter of levator ani muscle by rectal cancer (**Figure 20**). On T_1-weighted spin-echo images, rectosigmoid tumours appear as wall

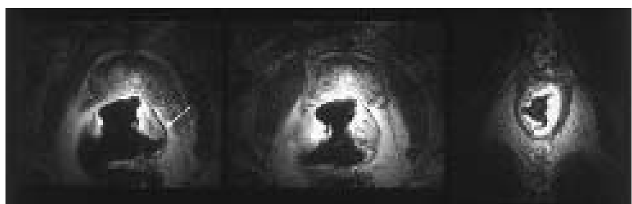


Figure 20 Rectal cancer: endorectal coil MRI shows a polypoid intraluminal tumour limited to the rectal wall as the peripheral dark stripe of serosa is intact (arrow).

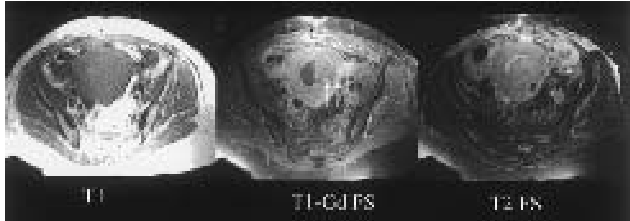


Figure 21 Endometrial cancer: polypoid intraluminal mass filling the endometrial canal is evident on T_1-weighted (left) and T_2-weighted (right) images. Contrast-enhanced T_1-weighted image (middle) shows true depth of myometrial invasion by the tumour (arrow).

thickening with a signal intensity similar to or slightly higher than that of skeletal muscle. Extension of the tumour beyond the bowel wall into perirectal fat (T3) can be well delineated on these T_1-weighted images, as the infiltrating tumour is of intermediate signal against the high signal of perirectal fat. The findings which suggest that a lesion is a T3 include extension of tumour beyond the contour of the interface between muscle and fat, with a rounded or nodular advancing margin. Gadolinium-enhanced T_1-weighted images improve definition of the tumour margins (de Lange *et al*., 1990). Coronal plane MRI is useful to establish the involvement of levator ani muscle and anal sphincter. Involvement of adjacent organs can be identified on axial and coronal MRI. Diagnosis of recurrent tumour is limited by inability to differentiate tumour, oedema, inflammation and fibrosis within 1 year, of radiation therapy. Tumour recurrence is suggested by a combination of high signal intensity on T_2-weighted images, round margins and the presence of >40% contrast enhancement.

CERVIX

MRI is the investigation of choice for local staging of the cervical cancer. The accuracy of MRI is superior to that of CT in assessing tumour size, defining parametrial and vaginal invasion and detecting regional lymph node involvement. Typically on T_2-weighted images the tumours appear brighter than the dark cervical stroma. Local spread into the surrounding fat is best appreciated on post-gadolinium fat saturated T_1-weighted images (Bragg and Hricak, 1993). Additionally, contrast administration is useful in distinguishing biopsy changes from residual tumour and in the evaluation of bladder and rectal invasion.

UTERUS

The overall prognosis in endometrial cancer depends on histological grade of the tumour, depth of myometrial invasion and lymph nodal involvement. When the tumour is confined to the uterus, determination of myometrial involvement is very important, as deep myometrial invasion is associated with a 40% incidence of lymph nodal involvement. This is better demonstrated on MRI than on CT. The indications for MRI in endometrial cancer include abnormal but nondiagnostic endocervical curettage, suspected advanced disease and patients at a poor risk for surgical staging.

Contrast-enhanced MR plays a major role in differentiating tumour margins from normal nonenhancing endometrium (Bragg and Hricak, 1993) (**Figure 21**). This also helps to determine tumour necrosis and volume of viable tumour. The staging accuracy of MRI in the assessment of patients with endometrial cancer has been reported to be in the range 83–92%. Lymph node involvement is judged on size criteria, and any node larger than 1 cm in the pelvis is considered abnormal unless proved otherwise. A staging classification has been developed based on MRI criteria. Breach in the integrity of the junctional zone places the tumour in category 1B. Extension into the cervix is best appreciated on sagittal T_1-weighted images. A high signal intensity on T_2-weighted images outside the uterus suggests extra-uterine spread (Hricak *et al*., 1987).

OVARIES AND ADNEXA

Ultrasound remains the initial imaging modality of choice for initial assessment of ovarian masses, and CT for subsequent tumour staging. The results of MRI are similar to those of CT in the characterization and staging of the ovarian tumours, although the multiplanar ability may be advantageous. Use of contrast agents (Gd–DTPA) increases the detection accuracy of MRI up to 92%, in the separation of benign from malignant tumours (Thurnher *et al*., 1990).

The MRI features suggesting ovarian malignancy include a cystic-appearing lesion with a wall thickness exceeding 3 mm, wall nodularity or the finding of a solid mass. Extension of the ovarian tumour is suggested by

indirect findings such as presence of ascites, nodular tumour implants on the peritoneal surface and/or loculated fluid collections. Post-Gd T_1-weighted fat-saturated imaging is considered to be the most sensitive imaging technique in detecting small peritoneal implants. Direct extension into adjacent organs can also be determined with MRI.

BREAST

The role of MRI in breast carcinoma is still evolving. At present it is a problem-solving modality and has no role as a screening modality. Current indications include invasive cancer in a breast with dense parenchyma, which is sub-optimally evaluated with film screen mammogram. It also provides information regarding vascularity, chest wall invasion (**Figure 22**) as noted, and MRI is devoid of ionizing radiation. The disadvantages include inability to detect microcalcification of early cancer, low specificity and high cost. Fat-suppressed contrast-enhanced MR imaging with special breast coils provides better images but still the tail of breast is difficult to image (Stelling, 1995). False-negative results of MR are due to carcinomas with minimal vascularity, partial volume averaging of small ($\leq$5 mm) lesions, smooth margins and motion artifacts. Some of the benign breast lesions may mimic carcinomas, by virtue of contrast enhancement. These include fibroadenoma, fibrocystic change and changes due to inflammation, recent surgery or recent radiotherapy. It is not possible to detect or exclude involvement of the nipple with Paget disease by contrast-enhanced MR as the nipple–areola complex normally shows brisk enhancement (Heywang *et al*., 1989). The future applications of MR in breast imaging include MR-guided biopsy and MR-directed therapy.

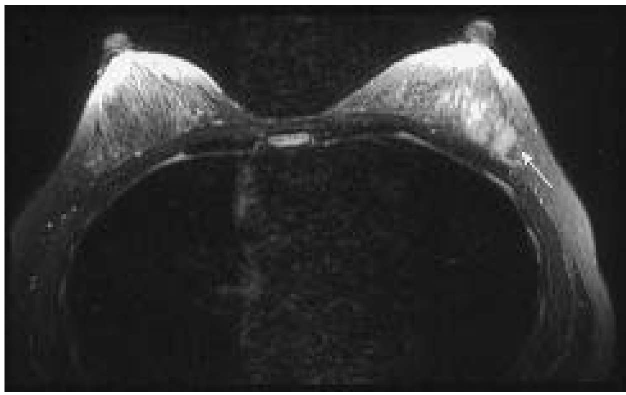


Figure 22 Breast cancer: Gd-enhanced fat saturated T_1 image shows a lobulated, poorly defined tumour suggestive of invasive cancer (arrow) seen in the left breast.

URINARY BLADDER

Treatment and prognosis of carcinoma of the urinary bladder depend on the depth of tumour infiltration into the bladder wall and the extent of the metastases. MRI and clinical staging complement each other in staging urinary bladder cancer. While clinical examination is sufficient in staging superficial bladder tumours, MRI is utilized in staging more invasive neoplasms (Fisher *et al*., 1985). MRI is superior to CT for staging of bladder cancer because of its multiplanar ability, which allows better imaging of trigone, prostate and seminal vesicles. The differentiation of T3a (muscular invasion) and T3b (invasion of perivesical fat) is better on MR than on CT, especially in tumours at the bladder dome or base. The reported accuracy of MRI for staging bladder cancers varies from 73 to 96% (Bryan *et al*., 1987). On T_1-weighted images the urine has a low signal, the bladder wall intermediate and the perivesical fat a high signal, thereby providing high inherent contrast. T_1-weighted images are good for detecting the perivesical infiltration of the tumour. On T_2 images the urine has a high signal, muscle a low signal and perivesical fat a high signal. The tumour has an intermediate signal on both T_1- and T_2-weighted images. Intraluminal tumours and bladder wall infiltration is better demonstrated on T_2-weighted images. After contrast administration, tumour, mucosa and lamina propria show earlier and greater enhancement compared with muscle layer and thereby visualization of small tumours ($\leq$7 mm) is improved. Delayed images after

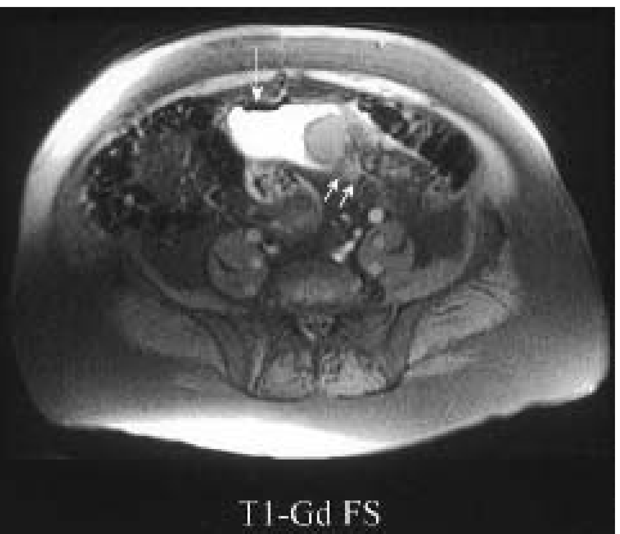


Figure 23 Bladder cancer: postcontrast T_1-weighted fat-saturated image demonstrates a lobulated intraluminal bladder mass along with contiguous extravesicular extension (small arrows) into the sigmoid colon. Presence of air in the bladder (large arrow) is indicative of colovesicle fistula.

contrast administration show the intraluminal tumour well against the high signal of excreted Gd in the bladder (**Figure 23**).

It is important to differentiate postsurgical fibrosis or granulation tissue from tumour recurrence. MR examination is usually performed 3 months after the procedure and T_2-weighted images show a high signal if it is tumour recurrence or a low signal if it is fibrosis.

PROSTATE

The primary role of MRI in imaging prostate cancer is to stage the disease, not for the detection of the disease. Currently, MRI is considered the imaging modality of choice to assess the extra-capsular spread of tumour (Bezzi *et al*., 1988). MR imaging of the prostate is best achieved with an endorectal coil as it provides higher resolution images than a body coil. Supplementary imaging with a body coil is required to image the abdomen and pelvis to detect lymphadenopathy. The T_1-weighted images are helpful in imaging the periprostatic fat, periprostatic veins, neurovascular bundles and lymph nodes. T_2-weighted images demonstrate the internal architecture of the prostate and seminal vesicles to better advantage.

Prostatic carcinoma appears in the peripheral zone as a low-signal area in the background of a normal high-signal peripheral zone on T_2-weighted images. The presence of seminal vesicle invasion is suggested by thickening of the tubules; other findings include a low signal mass, loss of normal seminal vesicle angle and low signal unilaterally or bilaterally (**Figure 24**). Coronal or sagittal images are the best for evaluating the seminal vesicles. The accuracy of MRI in the detection of transcapsular spread is 90% (Schnall *et al*., 1991). Direct invasion of the bladder and rectum can also be accurately detected on MRI. The involvement of lymph nodes is suggested by size criteria. Multiple clusters of nodes and size >1 cm are considered indicative of involvement. The accuracy of MR in detecting lymph nodes is equal to that of CT.

Figure 24 Prostate cancer: endorectal coil coronal (left) and axial T_2-weighted (right) MR image shows an ill-defined low signal intensity tumour in the superior aspect of the peripheral zone of the prostate (white arrow) with contiguous extension into the seminal vesicles (black arrow).

MR spectroscopy appears useful for noninvasive discrimination between benign prostatic hyperplasia and prostatic carcinoma. The ratio of citrate to choline peak areas and that of creatine to myo-inositol are currently considered the most useful to discriminate between advanced prostatic carcinoma (both ratios >1.0) and advanced benign prostatic hypertrophy cases (both ratios >1.0) (Garcia-Segura *et al*., 1999).

TESTIS

Ultrasonography represents the primary imaging modality to image the scrotum, because of its low cost and the effectiveness of imaging and accuracy achieved with high-frequency transducers. MRI is a supplemental imaging technique for tumour detection because of its multiplanar ability and excellent soft tissue contrast. In general, the tumours are isointense on T_1- and hypointense on T_2-weighted images relative to normal testicular tissue. MRI is useful in cases where the ultrasound is technically inadequate or the sonographic findings are ambiguous (Schnall, 1993).

MUSCULOSKELETAL

Soft-tissue Tumours

MRI is an excellent modality to evaluate soft tissue tumours. The primary role of MRI is to evaluate local extent and tumour stage. Tumour characterization is a secondary role as MRI is limited in its ability to differentiate benign lesions from malignant tumours, for which there are no absolute criteria. Benign tumours are usually small and well encapsulated with smooth margination, and exhibit signal homogeneity. Malignant tumours tend to be large (>3 cm) (Berquist, 1993), with irregular margination and invasion of adjacent organs and have heterogeneous signal intensity. Involvement of surrounding fat is well seen on T_1-weighted images whereas muscular invasion is better assessed on T_2-weighted images (**Figure 25**). Contrast-enhanced MRI is useful to delineate viable tumour and neurovascular bundle invasion and to differentiate oedema from tumour tissue. Involvement of adjacent bone is identified as an altered signal in the bone marrow. At times MRI may fail to identify subtle cortical involvement and periosteal reaction.

MRI features of some of the tumours are pathognomonic: a lipoma, for example, demonstrates a homogeneous bright signal on T_1-weighted images. However, it is not always possible to differentiate a benign lipoma from a low-grade liposarcoma. Other tumours can be difficult to characterize. Calcification within tumours can be missed on MRI and this may require correlation with plain X-ray or CT examination.

Bone Tumours

MRI plays an important role in the evaluation of malignant bone tumours, especially with the advent of limb-sparing surgery. The primary role of MRI is to stage the disease with regard to its local extent. MRI should be utilized in conjunction with plain radiography for characterizing bone tumours as the latter is a better predictor of histological diagnosis and more efficient in the differentiation of benign from malignant lesions.

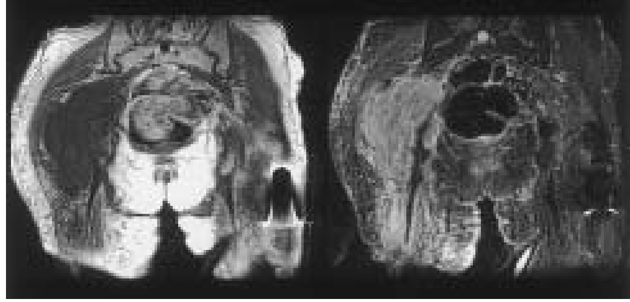


Figure 25 T_1 (left) and T_2 (right) weighted coronal MR images of a soft tissue sarcoma in the right gluteal compartment. Note the bright signal of the tumour in comparison to the surrounding musculature.

Figure 26 Bone sarcoma: contrast-enhanced T_1-weighted coronal image with fat suppression shows tumour involving distal right femur with associated pathological fracture and presence of a skip lesion in the proximal shaft (arrow).

The role of contrast agents is controversial, and currently routine use is not indicated. For selected tumours, such as osteosarcoma, gadolinium enhancement offers the potential for determining the efficacy of chemotherapy, by evaluating tumour necrosis prior and subsequent to chemotherapy. In certain situations, especially sarcomas close to joints, the use of gadolinium enhancement may help to clarify whether tumour resection should be intra- or extra-articular (Sundaram, 1997). Further, some reports demonstrate differentiation of tumour from oedema on postcontrast studies, and others have suggested that peak enhancement plotted against time may be useful to differentiate benign from malignant tumours.

Sagittal or coronal T_1 images are the most useful for evaluating the longitudinal extent of marrow involvement and possible invasion of contiguous joints. It is important to image from joint to joint to exclude skip lesions (**Figure 26**). T_2-weighted images show soft tissue involvement as a hyperintense signal against the intermediate signal of muscle. Cortical destruction and subtle tissue invasion is best appreciated on T_2-weighted axial images. When fat surrounds the bone, T_1-weighted images may also show such extension. Tumour calcification cannot be

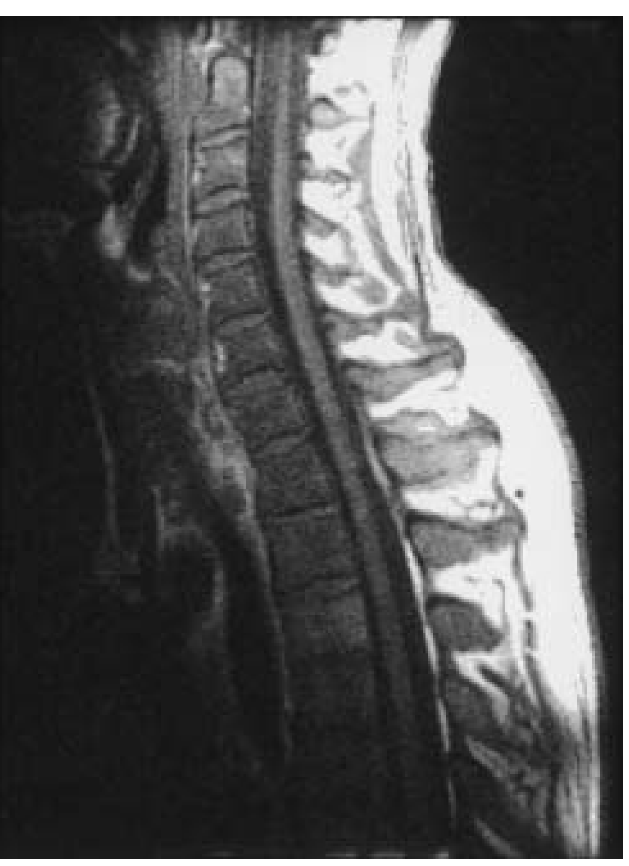


Figure 27 Spinal metastases from prostate carcinoma: diffuse alteration in the cervical and thoracic vertebral marrow signal on T_1-weighted sagittal image due to metastatic deposit.

identified unless it is gross. Involvement of neurovascular bundle should always be noted if present, as it precludes limb-sparing surgery. This also is best identified on axial T_2-weighted images. Tumour involvement is determined either by frank encasement of the neurovascular bundle or by loss of fat planes adjacent to it. MR angiography gives a preoperative map of the vessels and the tumour vascularity in certain cases. MRI plays an important role in follow-up of patients receiving radiotherapy or chemotherapy. The treatment response is assessed by evaluating for change in tumour volume and development of tumoural necrosis. This is best identified as central non-enhancing areas within the enhancing tumour on contrast-enhanced T_1-weighted images (Berquist, 1993).

The primary screening modality for bone metastases is planar bone scintigraphy. In the evaluation of spinal metastasis, however, MRI has sensitivity equal to or greater than that of scintigraphy, and it is possible to screen the entire spine on a single T_1-weighted sagittal sequence (**Figure 27**). It is also possible to ascertain the presence of impending vertebral collapse with MRI or the level of maximal spinal cord compression prior to performing decompression surgery. MRI can also be employed as a problem-solving modality when the scintigraphic findings are equivocal.

LYMPH NODE IMAGING

Lymph node imaging is commonly performed for staging lymphoma or metastatic disease. MRI and CT perform equally well in detecting lymph nodes. CT is preferred for a survey study, when more than one region must be evaluated as in staging lymphoma. MRI is reserved for a focused examination (e.g., for evaluation of pelvic nodes in prostatic or cervical malignancies) or in problem-solving situations. Size criteria are commonly utilized: usually lymph nodes >10 mm are considered pathological, except for axillary, groin and jugulo-digastric nodes, which can be as large as 15–20 mm in normal patients. The presence of bright signals and heterogeneity in the lymph node on T_2-weighted images are indicators of metastatic disease.

Recent clinical trials conducted with the reticuloendothelial system specific, iron oxide containing contrast agents have shown the ability to differentiate a benign from a malignant lymph node. This is because tumourous lymph nodes lack reticuloendothelial cells thereby reduced or heterogenous contrast uptake, in comparison with the normal homogeneous contrast uptake by normal or benign nodes.

REFERENCES

Adamson, A. J., *et al*. (1998). Focal brain lesions: effect of single-voxel proton MR spectroscopic findings on treatment decisions. *Radiology*, **209**, 73–78.

Arrington, J., *et al*. (1998). The central nervous system. In: Berman, C. G., *et al*. (eds), *Oncologic Imaging – A Clinical Perspective*. (McGraw-Hill, New York).

Berquist, T. H. (1993). Magnetic resonance imaging of primary skeletal neoplasms. *Radiology Clinics of North America*, **31**, 411–424.

Bezzi, M., *et al*. (1988). Prostatic carcinoma: staging with MR at 1.5 T. *Radiology*, **169**, 339–346.

Bragg, D. G. and Hricak, H. (1993). Imaging in gynecologic malignancies. *Cancer*, **71**, Suppl., 1648–1651.

Brant-Zawadzki, M., *et al*. (1984). Primary intracranial tumor imaging: a comparison of magnetic reasonance and CT. *Radiology*, **150**, 435–440.

Bryan, P. J., *et al*. (1987). CT and MR imaging in staging bladder neoplasms. *Journal of Computer Assisted Tomography*, **11**, 96–101.

Buchbinder, B. R. and Cosgrove, G. R. (1998). Cortical activation MR studies in brain disorders. In: Castillo, M. (ed.), *New Techniques in MR Neuroimaging: MRI Clinics of North America*, Vol. **6**, No. 2. 67–93 (W. B. Saunders, Philadelphia).

de Lange, E. E., *et al*. (1990). Preoperative staging of rectal carcinoma with MR imaging: surgical and histopathologic correlation. *Radiology*, **176**, 623–628.

Dillon, W. P. and Nelson, S. (1998). What is the role if MR spectroscopy in the evaluation and treatment of brain neoplasms? Editorial. *American Journal of Neuroradiology*, **20**, 2–3.

Drew, P. J., *et al*. (1999). Preoperative magnetic resonance staging of rectal cancer with an endorectal coil and dynamic gadolinium enhancement. *British Journal of Surgery*, **86**, 250–254.

Fernandez-del Castillo, C. and Warshaw, A. L. (1990). Diagnosis and preoperative evaluation of pancreatic cancer, with implications for management. *Gastroenterology Clinics of North America*, **19**, 915–933.

Ferrucci, J. T. (1994). Liver tumor imaging. Current concepts. *Radiology Clinics of North America*, **32**, 39–54.

Fisher, M. R., *et al*. (1985). Urinary bladder MR imaging Part II. Neoplasm. *Radiology*, **157**, 471–477.

Garcia-Segura, *et al*. (1999). *In vivo* proton magnetic resonance spectroscopy of diseased prostate: spectroscopic features of malignant versus benign pathology. *Magnetic Resonance Imaging*, **17**, 755–765.

Guibaud, L., *et al*. (1995). Bile duct obstruction and choledocholithiasis: diagnosis with MR cholangiography. *Radiology*, **197**, 109–115.

Hansberger, H. R. and Dillon, W. (1988). The radiologic role in diagnosis, staging, and follow-up of neoplasia of the brain, spine and head and neck. *Seminars in Ultrasound CT and MR*, **10**, 431–452.

Heywang, S. H., *et al*. (1989). MR imaging of the breast with Gd-DTPA: use and limitations. *Radiology*, **171**, 195.

Hricak, H., *et al*. (1987). Endometrial carcinoma staging by MR imaging. *Radiology*, **162**, 297–305.

Kabala, J., *et al*. (1992). Magnetic reasonance imaging of extracranial head and neck tumours. *British Journal of Radiology*, **65**, 375–383.

Lev, M. H. and Curtin, H. D. (1998). Layrnx. In: Gentry, L. R. (ed.) *Normal Anatomy of the Head and Neck: Neuroimaging Clinics of North America*, Vol. **8**, No. 1. 235–256. (W. B. Saunders, Philadelphia).

Lev, M. H. and Rosen, B. R. (1999). Clinical applications of intacranial perfusion MR imaging. In: Heiserman, J. E. (ed.), *Fast Scan and Echo Planar MR Imaging: Neuroimaging Clinics of North America*, Vol. **9**, No. 2. 309–331. (W. B. Saunders, Philadelphia).

Mergo, P. J., *et al*. (1997). Pancreatic neoplasms: MR imaging and pathologic correlation. *Radiographics*, **17**, 281–301.

Outwater, E. K., *et al*. (1996). Adrenal masses: correlation between CT attenuation value and chemical shift ratio at MR imaging with in-phase and opposed-phased sequences. *Radiology*, **200**, 749–752.

Schnall, M. (1993). MR imaging of the scrotum. *Seminars in Roentgenology*, **1**, 19–30.

Schnall, M. D., *et al*. (1991). Prostatic cancer: local staging with endorectal surface coil MR imaging. *Radiology*, **178**, 797–802.

Semelka, R. C., *et al*. (1993). Renal cancer staging: comparison of contrast enhanced CT and gadolinium enhanced fat suppressed spin echo and gradient echo MR imaging. *Journal of Magnetic Resonance Imaging*, **3**, 597–602.

Som, P. M. (1987). Lymph nodes of the neck. *Radiology*, **167**, 803–808.

Stelling, C. B. (1995). MR imaging of the breast for cancer evaluation – Current status and future directions. *Radiology Clinics of North America*, **33**, 1187–1203.

Sundaram, M. (1997). The use of gadolinium in the MR imaging of bone tumors. *Seminars in Ultrasound, CT and MR*, **18**, 307–311.

Tabor, E. K. and Curtin, H. D. (1989). MR of the salivary glands. *Radiology Clinics of North America*, **27**, 383–385.

Taveras, N. J., *et al*. (1990). MR imaging of bladder neoplasms: correlation with pathological staging. *Urology and Radiology*, **12**, 27–33.

Thurnher, S., *et al*. (1990). Gadolinium-DTPA enhanced MR imaging of adnexal tumors. *Journal of Computer Assisted Tomography*, **14**, 939–949.

Van Beers, B. E., *et al*. (1997). Contrast-enhanced MR imaging of the liver. *Radiology*, **203**, 297–306.

White, C. S., *et al*. (1993). Imaging of lung cancer. *Seminars in Oncology*, **20**, 142–152.

Wittenberg, J., *et al*. (1988). Differentiation of hepatic metastases from hemangioma and cysts by MR. *American Journal of Roentgenology*, **15**, 79–84.

Yuh, W. T., *et al*. (1992). Experience with high dose gadolinium MR imaging in the evaluation of brain metastases. *American Journal of Neuroradiology*, **13**, 335–345.

FURTHER READING

Barret, C. P., *et al*. (1994). *Primer of Sectional Anatomy with MRI and CT Correlation*, 2nd edn (Williams & Wilkins, Baltimore).

Beltran, J. (1997). *Current Review of MRI*, 1st edn (McGraw-Hill, New York).

Brown, M. A. and Semelka, R.C. (1995). *MRI: Basic Principles and Applications*, 1st edn (Wiley, New York).

Buthiau, D. and Khayat, D. (1998). *CT and MRI in Oncology*, 1st edn (Springer, New York).

Gerhardt, P. and Frommhold, W. (1988). *Atlas of Anatomic Correlations in CT and MRI*, 1st edn (Thieme Medical Publishers, Stuttgart).

Heuser, L. and Oedkerk, M. (1996). *Advances in MRI*, 1st edn (Blackwell, Oxford).

Kim, E. and Jackson, E. F. (1999). *Molecular Imaging in Oncology: Pet, MRI, and Mrs* (Springer, New York).

Lee, J. K. T., *et al*. (1997). *Computed Body Tomography with MRI Correlation*, 3rd edn (Lippincott-Raven, Philadelphia).

Lee, S. H., *et al*. (1999). *Cranial and Spinal MRI and CT*, 4th edn (McGraw-Hill, New York).

Ros, P. R. and Lee, S. (1997). *CT and MRI of the Abdomen and Pelvis: A Teaching File*, 1st edn (Williams & Willkins, Baltimore).

Chapter 79

Nuclear Medicine

Petra J. Lewis and Alan Siegel
Dartmouth Hitchcock Medical Center, Lebanon, NH, USA

CONTENTS

- Introduction to Nuclear Imaging
- Bone Scanning
- Iodine Scanning
- Thallium-201 Scanning
- Gallium Scanning
- Monoclonal Antibodies
- Peptides
- Scintimammography
- MIBG
- Positron Emission Tomography
- Sentinel Node Localization
- Conclusion

INTRODUCTION TO NUCLEAR IMAGING

Nuclear medicine involves the imaging of radioactive isotopes which are attached to tracer substances (hence radiotracers). These radiotracers are incorporated into physiological processes within the body depending on the type of tracer. In this way, nuclear medicine is a 'functional' as opposed to 'anatomical' imaging technique in contrast to many radiological methods such as plain films, computed tomography (CT), ultrasound and magnetic resonance imaging (MRI). The most commonly used isotope is ^{99m}Tc, but a wide variety of other isotopes are used, and many of these can be attached to multiple tracers depending on the imaging requirements (see **Table 1**). Radiotracers are usually administered to the patient intravenously, but on occasion the oral (e.g. ^{131}I capsules for thyroid imaging), inhaled (e.g. nebulized [^{99m}Tc]DPTA for lung ventilation studies), intrathecal (e.g. [^{111}In]DPTA for cerebrospinal leak studies) or intracavitary (e.g. ^{32}P for treatment of peritoneal malignancies) routes are used.

Imaging is performed on a gamma camera, which incorporates a large scintillation crystal that converts the gamma rays produced by the isotopes into electrical signals. The remainder of the camera apparatus changes these electrical signals into a visual display of the distribution of the radiotracer along with various analyses of the energy spectrum of the gamma rays, which improve image quality by reducing scatter and other artifacts. Further details of gamma camera technology are beyond the scope of this chapter. Much nuclear imaging is planar, i.e. producing two-dimensional images from three-dimensional data, but multiplanar images can also be produced using single photon emission computed tomography (SPECT). These images are produced through the gamma camera rotating around the patient and result in a volume of information which can be sliced in any plane (e.g. sagittal, coronal and transaxial images can all be produced). This markedly improves the detection of smaller abnormalities, especially those located centrally within the body, as well as improving the localization of these abnormalities. It does increase imaging time considerably, and is difficult to perform in certain studies when the count rate is very low (e.g. ^{131}I). Conventional gamma cameras are optimized to image gamma rays with energies of 140 keV – that of ^{99m}Tc – but acceptable images are obtained over a wide range of energies. Positron emission tomography (PET) cameras use a unique imaging system and are described separately later in this chapter.

Nuclear medicine has occupied an important position in the evaluation of patients with suspected or known malignancy since its inception in the 1960s (in fact, ^{131}I was first used in the 1930s). With the advent of PET there has been a resurgence of interest in nuclear techniques and it is likely that PET scans will become as routine a part of the management of oncology patients as CT has. The principle uses of nuclear medicine are for the initial diagnosis of patients (differentiating benign from malignant masses) and for staging and for treatment follow-up. It is used both in isolation and in combination with other imaging techniques. This chapter is divided by scan type rather than disease type due to the considerable overlap of the latter.

Table 1 Commonly used isotopes and tracers in nuclear oncological imaging (isotopes in italics are used for positron emission tomography)

Isotope	Energy (keV)	Half-life	Tracer	Oncological indications
^{99m}Tc	140	6 h	MDP	Bone scanning
			MIBI	Parathyroid, brain and soft-tissue tumours, e.g. breast
			Nofetumomab Merpentan	Small-cell lung cancer
			Arcitumomab	Colorectal cancer
			DMSA (V) (in Europe)	Medullary cell carcinoma thyroid
^{111}In	173	67 h	Octreotide	Neuroendocrine tumours including medullary cell cancer, small-cell lung cancer
			Capromab Pendetide	Prostate cancer
^{201}Tl	80	74 h	Thallium (I) chloride	Brain and musculoskeletal tumours, undifferentiated thyroid tumours
^{131}I	364	8 days	Sodium iodide	Thyroid cancer
			MIBG	Phaeochromocytoma, neuroblastoma
^{67}Ga	93,184,296	68 h	Gallium citrate	Lymphoma
^{18}F	512	110 min	Fluorine	Bone metastases
			FDG	Brain and soft-tissue tumours
			Fluorodopa	Pituitary tumours
			Oestradiol	Breast tumours
			5-FU	Colorectal tumours
^{11}C	512	20 min	Methionine	Brain tumours, other soft-tissue tumours
^{15}O	512	2 min	H_2O	Non-specific tumour bloodflow

BONE SCANNING

Skeletal scintigraphy for the detection of bony metastatic disease is probably the most widely used nuclear medicine procedure in oncology patients.

Basic Principle

In the majority of cases of skeletal metastases, there is a derangement in bone turnover that occurs before a detectable change in the macroscopic anatomy. Up to 50% of bone mass must be lost before changes are seen on plain X-ray films. Bone scanning is performed with isotope labelled phosphate compounds that are accumulated in the skeleton in relation to bone turnover. Most metastases will institute a focal increase in bone turnover due to the tumour-related destruction. Bone scanning is among the most sensitive modalities available for the detection of skeletal lesions.

Indications

The majority of bone scans are performed in patients with known cancer for staging or for follow-up of known or suspected bone lesions. The most common malignancies examined are prostate and breast cancer – two diseases that readily metastasize to bone. In some instances, studies are performed for the evaluation of a focal complaint, such as pain, to determine if malignancy may be the cause. Scans may also be ordered to evaluate primary malignancies of the bone. Because osteosarcoma is a bone-forming tumour, bone scans may also be requested to detect soft tissue metastases in this disease entity.

Procedure

The most commonly used pharmaceutical for bone scanning is methylene diphosphonate (MDP) although several other similar compounds are available. Technetium-99m (^{99m}Tc) is an isotope that is virtually ideal for nuclear scintigraphy and is used as the label for all of these compounds. The agent is administered intravenously in a dose of 25 mCi for adults (200 μCi kg^{-1} for children). Over a period of time, about half of the radiolabelled phosphate adsorbs on bone and calcium while much of the remainder is excreted through the kidneys. Peak target to background (bone to soft tissue) ratio occurs about 3 h after the injection and it is at this time that images are acquired. Most often, scans of the entire skeleton are obtained. Sometimes the study is supplemented by tomographic imaging (SPECT). SPECT may increase the sensitivity for lesion detection in certain areas and also improves the ability to locate the lesion in three dimensions.

Image Interpretation

Multiple foci of increased activity located predominantly in the axial skeleton with little involvement of the distal extremities is the classic appearance of skeletal metastases (**Figures 1** and **2**). The more numerous the lesions, the more highly predictive is the scan. It is possible, of course, for a patient to present with a solitary metastatic lesion. Studies have shown that 10–30% of solitary bone lesions in patients with cancer represent metastases.

Metastases need not be focal but may involve more expansive areas of bone and even lead to diffuse involvement. The 'superscan' occurs with diffusely increased bone turnover (seen with not only metastases but with other skeletal pathology such as hyperparathyroidism) to the extent that there is little remaining MDP to be excreted through the kidneys leading to absence of bladder activity.

Bone scanning is commonly used to follow the course of patients being treated for malignancy. It must be kept in mind that these scans image the bone reaction to the destructive tumour. Shortly after the onset of therapy, it is possible that bone lesions may become more intense owing to the increase in repair. This simulates progression of disease and has been termed the 'flare' phenomenon. A series of scans is necessary for the true evaluation of the course of disease.

Value and Pitfalls

The strength of bone scans is their extremely high sensitivity. It is not unusual for a bone scan to depict metastatic lesions that cannot be detected by plain X-rays. A negative bone scan is a highly accurate means for excluding disease (McNeil, 1984).

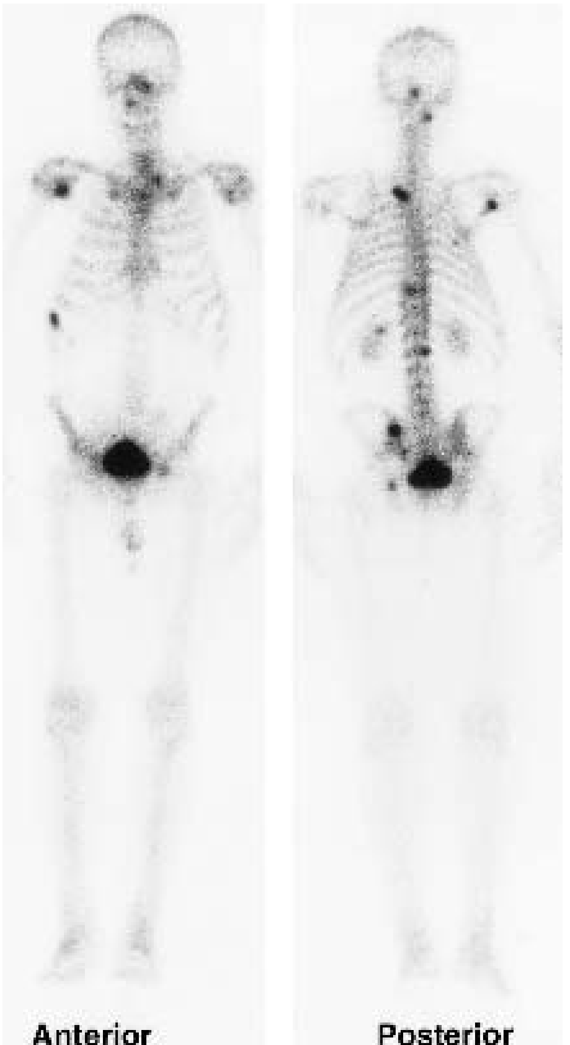


Figure 1 Whole-body bone scan of a 72-year-old male with prostate cancer showing widespread bony metastases.

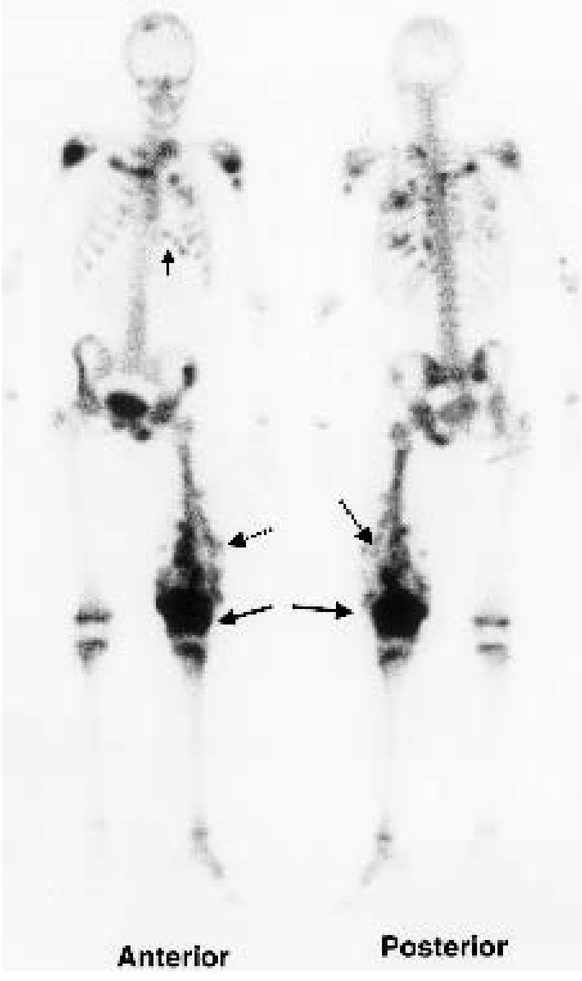


Figure 2 Whole-body bone scan of an 18-year-old boy with primary osteosarcoma of the distal left femur (solid arrow) with multiple bony metastases as well as metastases to his lung (small arrow) and soft tissues of the left thigh (dashed arrow).

As might be expected, the high sensitivity necessitates a less than optimal specificity. Since a wide spectrum of pathology can cause a resultant derangement of bone turnover, lesions that can have the appearance of metastatic disease may be caused by other processes. Probably the most common benign cause of focally increased ^{99m}Tc uptake is degenerative joint disease, a very common problem in the cancer patient age group. When lesions are isolated to the regions of joint spaces (particularly the spine), it is important that the study be interpreted with care. Correlative images such as plain film X-rays, or computed tomography (CT) are frequently used to aid bone scan interpretation. These are used to confirm the presence of a benign aetiology for increased bone turnover such as osteoarthritis, as a negative anatomical study does not exclude bone metastases. Another not uncommon cause of increased focal activity is previous trauma, especially rib fractures. Foci of increased activity at the costo-chondral junction or in a colinear distribution in adjacent ribs is highly suggestive of a traumatic aetiology.

False-negative studies can also occur with bone scans. Highly osteolytic metastases may not have enough of an osteoblastic (bone-forming) response to allow detection by bone scanning. The most common primary tumour to do this is multiple myeloma, wherein it has been estimated that as many as 50% of lesions may not be detectable. For this reason, plain film radiography rather than bone scanning is the modality of choice for staging myeloma patients. Renal cell carcinoma and thyroid cancer also have significantly high false-negative rates.

IODINE SCANNING

Basic Principle

Iodine is taken up by normal and, to a varying extent, neoplastic thyroid tissue as well as by several other organs such as the salivary glands. Radioactive iodine accumulates exactly as nonradioactive iodine does, and hence will show the distribution of iodine-avid tissue in the body. Neoplastic thyroid tissue takes up significantly less iodine than normal tissue, and hence appears as a 'cold' area on a radioiodine scan. A similar appearance is seen on a [^{99m}Tc]pertechnetate scan, which acts as an iodine analogue. Following total thyroidectomy, the neoplastic tissue can be stimulated to accumulate iodine by increasing the TSH (thyroid-stimulating hormone) levels as described below. Thyroid scanning can be performed using several isotopes – [^{99m}Tc]pertechnetate, ^{123}I and ^{131}I. The first two radiotracers are usually restricted to the evaluation of thyroid nodules. The image quality of [^{99m}Tc]pertechnetate and ^{123}I is considerably better than that of ^{131}I owing to the lower energies of the gamma emissions, and the radiation dose to the patient is much less (with the pertechnetate scans having the lowest radiation dose). In patients with known thyroid malignancies the isotope of choice is ^{131}I, which is used both for diagnosis and therapy.

Indications in Oncology

In the evaluation of a patient with a palpable thyroid nodule, either ^{99m}Tc pertechnetate or ^{123}I is used to assess for activity of the nodule. Nodules which are 'warm' or 'hot' (i.e. accumulate the tracer) on pertechnetate or iodine scans are very unlikely to be malignant, whereas those which are 'cold' (i.e. do not accumulate tracer) have a significantly higher likelihood of being malignant – about 10–20% of such nodules (see **Figure 3**). In the post-thyroidectomy patient ^{131}I scanning is used to assess the extent of the thyroid remnant as well as metastatic spread. ^{131}I therapy in much higher doses is then given to ablate the thyroid remnant as well as metastases, if present (Harbert, 1996).

Procedure

For the assessment of a suspicious thyroid nodule, either 300 μCi of ^{123}I is administered orally or 5 mCi of [^{99m}Tc]pertechnetate intravenously, with imaging 24 h or 20 min later, respectively.

Whole-body ^{131}I Scanning

Whole-body ^{131}I scans are only performed after patients have undergone a total thyroidectomy, as otherwise the intense uptake by the patient's normal thyroid tissue will prevent visualization of potential thyroid metastases. Following an optional preablation scan, a therapeutic dose of ^{131}I is administered. The dose given depends on the clinical and radiographic staging of the tumour but is usually in the range 100–250 mCi. The patient has to be admitted and monitored until the patient's radiation emissions are within an acceptable range. In some instances an outpatient dose of >30 mCi may suffice. Seven to ten days following this dose the whole body scan is usually repeated: the larger dose of radioiodine may show the presence of smaller or less iodine-avid metastases. Six to twelve months after thyroid remnant ablation, the diagnostic whole-body ^{131}I scan is repeated to evaluate the success of the ablation as well as for staging. Scans may then be repeated at 6–12-month intervals depending on the clinician preference and the tumour staging as well as the results of other monitoring studies such as the serum thyroglobulin level (e.g. an increase in serum thyroglobulin may stimulate a repeat study). An example of a whole body thyroid scan is shown in **Figure 4**.

To obtain sufficient uptake of ^{131}I into thyroid metastases, a high circulating level of thyroid-stimulating hormone (TSH) is required. As all patients are on exogenous T4 or T3 therapy following total thyroid ablation therapy,

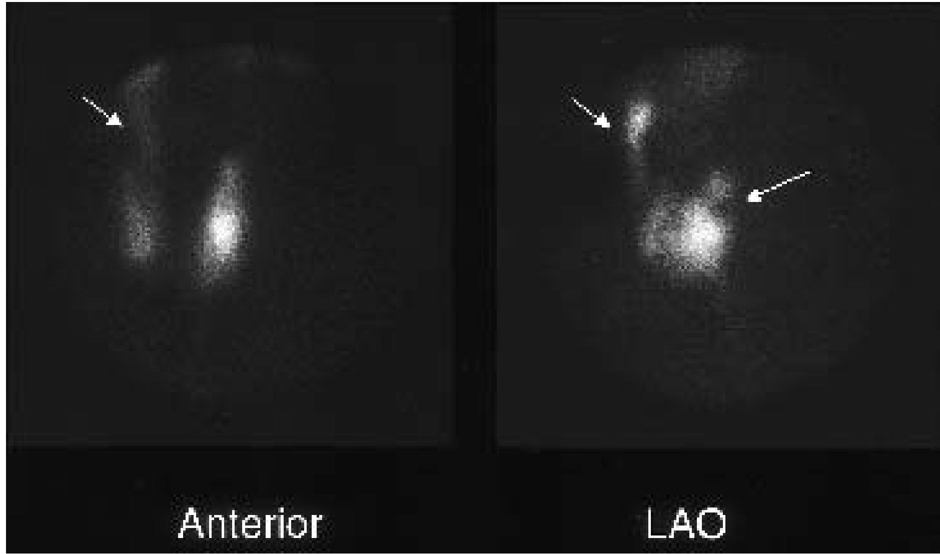


Figure 3 [^{99m}Tc]pertechnetate scan of the thyroid in a 43-year-old woman with a palpable nodule in the left lobe of the thyroid. A cold defect (arrow) is only seen clearly on the left anterior oblique (LAO) view corresponding to the palpable area. Biopsy indicated papillary carcinoma. The dashed arrow indicates a normal pyramidal lobe.

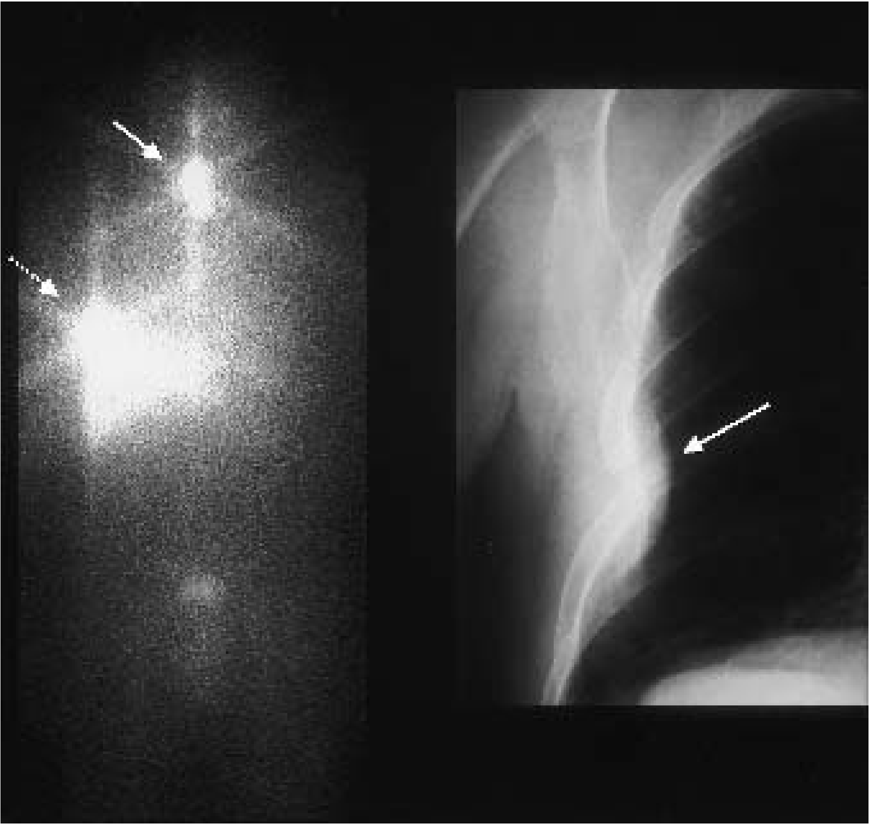


Figure 4 ^{131}I whole-body scan in a 50-year-old male with known follicular carcinoma of the thyroid, 2 years s/p total thyroidectomy and ^{131}I ablation therapy. A focus of intense activity in the neck (solid arrow) represents cervical nodal recurrence. An additional focus in the right lateral thorax (dashed arrow) was due to a rib metastasis, confirmed on the plain film (second image).

their TSH is suppressed. Until recently, therefore, all patients undergoing whole-body ^{131}I scanning for malignancy had withdrawal of T4 for 4–6 weeks or T3 for 2–3 weeks prior to the study, aiming for a TSH level of 30 IU L^{-1} or higher. The US Food and Drug Administration (FDA) has recently approved the administration of recombinant TSH (rTSH) to stimulate ^{131}I uptake in this group of patients. This does not require the withdrawal of thyroid hormone – which is unpleasant and greatly disliked by the patients – with rTSH being given intramuscularly on two consecutive days starting 2 days prior to the ^{131}I administration. The initial trials seem to suggest a slightly reduced sensitivity of this method compared with the withdrawal method, but others feel that in combination with thyroglobulin measurements it has adequate sensitivity (Schlumberger *et al*., 1999).

Image Interpretation

An initial scan following thyroidectomy will show the extent of any thyroid remnant (even after the most meticulous surgical thyroidectomy there is a small amount of normal thyroid tissue remaining) as well as other areas of nonthyroid iodine accumulation. Since iodine is also taken up by several other physiological processes, uptake is seen within the salivary glands, colon, nasal mucosa, urine, sweat and liver (the latter when circulating thyroid hormone is present). No normal uptake above background tissue should be seen outside the thyroid bed in the neck, in the lungs or in the bones, all common sites of metastases. Follow-up scans have several purposes: to demonstrate the success of ablation therapy, to indicate unsuspected metastases, to show the extent of metastases suspected on the basis of increasing thyroglobulin levels (see **Figure 4**), and to indicate the iodine avidity of the metastases for planned radioiodine therapy.

Value and Pitfalls

Most reasonably well-differentiated thyroid malignancies (papillary and follicular) and metastases accumulate radioiodine, but the presence of a large thyroid remnant will reduce sensitivity. Very anaplastic tumours may not accumulate iodine and medullary cell carcinoma is also not iodine avid. In anaplastic tumours thallium or ^{18}FDG PET may be used as alternative imaging methods (see below). Octreotide (see Peptides) as well as pentavalent DMSA (in Europe) have been used to image medullary cell tumours. Contamination of the study by sweat, saliva or nasal secretions is a common cause of a false-positive study. Inflammatory arthritis has also been reported to produce local ^{131}I accumulation. Gut activity has the potential to mask intraabdominal metastases, but these are rare outside the liver.

THALLIUM-201 SCANNING

Basic Principles

^{201}Tl is an isotope with a relatively long half-life of 74 h and low energy emissions (approximately 80 keV). Both of these properties impair the quality of gamma camera images obtained with this isotope. It is taken up into cells via the Na^+/K^+ pump, acting as an analogue of potassium. Because cell viability is vital to the integrity of this pump, thallium will only accumulate in viable cells. It has been extensively used in the past as a marker of myocardial perfusion and viability and, because it also accumulates in increasing amounts in neoplastic tissue relative to normal tissue, it is also used as a nonspecific tumour marker. In this way, it is very similar to the use of sestamibi, which is described below, but operates through a different mechanism. The poorer image quality and higher radiation dose of thallium also limit its usefulness.

Indications

Thallium has been used to image a wide variety of tumours, but most data are available for brain tumours and musculoskeletal tumours.

Brain Tumours

Thallium has been used in two major areas within neuro-oncology – differentiating recurrent tumour from radiation necrosis, and differentiating cerebral lymphoma from toxoplasmosis in AIDS patients. It has also been used as a noninvasive way of grading cerebral gliomas and predicting survival. In this manner it has been used as a ‘poor man’s FDG PET scan’ (see below). With anatomical imaging methods such as CT and MRI, radiation necrosis and tumour recurrence can appear very similar – both as enhancing lesions within the radiation field, frequently surrounding the surgical biopsy site. This is because contrast enhancement (whether with iodine- or gadolinium-based agents) is purely an indicator of blood–brain barrier breakdown, which occurs in both conditions. As thallium is a marker of cell viability, it will be taken up by the recurrent tumours but not by necrotic tissue from radiation necrosis (Slizofski *et al*., 1994). It also accumulates roughly proportionally to the degree of anaplasia of the tumour, but is not specific for tumour types (Black *et al*., 1989; Slizofski *et al*., 1994). The intensity of thallium uptake has been shown to correlate with glioma grading and inversely with prognosis in several studies (Black *et al*., 1989; Kosuda *et al*., 1994). In AIDS patients, both cerebral lymphoma and toxoplasmosis are common (and indeed may co-exist). Differentiating between the two may again be difficult on CT or MRI, but lymphoma is thallium avid, where toxoplasmosis and progressive multifocal leucoencephalopathy are not. A number of studies have

shown high sensitivities (92–100%) and reasonable specificities (80–94%) (Kessler *et al.*, 1998) for differentiating malignancy from infection or inflammation in AIDS patients, potentially reducing the number of cerebral biopsies required. [^{99m}Tc]MIBI has also been used with similar results to assess intracerebral neoplasms. Some difficulty can occur from high physiological uptake by the choroid plexus with [^{99m}Tc]MIBI.

Procedure for Brain Tumour Imaging

No patient preparation is necessary. A 2–3-mCi dose of ^{201}Tl chloride is administered intravenously and imaging commences 30–60 min later with SPECT imaging performed.

Image Interpretation

The images should always be interpreted alongside the results of a current anatomical study (CT or MRI) if available. Images are assessed for the presence of focal areas of increased thallium uptake (see **Figure 5**). The specificity of the study can be improved by obtaining a ratio of the count rate obtained in the abnormal area to a normal area (e.g. contralateral hemisphere in the brain). Cut-off ratios of 1.5–2 have been used to differentiate malignancy from benign processes, with the ratio being roughly proportional to the grade of malignancy. This ratio has also been shown to be inversely proportional to prognosis.

Value and Pitfalls of Thallium Brain Scanning

Thallium scanning provides a valuable way of identifying tumour recurrence, identifying potential areas of higher grade tumour for localizing biopsies and evaluating the response to therapy. It has been shown to be more accurate than CT for assessing recurrence. Low-grade tumours may not accumulate thallium and thus result in a false-negative study, as may very small tumours owing to the inherently poor resolution.

Studies have shown that tumours smaller than 1–1.6 cm have a high rate of false-negative thallium scans. These studies will not differentiate between lymphoma and other intracerebral malignancies such as metastatic disease or gliomas. The results in paediatric brain tumours have been mixed, with some studies showing a reduced sensitivity, although maintaining high specificities.

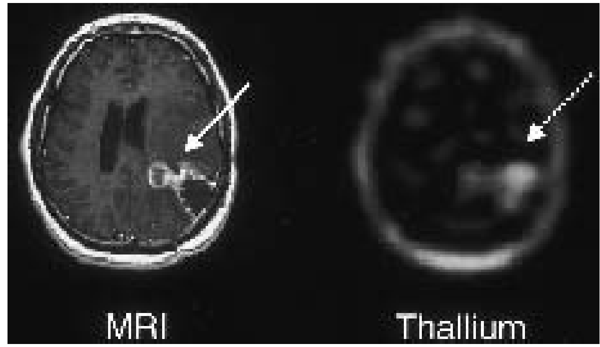


Figure 5 Brain MRI (gadolinium-enhanced T1 axial image) and thallium scan in a 34-year-old man with glioblastoma multiforme previously treated with surgery and radiotherapy. Postradiotherapy rim enhancement was seen at the surgical site (solid arrow). The thallium SPECT scan revealed markedly increased radiotracer accumulation at the same site (dashed arrow) confirming tumour recurrence rather than radiation necrosis. The patient died 4 months later and tumour recurrence was confirmed at autopsy.

Musculoskeletal Tumours

Through its properties as a nonspecific tumour agent, thallium has been used as a noninvasive way to differentiate benign from malignant bone and soft tissue sarcomas and to assess for metastases and recurrence (Sato *et al.*, 1998). While it does have a role to play in these areas in certain patients, e.g. it has been used to evaluate for metastases in osteosarcoma and Kaposi sarcoma with some success, its greatest potential is in assessing response to therapy. Studies in both osteosarcoma and soft tissue sarcomas have shown that changes in thallium uptake in response to preoperative chemotherapy are highly predictive of tumour necrosis percentages at surgery, which in itself is an important prognostic factor. Here, thallium is acting again as a viability marker, and a mid-course thallium scan is predictive of the final postchemotherapy response (Sumiya *et al.*, 1998). A less than 20% decrease in the thallium index (tumour-to-contralateral normal bone/soft tissue uptake) indicates a minimal response to chemotherapy. This is much more helpful than [^{99m}Tc]MDP bone scans in the same situation.

Procedure for Musculoskeletal Tumour Imaging

No patient preparation is necessary. A 2–3-mCi dose of ^{201}Tl chloride is administered intravenously and imaging commences 30–60 min later. SPECT imaging is optional, but patients benefit from SPECT imaging of the local tumour site.

Image Interpretation

In a similar way to brain tumours, the activity in the tumour can be compared with the contralateral limb or body region. This semiquantitative information is vital for assessing treatment response, and may be of some help in differentiating benign from malignant lesions, although there are limitations as indicated below.

Pitfalls of Thallium Musculoskeletal Imaging

Unfortunately, the significant incidence of both false-positive and -negative studies when attempting to use

thallium to differentiate benign from malignant musculoskeletal processes limits its use in this regard. Thallium accumulation has been reported in histiocytosis X, giant cell tumours, stress fractures and myositis ossificans. Lack of uptake has been reported in liposarcomas and myxoid tumours.

Other Potential Indications

Thallium has been investigated in several other soft tissue tumours including Kaposi sarcoma (where it is useful in combination with gallium scanning in distinguishing between intrathoracic lymphoma, Kaposi sarcoma and infection) and thyroid cancer. It is less sensitive than ^{131}I in detecting thyroid metastases in well-differentiated thyroid cancer, but has a potential role in the ^{131}I-negative, thyroglobulin-positive patient (i.e. less well differentiated thyroid malignancies). In medullary cell carcinoma of the thyroid, both pentavalent ^{99m}Tc–DMSA and octreotide appear more sensitive.

GALLIUM SCANNING

Basic Principles

Gallium-67 citrate is an agent that has been used for bone scanning and the detection of areas of inflammation or infection, but its greatest role has come in the field of oncology. Many tumours are known to accumulate gallium but it has been used most commonly in cases of lymphoma.

The exact reason for the concentration of gallium in various tumours is not clearly defined. Gallium is an analogue of iron and is bound in the plasma to transferrin and lactoferrin. It has been suggested that gallium binds to intratumoural ferritin stores.

Indications

Most lymphomas do accumulate gallium, to a greater extent in Hodgkin disease than in non-Hodgkin lymphoma (Turner *et al*., 1978). In these instances, gallium scans can play two roles: detection of sites of disease for staging purposes and follow-up after therapy. Since gallium scans have poor anatomical resolution, interpretation of scans is aided by a correlative study such as CT.

It is possible for gallium scans to detect sites of disease not seen by CT. This is especially true in the abdomen, where the presence of bowel can be limiting, as well as in the spleen. A baseline gallium scan is also helpful in ascertaining that the tumour in a patient who will be followed is initially gallium avid. The greatest value of gallium comes in its role in the follow-up of disease. Soft-tissue masses detected by CT may remain after treatment even when all active disease has been eradicated. Gallium is more highly specific: only active lymphoma itself will accumulate the radiotracer.

Procedure

It is now standard technique to use a higher dose of gallium for the detection of tumour than was previously used in cases of infection. This is called the ‘high-dose’ gallium scan and involves the intravenous administration of 10 mCi of ^{67}Ga citrate. Gallium has a physical half-life that is longer than that of technetium (78 h) and imaging is carried out over a period of several days. Initial images are typically acquired about 2–4 days after the administration of the dose. Since normal bowel excretion can complicate the interpretation of the images of the abdomen, delayed images are often obtained. Bowel contents will change in position and fade whereas tumour will retain its configuration and if it changes at all will probably show an increase in the target-to-background ratio. Many sites also perform tomographic (SPECT) imaging of the chest, abdomen and pelvis. The use of laxatives to clear colonic activity is somewhat controversial. It has been suggested that laxatives may inflame the colon and cause a resultant increase in bowel activity.

Image Interpretation

The normal distribution of gallium differs somewhat from patient to patient but, in general, activity will normally be seen in the liver, skeleton (and bone marrow), spleen, salivary glands, lacrimal glands and breast tissue in females. As mentioned above, gallium is excreted in the bowel giving rise to variable intestinal activity. Gallium is excreted through the kidneys during the first 24 h; renal and bladder activity is minimal by 72 h.

In patients with lymphoma an abnormal study is manifested by increased radiotracer activity at sites of adenopathy. It is common to see disease in the mediastinum, hilar, neck, retroperitoneum and pelvis (see **Figure 6**). Lymphoma may also involve other sites, including the spleen, bones and kidneys.

Pitfalls

Diffuse uptake within the lungs themselves may be due to lymphoma but gallium also accumulates in areas of inflammation. Diffuse uptake, therefore, may also be due to pulmonary toxicity from prior chemotherapy or to a superimposed pneumonia in an immunocompromised patient. Normal physiological uptake of gallium in the hilar is common. It is faint and symmetric and is not accompanied by mediastinal activity or corresponding abnormalities on a CT scan. Bowel uptake mimicking intraabdominal adenopathy has already been described and can be reduced by delayed imaging.

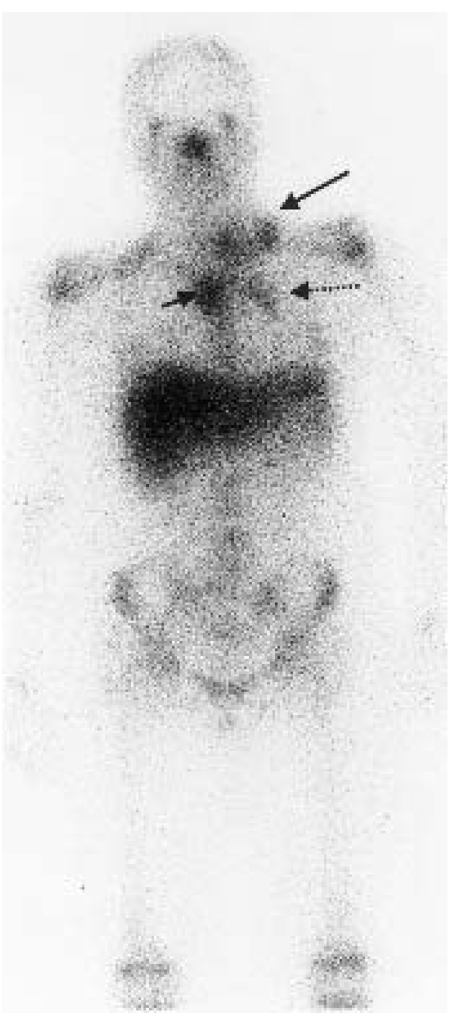


Figure 6 Gallium-67 scan in a 19-year-old woman with newly diagnosed Hodgkin lymphoma showing intense gallium accumulation in the left supraclavicular nodes (solid arrow), mediastinum (small arrow) and left hilum (dashed arrow). These findings were confirmed on the SPECT images (not shown).

Other Roles

Areas of increased activity can also be seen in other forms of tumour. Gallium is avidly taken up at sites of melanoma and has been used in the diagnosis of hepatocellular carcinoma.

MONOCLONAL ANTIBODIES

Within the past several years, several monoclonal antibodies with radioisotope labels, long investigated for the detection of tumour, have received FDA approval and become commercially available in the United States. Now in use are satumomab pendetide for colorectal and ovarian carcinoma (Oncoscint from Cytogen), capromab pendetide for prostate carcinoma (Prostascint from Cytogen), nofetumomab merpentan for small cell lung carcinoma (Verluma from DuPont) and arcitumomab for colorectal carcinoma (CEA-scan from Immunomedics).

Basic Principles

Each of these monoclonal antibodies will bind to an antigen that is found in fairly high abundance on the tumours in question. These antibodies are produced in cell cultures from murine (mouse) lymphocytes that are combined (hybridized) with myeloma cells. Specific cell lines are then screened, isolated and the antibody-producing cells are cloned. The antibodies are then radiolabelled with indium-111 (Oncoscint and Prostascint) or with the superior imaging agent technetium-99m (Verluma and CEA-scan). Scanning with these agents is generally used in situations wherein there is the suspicion of tumour recurrence or metastases (e.g. an elevated CEA or PSA) but other imaging studies (e.g. CT or bone scanning) have failed to detect disease.

Imaging Protocol

Imaging acquisitions differ slightly for each of the agents utilized. The indium-labelled agents are imaged starting 2–4 days after the dose administration. Owing to variable and potentially confusing bowel excretion, the studies are coupled with bowel preparation with laxatives and possibly delayed sets of imaging. It is typical to image Prostascint along with a blood pool label (technetium-99m) in a dual window setting to separate out normal Prostascint activity within the vasculature. Technetium labels provide a higher count rate and, therefore, can be imaged earlier. CEA-scan is performed at 4 h and Verluma at 16 h postinjection.

Scan Interpretation

Areas of metastatic or recurrent prostate cancer in Prostascint scans usually appear as focal abnormalities, most often in the prostate bed or pelvic lymph nodes (see **Figure 7**). Great care must be taken to separate normal from abnormal activity. As mentioned above, normal activity within bowel and blood pool is common. It is also normal to see activity in the liver, spleen, kidneys and bladder as well as bone marrow. Faint uptake in a normal prostate does occur. Prior irradiation may also be a cause of increased activity. Correlation with other imaging studies, such as CT, is often very helpful for delineation of normal anatomy. Other antibodies have a similar appearance; all have relatively high background activity.

Accuracy of Imaging

At present, monoclonal antibody imaging is not indicated as a first-line imaging procedure for diagnosis or staging.

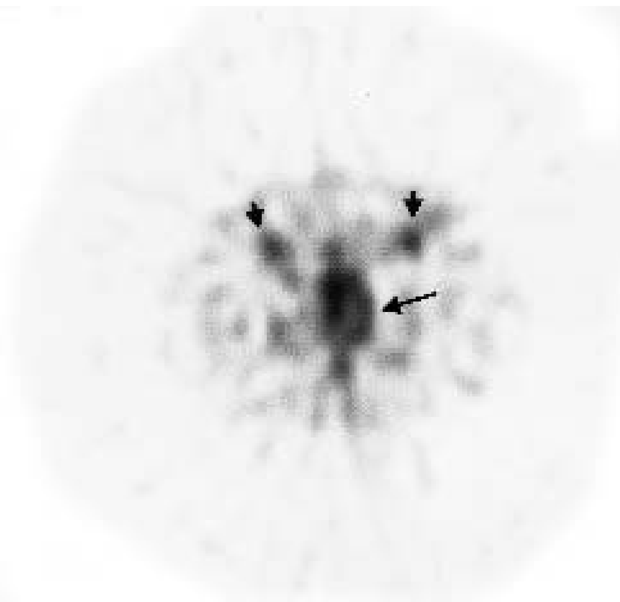


Figure 7 [^{111}In]Prostascint (capromab pendetide) SPECT scan in a 68-year-old man with a history of prostate cancer s/p radical prostatectomy with a rising PSA. Transaxial image through the pelvis shows evidence of recurrence in the prostate bed (solid arrow). The two anterior foci of radiotracer uptake (short arrows) reflect blood pool activity within the illiac vessels.

Because these agents are typically used to detect occult disease not uncovered by other modalities, the sensitivity is not extremely high. The sensitivity of Oncoscint for detecting recurrent colon carcinoma is reportedly 75% in the pelvis and 67% in the extrahepatic abdomen and for recurrent ovarian carcinoma 59%. The accuracy of Prostascint for the detection of pelvic lymph node metastases in presurgical patients has been reported to be 76% (Babaian *et al*., 1994). This is still considerably higher than with CT. With care taken to exclude false-positive studies, especially by paying attention to bowel and blood pool activity, the specificity of these studies can be fairly good. Imaging with monoclonal antibodies has been shown, in multiple studies, to be of value in the detection of occult disease (Manyk, 1998).

Precautions

Monoclonal antibodies are proteins that are, technically speaking, foreign to the human body and, as such, can cause allergic reactions. Serious side effects are few but these agents need to be administered with care. Also, because of the nature of these proteins, patients may form antibodies against these agents. These are known as human anti-mouse antibodies (the 'HAMA' reaction). This reaction occurs with variable frequency but the significance of the production of HAMA is that it may lead to reduced sensitivity of subsequent studies performed with murine antibodies.

Future Directions: Therapy

By labelling monoclonal antibodies with potentially β-emitting isotopes such as iodine-131, it is possible that these agents can be useful therapeutically. The antibody will theoretically deliver a high dose of radiation directly to the tumour itself and avoid the side effects that occur with general systemic treatment. Owing to the variable uptake of the agents that exist today, this science should be considered to still be in the early stages of development. With present antibodies, the radiation dose to the bone marrow is frequently the dose-limiting factor.

PEPTIDES

Peptides are small compounds composed of sequences of amino acids. Only one peptide has been in common use in nuclear medicine, pentetreotide or Octreoscan (Mallinckrodt).

Basic Principles

Pentetreotide is an analogue of another peptide, somatostatin. Receptors to somatostatin are found in abundance on many neuroendocrine tumours, particularly gastrinoma, glucagonoma, small cell lung cancer and carcinoid. Somatostatin receptors are also frequently found on other tumours, including lymphoma, medullary cell carcinoma of the thyroid, breast cancer and meningioma.

Imaging Protocol

Pentetreotide is labelled with indium-111, and is administered intravenously. Images are acquired 4 and 24 h and sometimes 48 h later. SPECT imaging of the region of interest is normally performed. There is intense uptake in kidneys and spleen and normal activity is also seen in the liver and, in varying amounts, in the bowel.

Indications

Indications for performing of [^{111}In]pentetreotide imaging can be divided into several categories. First, patients with a known neuroendocrine tumour can be scanned for the presence of metastases. The converse can also be applied: patients with a known metastasis can be scanned in search for the primary lesion. Studies have shown that of [^{111}In]pentetreotide can add information beyond CT and MRI (Olsen *et al*., 1995). Finally, scans are performed in patients with known disease to determine the receptor status of a tumour (**Figure 8**). Somatostatin receptors

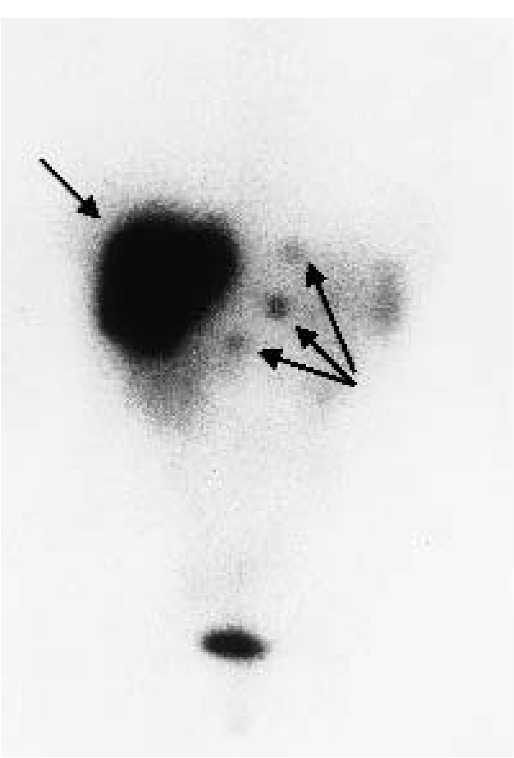


Figure 8 [^{111}In]pentetreotide scan in a 40-year-old woman with metastatic carcinoid showing multiple hepatic metastases (arrows).

occur in amounts that may vary from patient to patient. If uptake of [^{111}In]pentetreotide is confirmed with the imaging study, the patient can be considered a candidate for therapy with nonradioactive octreotide acetate, an inhibitor to neuroendocrine tumour growth.

Future Prospects

Other peptides are in development and play roles in tumour imaging. One of them, depreotide (Neotect, Diatide) has recently received approval from the FDA for small cell lung cancer. This agent is also a somatostatin analogue but is coupled to technetium, an isotope with imaging characteristics that are superior to those of indium.

As was mentioned with the monoclonal antibodies, there are potential therapeutic capabilities with peptides. The peptides may be labeled with β-emitters and studies are under way involving the administration of high-dose Octreoscan to evaluate tumour treatment with emitted conversion electrons.

SCINTIMAMMOGRAPHY

In recent years, scintigraphic techniques have been devised for the imaging of primary sites of breast cancer. Known most commonly as scintimammography, [^{99m}Tc]sestamibi has been the agent most frequently used for this purpose.

Principles and Indications

The mechanism for sestamibi concentration by breast carcinoma cells (and several other tumours) is not well understood. It is possible that the agent is localized in the intracellular mitochondria of metabolically active tumour cells. Nevertheless, uptake tends to be a good predictor of malignancy.

Scintimammography is not indicated as the first-line imaging study in patients with suspected breast cancer but is used in complicated situations where additional information is needed above what has been determined by conventional mammography or ultrasound. Indications include additional workup of the indeterminant mammographic lesion or of a palpable mass not detected by mammography, as may occur in patients with dense breasts.

Imaging Protocol

Sestamibi is administered intravenously in a dose of 20 mCi into the arm contralateral to the suspected lesion. This is done to avoid potentially confusing lymph node uptake that may occur normally if some of the injected dose is infiltrated in the arm. If there are suspicious lesions in both breasts, another site, such as the foot, should be used. Imaging commences about 10 min later and is typically performed in the anterior and lateral projections. Lateral images are improved with the patient positioned prone on a customized scintimammographic pad. This is a device with a cut-out that allows the breast to hang dependently away from the chest wall. Tomographic imaging may also be performed but its value has not been defined.

Scan Appearance and Utility

Uptake of the tracer within malignancy typically appears as a focal area of increased activity with a substantial target to background ratio (**Figure 9**). It is possible to find multiple lesions in addition to uptake within axillary adenopathy. Scintimammography has several advantages over conventional mammography with considerably higher specificity (89% compared with around 30% for mammography) (Khalkhali *et al*., 1995). In addition, imaging is not hindered by dense breasts.

Pitfalls

The greatest limitation of scintimammography is that it utilizes a technology with a relatively low spatial resolution. Lesions under 1 cm in size are difficult to visualize. For lesions >1 cm the sensitivity is >92%. False-positive uptake occurs in some benign lesions such as cellular fibroadenomas and active inflammation.

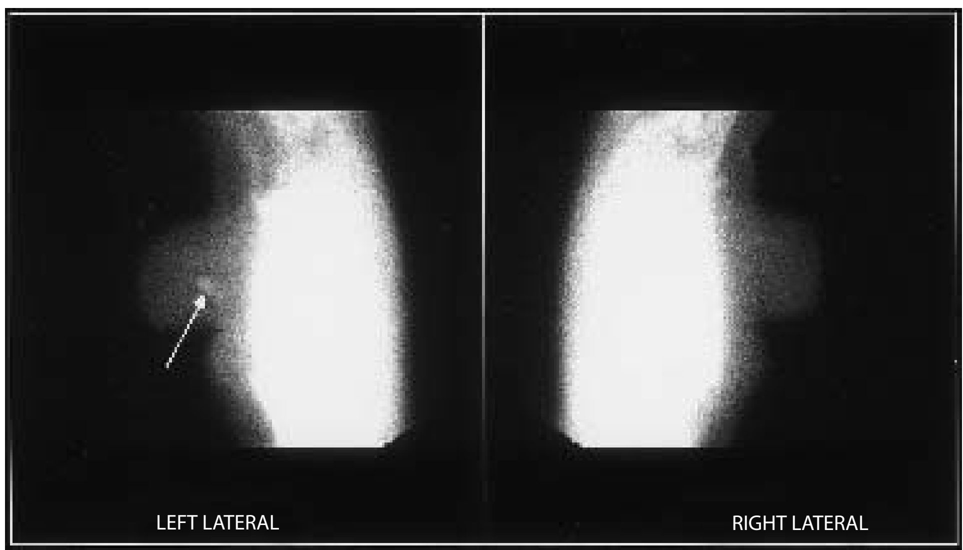


Figure 9 Scintimammogram using [^{99m}Tc]MIBI in a 55-year-old woman with an indeterminate left breast mass on mammogram and ultrasound showing intense focal uptake of tracer (arrow). Invasive ductal carcinoma was confirmed at biopsy.

MIBG

MIBG or Meta-iodobenzylguanidine is a pharmaceutical that has been utilized for a number of years for imaging phaeochromocytoma and neuroblastoma. It has only recently, however, received FDA approval and is now commercially available.

Principles

MIBG is an analogue of norepinephrine that is labelled with the iodine isotope ^{123}I or ^{131}I (although only the ^{131}I-labelled MIBG is produced commercially). MIBG is concentrated by the adrenal medulla and in adrenergic tumours, most significantly phaeochromocytoma and neuroblastoma. Most of the uptake of MIBG occurs as an active process with localization of the molecule in catecholamine hormonal storage vesicles. Iodine-131 emits β-particles and high-energy γ-rays. It has a long half-life and can only be utilized in the relatively low dose of 0.5–1.0 mCi. This results in a relatively low count study. Sites that are approved for use of the ^{123}I label find better image quality owing to the higher administered dose (^{123}I has no β-emissions and more ideal γ-energies and a shorter half-life than ^{131}I).

Indications

MIBG is used in patients who are suspected to have a phaeochromocytoma based on biochemical analysis but no detectable tumour in other imaging studies (e.g. CT, MR). This agent has the advantage of easily allowing scanning of the entire body. It should be kept in mind that 10% of phaeochromocytomas can occur in an extra-adrenal location. It can also be used in patients with known tumour for the detection of sites of metastases (10% are malignant) or in a search for multiple primary sites (as may occur in the multiple endocrine neoplasia syndrome). In patients with neuroblastoma, MIBG also may be used for the detection of metastases. MIBG also plays a role in treatment planning. Patients may be candidates for a therapeutic trial of MIBG if the tumour can be visualized with a diagnostic scan (see below).

Imaging Protocol

The exact protocol depends on whether ^{131}I or ^{123}I is utilized. The ^{131}I-labelled agent will be discussed. The pharmaceutical is administered intravenously with a slow push (to avoid the development of acute hypertension). Images are acquired 48 h later and should be tailored to the areas of concern. Planar images of the entire body can be obtained and it is possible to perform SPECT (although this is vastly superior with the higher dose of ^{123}I) of the chest or abdomen as needed. Patient preparation is needed for this examination. A number of medications can inhibit the uptake of MIBG by phaeochromocytoma and must be discontinued. These include Labetalol, calcium channel blockers and phenylephrine. It is also necessary to pretreat the patient with potassium iodide (oral SSKI or Lugol's solution) in order to block thyroid uptake of radioactive iodine and hence to reduce the radiation dose to the thyroid.

Accuracy and Limitations

Phaeochromocytoma (or neuroblastoma) will appear as a focal area of abnormal activity. Sensitivities for the detection of phaeochromocytomas with MIBG have been

estimated to be about 85%. Care must be taken to avoid confusing adrenal tumour with excreted activity in the renal pelvis. To clarify the position of the kidney relative to the suspected abnormal activity, a dose of a renal agent such as [^{99m}Tc]DTPA can be administered when the MIBG imaging is complete. Phaeochromocytoma can also occur in or near the bladder. It may be necessary to place a bladder catheter and flush the bladder to remove normally excreted activity (McEwan *et al*., 1985).

Future Directions

The use of MIBG for therapy remains in the early stages of development. Because this is an iodinated pharmaceutical, the β-emissions from ^{131}I can be used to administer a high dose of local radiation when administered in large amounts. MIBG has been used for the treatment of neuroblastoma in children.

POSITRON EMISSION TOMOGRAPHY

General Principles

Positron emission tomography uses both different isotopes and different camera systems to image the distribution of radioisotopes within the body, but as with conventional nuclear medicine techniques it is a functional imaging method. While most single photon-emitting isotopes such as ^{99m}Tc decay through the emission of a single γ-ray, positron-emitting isotopes decay through the emission of a positron of variable energy. This positron travels 1–5 mm through the tissues, the distance depending on the energy of the positron before colliding with a nearby outer-shell electron. This 'annihilation reaction' converts matter to energy and results in the production of two 512-keV γ-rays at almost exactly 180° to each other. It is these paired γ-rays which are detected by the PET camera, which consists of rings of multiple scintillation crystals (each like a tiny gamma camera only using different crystal technology) linked in pairs 180° across from each other. All PET images are multiplanar (in a similar way to SPECT). The physics of PET result in superior image resolution compared with SPECT and the production of the positron-emitting radiotracers is relatively easy owing to the smaller atom size. Many of the PET isotopes are radioactive isotopes of atoms normally present in many physiological compounds such as carbon, nitrogen and oxygen. This means that most PET radiotracers are much closer in composition to normal physiological substrates. On the other hand, ^{18}F, a commonly used isotope, can substitute for hydroxyl ions ($^{-}$OH). These factors enable PET to image a wide variety of physiological processes such as blood flow, metabolism (glucose, amino acid, fatty acid, protein, etc.) and receptor density *in vivo* with very good anatomical resolution. The oncological potential is extensive and only recently has become realized.

There are two major problems that have prevented the wide availability of PET. The first is the expense of the camera system, at $1–2 million at 2000 prices. The second is that most PET isotopes have to be produced by a cyclotron, an even more expensive piece of machinery with high upkeep demands in both manpower and cost. The half-life of PET isotopes is much shorter than those of single-photon isotopes: whereas the ^{18}F half-life is 110 min, ^{11}C is only 20 min, ^{13}N 10 min and ^{15}O only 2 min. This significantly limits the distribution of these isotopes from central cyclotrons. Recently, two advances have led to the much wider availability of PET. First, conventional SPECT-enabled gamma cameras can be upgraded at reasonable cost to detect the paired 512-keV γ-rays (so called 'coincidence detection') although with lower but acceptable spatial resolution than dedicated PET cameras. Second, several networks of cyclotrons have been set up to distribute commercially [^{18}F]Fluorodeoxyglucose (^{18}FDG), the most commonly utilized PET radiotracer. These two factors are currently leading to an explosion of PET across many countries in the world. It is likely that within a few years, ^{18}FDG PET will become an essential component of oncological patient management in most oncology centres. For the purposes of the following sections, dedicated PET and coincidence detection will be discussed as a single entity. Increasingly the technical advances in coincidence detection are bringing its accuracy closer and closer to that of dedicated PET.

In oncology, most clinical PET imaging uses ^{18}FDG, with other isotopes being mainly investigational and mentioned briefly in a later section. Imaging can be of a local region (e.g. the head or chest) or of the entire body in one session.

^{18}[F]Fluorodeoxyglucose Imaging

Principles of ^{18}FDG Imaging

[^{18}F]Fluorodeoxyglucose is an analogue of glucose. It enters the cell at rate proportional to that of glucose, using the same glucose transporter system. It also undergoes the first step in glycolysis to produce ^{18}FDG-6-phosphate (via hexakinase) but is not a substrate for the next enzyme in the cascade (glucose-6-phosphatase). ^{18}FDG-6-phosphate does not readily cross the cell membrane and hence remains trapped in the cell. The local concentration of ^{18}FDG-6-phosphate as recorded by the radioactivity 30+ min after ^{18}FDG injection is therefore a measure of the local glycolytic rate. This method has advantages over another PET method for measuring glycolysis, with [^{11}C]glucose, as the latter, by passing readily through glycolysis and beyond, results in imaging of a combination of glucose uptake, metabolism and metabolite distribution.

For decades, glycolytic rates have been known to increase in neoplasms and for many neoplasms this rate is proportional to the degree of malignancy (anaplasia) of the tumours. In benign processes, the glycolytic rates tend to be low. ^{18}FDG imaging therefore has the potential to distinguish between benign and malignant processes, grade tumours, identify metastases and diagnose tumour recurrence. ^{18}FDG is a very nonspecific tumour imaging agent, and it has been used in practically every tumour type for one or more of the above indications although currently only a handful are in routine clinical usage (typically in the USA those covered by Medicare). As of 2000, these include differentiating benign from malignant lung nodules, staging of non-small cell lung cancer, lymphoma staging, melanoma staging, oesophageal cancer staging, head and neck cancer and diagnosing colorectal tumour recurrences. These and some of the other more common indications not currently covered by Medicare are described below.

Solitary Pulmonary Nodule (Coleman, 1999)

Previously, CT and plain films along with biopsy have been the mainstays for solitary pulmonary nodule evaluation. ^{18}FDG PET has an important role to play, however. Malignant nodules have increased rates of glycolysis, and hence increased ^{18}FDG uptake relative to benign processes such as hamartomas and granulomas. The sensitivity of PET is 90–100%, with a specificity of 75–95% for differentiating benign from malignant nodules. Nodules which are positive on ^{18}FDG scanning should therefore be biopsied; those which are negative are generally safe to follow. By following this protocol, the number of potentially harmful and expensive lung biopsies can be reduced significantly. Some centres incorporate a semiquantitative measure of ^{18}FDG uptake into their evaluation – the standardized uptake ratio (SUR), which estimates the amount of the injected radiotracer taken up into the tumour relative to the injected dose and patient's body mass. This may help differentiate further between benign and malignant processes, although not all authorities agree. Several studies have indicated that the SUR is inversely proportional to survival times in primary lung cancer. False-positive scans do occur occasionally in inflammatory processes in particular and false-negative scans have been reported in slow-growing carcinoid and occasionally bronchoalveolar tumours. Continued monitoring over time of ^{18}FDG PET-negative tumours by CT or CXR should identify these patients.

Lung Cancer Staging (Coleman, 1999; Kalff *et al.*, 2001)

The identification of lymph node and distant metastases is a major component of the staging of non-small cell lung cancer. This staging determines the patient's potential operability. Both CT and MRI rely on the size of the lymph nodes as a criterion for malignancy, with a 1-cm short axis dimension conventionally being used as the cut-off for potential malignancy. Obviously smaller nodes than these may be malignant, and larger nodes benign. The sensitivity and specificity of CT for staging lung cancer have been reported to be as low as 50–70%. This leads to frequent mediastinoscopies and even patients being found to be inoperable (or at least incurable) at the time of surgery. By imaging function rather than anatomy, ^{18}FDG PET significantly adds to the preoperative work-up of these patients with accuracies of 80–100%, considerably superior to CT or MRI. PET has been found to change lung cancer staging in up to 30% of patients – even after CT scanning has been performed (**Figure 10**). The ability to image the whole body allows the additional identification of unsuspected distant metastases including evaluating the adrenal glands, a common site of lung cancer metastases. In this latter area PET has been reported to have a 100% sensitivity and 80% specificity for metastatic involvement. The addition of ^{18}FDG PET to the routine preoperative evaluation of patients with lung cancer has been shown to result in significant cost savings – calculated as $98 million per year across the USA if all potentially resectable patients underwent PET scans. ^{18}FDG PET has also been used to evaluate for local tumour recurrence following surgery and/or radiotherapy or chemotherapy. These patients are typically very difficult to assess by CT owing to the considerable postoperative or postradiotherapy changes. Focally increased ^{18}FDG uptake in these patients is indicative of tumour recurrence with extremely high accuracy.

Lymphoma Staging (Delbeke, 1999b)

Although ^{67}Ga is currently the major functional imaging method for the staging of lymphoma, ^{18}FDG PET has some significant advantages. The resolution is significantly better, not all lymphomas are gallium avid whereas most are ^{18}FDG avid and the imaging can be performed on a single day rather than stretched out over several days as with gallium (Moog *et al.*, 1997). The uptake of ^{18}FDG is proportional to the grade of lymphoma, and hence inversely proportional to prognosis. Bone marrow involvement will show on the PET scan better than on a gallium scan. Very low-grade lymphomas may be ^{18}FDG negative. In the posttherapy patient, the advantages become more marked with ^{18}FDG PET predicting the response to treatment, even in the presence of residual masses on CT. Of importance, ^{18}FDG PET like gallium cannot be used to differentiate lymphoma from sarcoidosis (a not uncommon problem), as both are ^{18}FDG avid.

Melanoma Staging (Delbeke, 1999b)

Melanoma is a highly malignant tumour that may metastasize to any organ of the body. For this reason, staging and monitoring these patients is difficult, frequently requiring several different imaging modalities and metastases to unexpected areas such as the bowel and extremity subcutaneous tissues may frequently be missed. Owing to

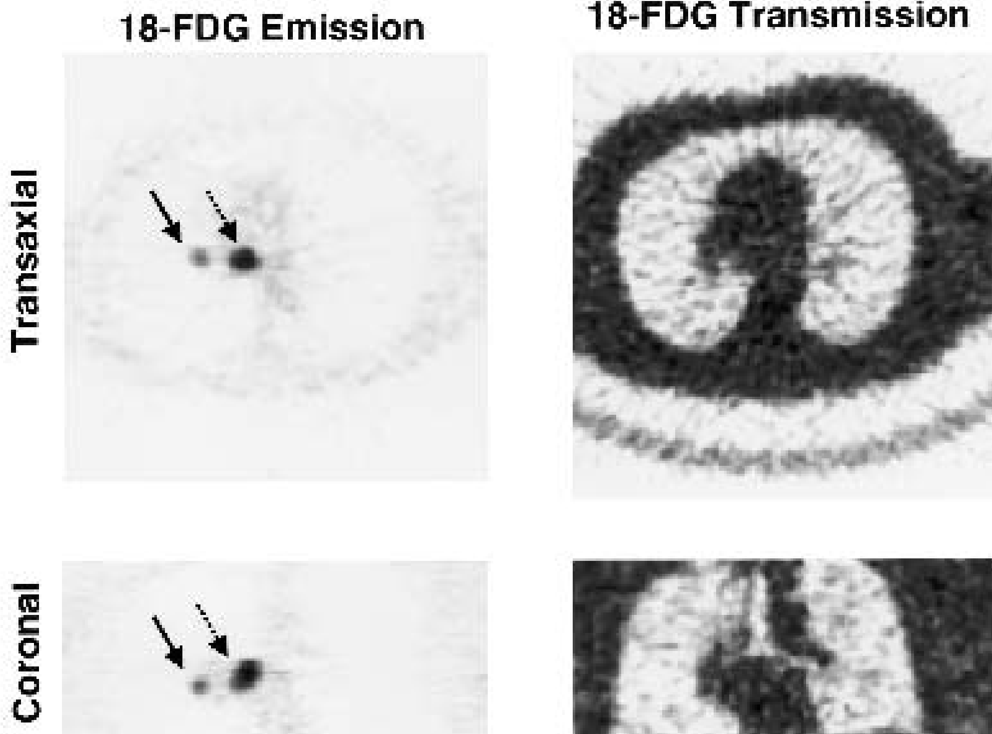


Figure 10 Transaxial ^{18}FDG emission scan through the thorax at two levels with corresponding transmission scan at the same level to show localization in a 62-year-old man with known lung cancer (primary tumour not shown). Focal uptake of ^{18}FDG in the mediastinum (dashed arrow) and right hilum (solid arrow) were due to nodal metastases confirmed at surgery. (Images courtesy of the Clinical PET Centre, United Medical Schools of Guy's, St. Thomas' and King's Hospitals, London, UK.)

its high degree of malignancy, even small melanoma metastases are typically intensely ^{18}FDG avid and the whole-body imaging capability of PET enables a single scan to assess for metastases literally from head to toe. ^{18}FDG PET scanning has been shown to be an accurate (88–100%) and cost-effective method of staging melanoma (Rinne *et al.*, 1998).

Colorectal Tumour Recurrence (Delbeke et al., *1997)*

The postoperative evaluation of colorectal tumours by CT is hampered particularly in the surgical site by postsurgical and/or radiotherapy changes. Small hepatic metastases may also be missed. Serum markers such as CEA may indicate a recurrence but these are nonspecific and nonlocalizing. Monoclonal studies such as CEA scan (described above) have limited sensitivity and specificity for detection owing to high background activity and poor resolution. ^{18}FDG, on the other hand, is not taken up by postsurgical and chronic radiation changes, and has lower background activity and higher resolution than monoclonal studies. Activity within the bladder and normal bowel activity must, however, be carefully distinguished from tumour recurrence. ^{18}FDG PET has been shown to be highly accurate (92%) for diagnosing and localizing local and distant colorectal tumour recurrence, including hepatic metastases, changing surgical management in 28% of patients after CT (Delbeke *et al.*, 1997). In patients with negative CT scans but rising CEA levels, PET identified the site of metastases in 87% of patients (Maldonado *et al.*, 2000). Scans performed within 6 months of radiotherapy may result in false-positive results owing to local inflammation. The response of hepatic metastases to chemotherapy can be predicted by ^{18}FDG PET scans 4–5 weeks into treatment.

Brain Tumours (Delbeke, 1999b)

Some of the earliest ^{18}FDG PET studies investigated brain tumour metabolism. The uptake of ^{18}FDG was found to be highly correlated with tumour grade (Jolesz, 1983), and in fact a better marker of prognosis than histology. ^{18}FDG PET has therefore been used as a prognostic marker, as a method of indicating the areas of highest malignant potential (hence highest ^{18}FDG uptake) for biopsy and for diagnosing tumour recurrence. The primary energy source for the brain is glucose, and this leads to high physiological cortical uptake of ^{18}FDG. Tumours which have lower metabolic activity than the surrounding cortex (such as grade 1 and 2 gliomas) may be difficult to distinguish from surrounding cortex. Despite this, ^{18}FDG PET has an

important role to play in differentiating tumour recurrence from radiation necrosis in a similar manner to thallium, particularly in the higher grade tumours. ^{18}FDG PET has been used to differentiate lymphoma from toxoplasmosis in AIDS patients, again similar to thallium with very comparable results.

Breast Cancer (Delbeke, 1999a)

^{18}FDG PET has been used for the diagnosis, staging and evaluation of recurrence in breast cancer. The sensitivity and specificity for primary diagnosis ranges from 80–100%, with false negatives occurring in <1-cm tumours and slow growing tumours (Avril *et al*., 1996). It may be useful in patients with dense breasts, after augmentation mammoplasty and those with equivocal mammography. The sensitivity for axillary lymph node metastases is around 95%, but it will miss micrometastases so the role here is less clear (**Figure 11**). It will, however, indicate the presence of internal mammary and distant metastatic disease. Osteolytic bone metastases are more easily detected on ^{18}FDG PET than with routine bone scanning.

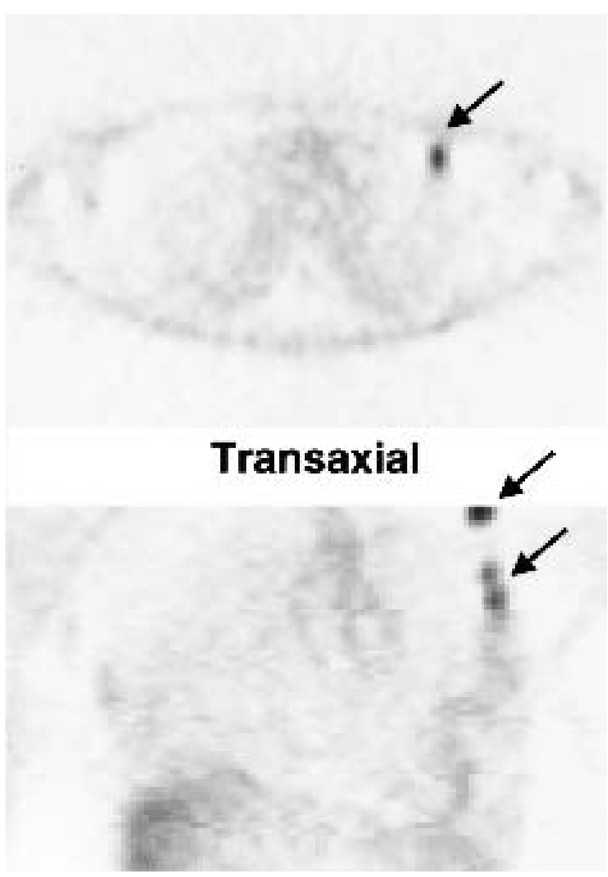


Figure 11 Transaxial and coronal ^{18}FDG emission scan through the thorax in a 58-year-old woman with left breast cancer. Multiple foci of left axillary radiotracer uptake (arrows) were due to axillary metastases. (Images courtesy of the Clinical PET Centre, United Medical Schools of Guy's, St. Thomas' and King's Hospitals, London, UK.)

^{18}FDG PET also has a potential role in the evaluation of breast cancer recurrence, particularly in the brachial plexus region where anatomical studies are particularly difficult to interpret.

Other Tumours (Delbeke, 1999a)

As indicated previously, ^{18}FDG PET scanning has been investigated for the diagnosis and staging of many tumour types. Areas that show particular promise are in the staging of head and neck tumours where improved accuracy over CT and MRI has been shown in several studies. In pancreatic cancer ^{18}FDG PET distinguishes between benign and malignant pancreatic processes with 85–93% accuracy, altering the surgical management of patients with pancreatic masses in 41% of patients in one study. ^{18}FDG PET has been of value in distinguishing between benign and malignant musculoskeletal tumours, guiding biopsy to areas of greatest malignant potential, as well in predicting the preoperative response to chemotherapy. Anaplastic ^{131}I-negative thyroid tumours may be successfully staged with ^{18}FDG PET. ^{18}FDG PET has been shown to be more sensitive than CT in detecting recurrent ovarian cancer. Considerable promise has also been shown in oesophageal tumours where again CT is limited.

Protocols

Patients are typically injected with 5–10 mCi of ^{18}FDG intravenously after a 6–12 h fast (the latter to decrease cardiac and soft tissue uptake). The patient should be resting and not talking during the uptake period. Imaging usually takes place 30–120 min later. There is some evidence that delayed imaging may improve lesion detection. Imaging may be performed of the local tumour area alone (e.g. a head and neck tumour), of the whole body (e.g. melanoma staging) or both (e.g. lung cancer staging). Local scans usually correct for the attenuation of the radiation by the body tissue through the use of a 'transmission scan,' in reality a quick and dirty CT scan using a radioactive source. This improves the resolution of deep structures and enables semiquantitative measures such as SUR measurements to be performed (see above). The studies are typically reconstructed in multiple imaging planes.

Image Registration

Image registration or image fusion is the production of an image that combines both the patients anatomical (CT or MRI) scan with the PET study. This is usually produced by some form of software manipulation with or without markers placed on the patient. This combines the anatomical information from the CT, with the functional information from the PET scan. Thus the exact localization of abnormal areas of ^{18}FDG uptake can be identified. Image registration has the potential to add considerably to PET image interpretation (**Figure 12; see colour plate section**). Recently available are combined CT and PET

scanners where both sets of scans may be obtained sequentially without moving the patient resulting in inherently registered CT and PET images.

Interpretation and Pitfalls

PET scans are evaluated for abnormal areas of significantly increased local ^{18}FDG uptake. Wherever possible, PET ^{18}FDG scans should be interpreted with knowledge of the patient's history and current anatomical studies. This is aided by image registration as described above. Physiologically intense uptake of ^{18}FDG occurs in the brain, heart (variable amounts) and urine. Moderate uptake may be seen in Waldeyer's ring, muscle and bowel. The thyroid and thymus may accumulate moderate amounts of ^{18}FDG, the latter in children and postchemotherapy patients. Very low activity is seen in fat, lung and bone, but diffuse bone marrow activity can be seen immediately after chemotherapy courses. Muscle activity (even through tension and talking) during the uptake period can significantly increase local muscle activity.

False-negative studies may occur in low-grade tumours such as low-grade lymphoma and sarcoma, in tumours with extensive necrosis and in very small lesions, particularly if of lower malignant potential (<1 cm, although smaller lesions are frequently seem when the glycolytic activity is high). Sclerotic bone metastases from prostate cancer are frequently under-diagnosed. Hyperglycaemic diabetic patients may produce false-negative studies, presumably through competition for glucose uptake. False-positive studies have been reported in inflammation and infection (such as abscesses, tuberculosis), sarcoidosis and occasionally in inflammatory nodes. With the exception of sarcoidosis, the uptake of ^{18}FDG is usually mild to moderate in these cases, and it is possible that the use of SURs or similar ratios may help in diagnosis. Scans performed too soon (<6 months) after radiotherapy may show ^{18}FDG uptake into areas of inflammation. Extravasated tracer taken up into local lymph nodes may be mistaken for metastases. Physiological activity in the bowel can sometimes be relatively intense, and the evaluation of pelvic malignancies may be impaired by urine activity in the bladder. In these patients, catheterization and possibly bladder irrigation can be helpful. Prostate cancer is particularly problematic with PET, as both primary and nodal metastatic lesions are obscured by urinary activity, and frequently the primary tumour has a low relative uptake of ^{18}FDG except in more aggressive cases.

Other Positron-emitting Radiotracers

Several other PET radiotracers are currently under investigation in oncology.

Protein Synthesis Markers

^{11}C-labelled methionine, leucine and others have been used to evaluate tumour protein (amino acid) metabolism (typically raised relative to normal tissue). Most work has been done in brain tumours, where they have the advantage over ^{18}FDG of lower background activity, and hence easier visualization of lower grade tumours from normal cortex (Kaschten *et al*., 1998) (**Figure 13**). The short half-life of these tracers (20 min) limits geographic availability, and there is some debate as to whether they image protein metabolism or a more complex combination of metabolic processes.

DNA Synthesis Markers

^{11}C-labelled thymidine and uridine have been used to assess for rapid DNA synthesis in tumours. As yet, the role of these radiotracers have to play is unclear.

Fluoride Bone Scans

$^{18}F^-$ acts as an analogue of hydroxyl ions, being incorporated into the bone matrix as a marker of osteoblastic activity. In this way a bone scan similar physiologically to [^{99m}Tc]MDP is produced. The advantages over conventional bone scans are a superior resolution and the inherently multiplanar images leading to improved accuracy for the detection of bone metastases. Its main disadvantage is the significantly increased cost.

Blood Flow Studies

Tumour blood flow (typically, but not inevitably, increased over normal tissues and benign processes) can be measured using radioactive water, $H_2{}^{15}O$. This has not yet found clinical utility, partially because of the poor resolution of these studies and the extremely short half-life of ^{15}O (2 min).

Drug Markers

By attaching a positron-emitting radioactive label to a chemotherapeutic agent, PET allows one to assess *in vivo* if a particular tumour or metastasis in an individual patient accumulates and metabolizes that drug. This has been used in colorectal tumours by labelling 5-fluorouracil (5-FU) with ^{18}F, where it appears in preliminary studies to give useful prognostic information regarding the effectiveness of this agent. These types of studies have the potential to individualize chemotherapy to optimize gains rather than waiting until after several courses of chemotherapy to assess tumour response.

Neuroreceptor Markers

PET has been used extensively in neuroimaging to assess neuroreceptor distribution and density. [^{18}F]fluorodopa is an analogue of dopamine and is taken up by the pituitary in low concentrations in the normal brain, being converted into dopamine. In functioning pituitary adenomas, markedly increased [^{18}F]fluorodopa uptake is seen. This uptake is rapidly blocked by the administration of bromocriptine

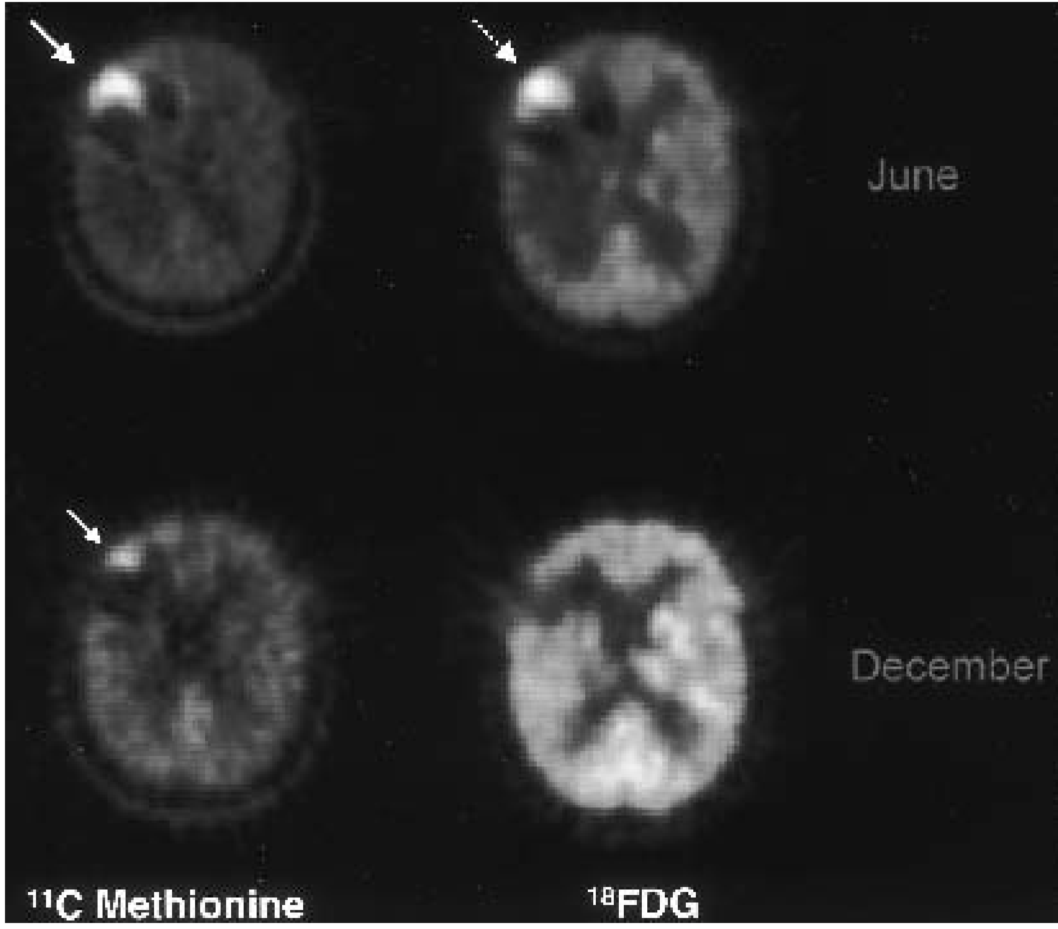


Figure 13 [^{11}C]methionine and ^{18}FDG scans in a patient with a grade 3 glioma of the right frontal lobe at two dates 6 months apart. In June (preradiotherapy), increased uptake at the tumour site is seen on both the [^{11}C]methionine (solid arrow) and ^{18}FDG (dashed arrow) studies. Postradiotherapy (December), whereas the [^{11}C]methionine scan still shows uptake consistent with residual tumour (small arrow), the tumour uptake on the ^{18}FDG scan is equal to the background cortical activity and cannot be distinguished. (Images courtesy of the Clinical PET Centre, United Medical Schools of Guy's, St. Thomas' and King's Hospitals, London, UK.)

(within hours) in those patients in whom a response to bromocriptine is later seen on MRI. In this way the PET scan predicts the success of bromocriptine therapy.

SENTINEL NODE LOCALIZATION

Principles and Indications

In recent years techniques have been developed which aim to avoid full regional lymph node dissections in patients with breast cancer and melanoma. The 'sentinel' node(s) – the node or nodes to which the area of the tumour initially drains – is identified and excised. Only those patients in whom the sentinel node contains tumour cells undergo full loco-regional lymph node dissections. The sentinel node may be identified by two methods: injecting isosulfan blue dye at the tumour site and tracing back blue lymphatics to the first draining blue node(s) or by injecting ^{99m}Tc sulfur colloid at the tumour site and using a handheld gamma probe to localize the hot nodes in the operating room at the time of surgery. In many centres both procedures are performed to maximize the sensitivity of node localization. Only the gamma probe procedure will be addressed here.

Protocol

The techniques differ between melanoma and breast cancer patients.

Melanoma

In melanoma, a total dose of 200 μCi of ^{99m}Tc sulfur colloid which has been filtered to reduce the particle size to <0.2 μm is injected intradermally at four sites around the tumour or incision site, usually using no more than 0.1 mL per injection. As the draining nodal basin is not always

predictable, especially in truncal lesions, imaging is performed immediately for 15 min, followed by delayed images at 2 h. A transmission scan is usually performed to provide a body contour for better localization. The site of the first visualized node can be marked on the patient's skin to aid localization in the operating room using the gamma probe. Usually a node with a count rate greater than 10 times the background counts when measured *ex vivo* is acceptable.

Breast Cancer

Sentinel node procedures are still controversial in breast cancer, although a recent multicentric study with over 500 patients appears to validate the technique (Tafra *et al*., 2001). There has been as yet no uniform agreement on the method of injecting the radiotracer in breast cancer. Studies have shown that unfiltered ^{99m}Tc sulfur colloid is more sensitive (with a larger particle size), but methods have included intradermal injections in the skin above the tumour, subareolar injections and injections around the tumour or excision site within the breast. Peritumoural and subareolar injections usually use a larger volume of injectate (4–5 mL), and injected doses range from 0.1 to 1 mCi. The intradermal injections are similar to those done in melanoma patients. The tumour is localized by ultrasound or palpation usually. Surprisingly, these different methods, which are sometimes performed in combination, appear to give very similar results. Imaging may or may not be performed as the drainage routes are less variable than in melanoma. Imaging may be of value in medial lesions to assess for sentinel nodes in the internal mammary chain, but its role has yet to be defined.

Accuracy and Limitations

In melanoma patients, the combined technique results in an identification rate of a sentinel node in >98% of patients, with a false-negative rate (compared with full node dissections) around 1% (Lange, 2000). In breast cancer patients various studies have shown localization rates of 87–97%, with a false-negative rate of 4.3% in the recent multicentric study (Tafra *et al*., 2001). Surgical experience in localizing the node is important.

CONCLUSION

Nuclear medicine, by virtue of its ability to image function, plays an integral role in the evaluation of patients with tumours or suspected tumours. This includes diagnosis (differentiating benign from malignant processes), staging and the evaluation of recurrence. In most cases, it acts as an adjunct to anatomical imaging methods such as plain films, CT or MRI, but on occasions it is the primary imaging method of choice (e.g. bone scanning, ^{131}I scanning). Frequently, the combination of the anatomical and functional information has a synergistic effect and this often overcomes the inherently poorer resolution of nuclear imaging. As PET becomes more widely available, it is likely that it will play an increasingly vital role in the management of oncology patients.

REFERENCES

Avril, M., *et al*. (1996). Metabolic characteristics of breast tumours with positron emission tomography using F-18 flurodeoxyglucose. *Journal of Clinical Oncology*, **14**, 1848–1857.

Babaian, R. J., *et al*. (1994). Radioimmunoscintigraphy of pelvic lymph nodes with 111-indium-labelled monoclonal antibody CYT-356. *Journal of Urology*, **152**, 1952–1955.

Black, K. L., *et al*. (1989). Use of thallium-201 SPECT to quantitate malignancy grade of gliomas. *Journal of Neurosurgery*, **71**, 342–346.

Coleman, R. E. (1999). PET in lung cancer. *Journal of Nuclear Medicine*, **40**, 814–820.

Delbeke, D. (1999a). Oncological applications of FDG PET imaging. *Journal of Nuclear Medicine*, **40**, 1706–1715.

Delbeke, D. (1999b). Oncological applications of FDG PET imaging: brain tumours, colorectal cancer, lymphoma and melanoma. *Journal of Nuclear Medicine*, **40**, 591–603.

Delbeke, D., *et al*. (1997). Staging recurrent metastatic colorectal carcinoma with PET. *Journal of Nuclear Medicine*, **38**, 1196–1201.

Harbert, J. C. (1996). Radioiodine therapy of differentiated thyroid carcinoma. In: Harbert, J. C., Eckelman, *et al*. (eds), *Nuclear Medicine. Diagnosis and Therapy*, 975–1019 (Thieme Medical Publishers, New York).

Jolesz, F. A. (1983). Functional imaging of the brain. *Medical Instrumentation*, **17**, 59–62.

Kalff, V., *et al*. (2001). Clinical impact of (18)F fluorodeoxyglucose positron emission tomography in patients with non-small-cell lung cancer: a prospective study. *Journal of Clinical Oncology*, **19**, 111–118.

Kaschten, B., *et al*. (1998). Preoperative evaluation of 54 gliomas by PET with fluorine-18-fluorodeoxyglucose and/or carbon-11-methionine. *Journal of Nuclear Medicine*, **39**, 778–785.

Kessler, L. S., *et al*. (1998). Thallium-201 brain SPECT of lymphoma in AIDS patients: pitfalls and technique optimization. *American Journal of Nuclear Radiology*, **19**, 1105–1109.

Khalkhali, I., *et al*. (1995). Scintimammography: the complementary role of Tc-99m sestamibi prone breast imaging for the diagnosis of breast cancer. *Radiology*, **196**, 421–426.

Kosuda, S., *et al*. (1994). Prediction of survival in patients with suspected recurrent cerebral tumors by quantitative thallium-201 single photon emission computed tomography. *International Journal of Radiation Oncology and Biological Physics*, **30**, 1201–1206.

Lange, J. R. (2000). The current status of sentinel node biopsy in the management of melanoma. *Dermatological Surgery*, **26**, 809–810.

Maldonado, A., *et al*. (2000). FDG-PET in the detection of recurrence in colorectal cancer based on rising CEA level. Experience in 72 patients. *Clinical Postron Imaging*, **3**, 170 (abstract).

Manyk, M. (1998). Clinical applications of radioimmunoscintigraphy with prostate-specific antibodies for prostate cancer. *Cancer Control*, **5**, 493–499.

McEwan, A., *et al*. (1985). Radio-iodobenzylguanidine for the scinitigraphic location and therapy of adrenergic tumours. *Seminars in Nuclear Medicine*, **15**, 132–153.

McNeil, B. J. (1984). Value of bone scanning in neoplastic disease. *Seminars in Nuclear Medicine*, **14**, 277–286.

Moog, F., *et al*. (1997). Lymphoma: role of whole-body 2-deoxy-2-[F-18]-fluoro-D-glucose (FDG) PET in nodal staging. *Radiology*, **203**, 795–800.

Olsen, J. O., *et al*. (1995). Somatostatin receptor imaging of neuroendocrine tumors with indium-111 pentetreotide (Octreoscan). *Seminars in Nuclear Medicine*, **25**, 251–261.

Rinne, D., *et al*. (1998). Primary staging and follow-up of high risk melanoma patients with whole-body ^{18}F-fluorodeoxyglucose positron emission tomography: results of a prospective study of 100 patients. *Cancer*, **82**, 1664–1671.

Sato, O., *et al*. (1998). Value of thallium-201 scintigraphy in bone and soft tissue tumors. *Journal of Orthopaedic Science*, **3**, 297–303.

Schlumberger, M. J., *et al*. (1999). The role of recombinant thyroid-stimulating hormone (rhTSH) in the detection and management of well-differentiated thyroid carcinoma: a roundtable discussion. *Journal of Endocrinology Investigations*, **22 (Suppl. 11)**, 35–41.

Slizofski, W. J., *et al*. (1994). Thallium imaging for brain tumours with results measured by a semiquantitative index and correlated with histopathology. *Cancer*, **74**, 3190–3197.

Sumiya, H., *et al*. (1998). Midcourse thallium-201 scintigraphy to predict tumor response in bone and soft-tissue tumors. *Journal of Nuclear Medicine*, **39**, 1600–1604.

Tafra, L., *et al*. (2001). Multicenter trial of sentinel node biopsy for breast cancer using both technetium sulfur colloid and isosulphan blue dye. *Annals of Surgery*, **233**, 51–59.

Turner, D. A., *et al*. (1978). Gallium-67 imaging in the management of Hodgkin's disease and other malignant lymphomas. *Seminars in Nuclear Medicine*, **8**, 205–218.

FURTHER READING

Freeman, L. M. and Blaufox, M. D. (eds) (1995). Nuclear endocrinology. *Seminars in Nuclear Medicine*, **25**, 205–286.

Freeman, L. M. and Blaufox, M. D. (eds) (1997). The role of nuclear medicine in oncological diagnosis (Part 1). *Seminars in Nuclear Medicine*, **27**, 3–81.

Freeman, L. M. and Blaufox, M. D. (eds) (1997). The role of nuclear medicine in oncological diagnosis (Part 2). *Seminars in Nuclear Medicine*, **27**, 1–80.

Freeman, L. M., *et al*. (eds) (2000). Sentinel node localization. *Seminars in Nuclear Medicine*, **30**, 4–64.

Maisey, M. N., *et al*. (1999). *Atlas of Clinical Positron Emission Tomography* (Oxford University Press, New York and Arnold, London).

Chapter 80

Breast Imaging and Diagnosis

Elsie Levin
Boston University School of Medicine, Boston, MA, USA

CONTENTS

- Introduction
- Screening
- Diagnostic Mammography
- Galactography (Ductography)
- Conclusion

INTRODUCTION

Mammography, a low dose X-ray examination of the breast, is performed to screen asymptomatic women for breast cancer and to diagnose and manage benign and malignant breast disease in screened and symptomatic patients. High-quality mammography is dependent on properly trained technologists using dedicated mammography equipment and physicians who are educated and trained in the subtleties of mammographic interpretation (Smith and D'Orsi, 2000). These issues have been addressed in the United States with the passage of the Mammography Quality Standards Act in 1992 which requires mammography facilities to pass periodic rigorous inspections.

Mammography is the only reliable method of consistently detecting nonpalpable breast cancer, but some cancers may not be detected owing to observer error, the location of the lesion such that it may be difficult to include on the film, a dense parenchymal pattern which can obscure a lesion or because the lesion fails to produce a recognizable abnormality (Sadowsky and Levin, 1994). It is therefore important that the mammogram be interpreted with all the available clinical information.

The field of breast imaging continues to evolve as mammographic technology improves and new technologies such as ultrasound, contrast-enhanced magnetic resonance imaging (MRI) and percutaneous biopsy procedures have been adopted. Full-field digital mammography has recently become available for clinical use. The traditional film screen cassette is replaced with a specialized detector which converts the latent X-ray image into an electronic digital image. The major advantages of digital mammography include the acquisition speed and also the ability to optimize the image by adjusting contrast and brightness, which may enhance lesion detection, particularly in the dense breast (Feig and Yaffe, 1995). The digital format allows for teleradiology, long-distance consultation and computer analysis. Studies have shown that double reading of screening mammograms can increase the number of cancers detected by 4–5%, but this is not always practical or feasible (Burhenne *et al*., 2000). A computer can function as a second reader and potentially reduce the false-negative rate of screening mammography (Burhenne *et al*., 2000).

SCREENING

Mammographic screening is the most reliable method to find early breast cancer in large populations and consists of a craniocaudal and mediolateral oblique view of each breast. In 1997, the American College of Radiology recommended annual mammographic screening for asymptomatic women beginning at age 40 years, since there was mounting evidence that younger women have a shorter lead time for mammographic detection of breast cancer (Feig *et al*., 1998). Women at high risk of breast cancer (genetic predisposition, significant family history or those receiving chest irradiation for Hodgkin disease) should begin screening before age 40 years.

Controlled studies have shown that breast cancer mortality rates can be reduced by 25–30% by screening with mammography and physical examination. The HIP (Health Insurance Plan of New York) study of 1963–1969 was the first prospective randomized controlled trial designed to evaluate the efficacy of breast cancer screening with mammography and clinical breast examination. Ten years after entry into the study, there were 30% fewer breast cancer deaths in the study group compared with the controls (Smith and D'Orsi, 2000).

The largest randomized trial of breast cancer screening is the Swedish Two-county study which included 133 000 women aged 40–74 years of whom 77 000 were invited for

screening. A single-view mammogram (MLO), without clinical breast examination, was performed every 24 months in women aged 40–49 years and every 33 months in women aged over 50 years. After 8, 11 and 14 years of follow-up, the study group had approximately 30% fewer breast cancer deaths (Smith and D'Orsi, 2000).

There has been controversy regarding the benefits of screening women in their forties with mammography. The Gothenburg randomized trial of breast cancer screening began in 1982 with approximately 52 000 women between ages 39 and 59 years, randomized to screening and control groups. After 11 years of follow-up, there was a statistically significant reduction of breast cancer mortality (44%) in women aged 39–49 years. The most recent updated reports of the Malmo trial also showed a statistically significant reduction of breast cancer mortality (36%) in women aged under 50 years at the time of randomization (Smith and D'Orsi, 2000). A meta-analysis of the screening trials was performed to overcome the limit of small sample sizes by combining age-specific results. The most current data showed 18% fewer breast cancer deaths when all trials were combined, 26% fewer deaths when only population-based trials are combined and 29% fewer deaths for all five Swedish randomized controlled trials (Smith and D'Orsi, 2000). The results of the Canadian randomized trial of breast cancer screening in women aged 40–49 years have been controversial owing to questions regarding the study design, quality of mammography and excess of advanced breast cancers in the screening group (Kopans and Feig, 1993).

Since the risk of developing breast cancer increases with age, screening of elderly women should continue at yearly intervals unless there are significant comorbid conditions. Women aged over 64 years account for almost 50% of newly diagnosed cases of breast cancer, which is the leading cause of cancer deaths in this age group (Field *et al*., 1998).

A retrospective review of 119 consecutive women aged 65 years and older with newly discovered breast cancer, who had a previous normal screening evaluation (mammography and clinical breast examination) 8–30 months before diagnosis, showed significantly smaller tumours, a higher incidence of minimal disease and DCIS (ductal carcinoma *in situ*), fewer palpable interval tumours and less axillary node involvement when elderly women were screened annually versus biennially (Field *et al*., 1998). Similar results were obtained in a retrospective review of women aged 40–79 years. In this study, annual screening resulted in the detection of smaller tumours which had a more favourable prognosis (Hunt *et al*., 1999). Annual screening mammography also resulted in lower rates of recall, for further definition of questionable findings, than when screening was performed every 2 years (Hunt *et al*., 1999). In another retrospective review of 103 malignant breast tumours detected in 102 women aged 65–74 years, there were statistically significant differences between the screened and nonscreened group with respect to the size of the lesions (11 vs 21 mm), the incidence of minimal tumours (63 vs 37%) and incidence of lymph node metastases (11 vs 30%) (Gabriel *et al*., 1997). The overall positive predictive value of mammography for the 65–74-year-old patients was 47% versus 28% in the general population (Gabriel *et al*., 1997). Annual mammographic screening also resulted in a higher likelihood of finding smaller DCIS lesions than biennial screening, and small, low grade noncomedo DCIS was more common in the annual mammographic screening group than in the biennial screening group in a retrospective review of 166 consecutive mammograms of women evaluated for DCIS (Carlson *et al*., 1999).

Although the benefits of screening mammography have been documented, it is still underutilized, especially in women aged over 65 years (Harrison *et al*., 1997). The most important barrier to screening mammography is the lack of a recommendation by a woman's physician (Rimer, 1997). Another common reason cited is that women do not need to have one in the absence of symptoms. Other less common reasons include concerns about pain and about radiation risk and anxiety about the results. It is important for physicians to recommend screening because this has been one of the most important predictors of mammography use.

Mammography is the primary imaging modality for breast cancer screening. The sensitivity of mammography decreases in women who have predominantly dense fibroglandular tissue because cancers can be obscured. Ultrasound can detect malignant lesions that are occult to mammography, but since it cannot routinely detect calcifications and it is operator dependent, it cannot be used for mass screening. It may be useful in a subset of high risk patients, particularly those with dense breast tissue on mammography (Gordon and Goldenberg, 1995). In a prospective study of screening ultrasound in women with dense breast tissue, the overall cancer detection rate increased by 17%, and the number of cancers detected with imaging alone increased by 37% (Kolb *et al*., 1998).

Contrast-enhanced MRI can also detect lesions occult to mammography and ultrasound, but it is not designed for mass screening because of cost and low throughput. It may be useful in screening of high-risk patients with dense breast tissue but current trials are in the preliminary stages (Kuhl *et al*., 2000).

DIAGNOSTIC MAMMOGRAPHY

A diagnostic mammogram is a problem-solving examination performed in patients with either clinical findings or abnormalities detected on screening mammography. It is a comprehensive evaluation that should be monitored by a breast imaging radiologist who can perform ancillary

procedures such as ultrasound, contrast-enhanced MRI or percutaneous needle aspiration or biopsy in order to provide specific management recommendations (see **Figures 1** and **2**).

A diagnostic mammogram begins with the routine craniocaudal and mediolateral oblique views. Any clinical findings should be marked on a diagram and denoted with a radiopaque marker placed on the skin over the area in

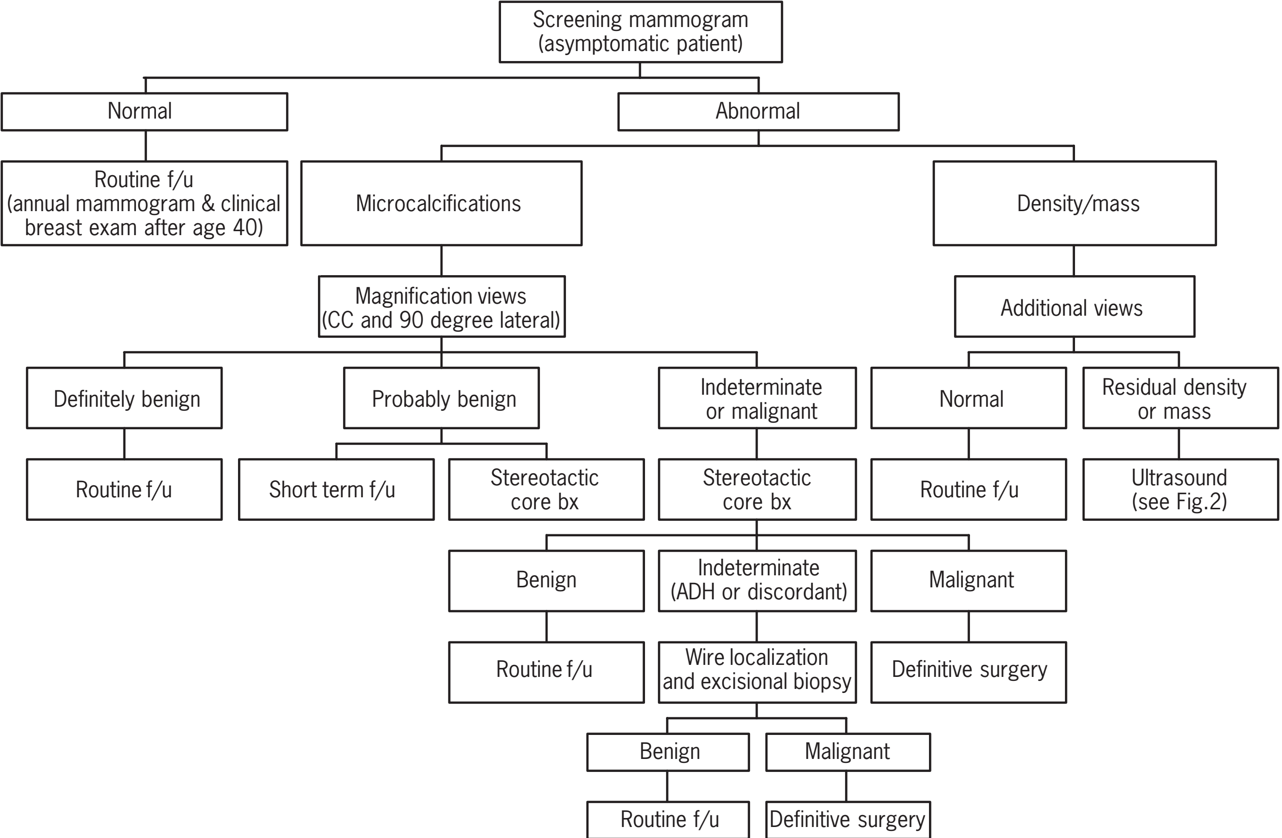


Figure 1 Breast imaging evaluation of the asymptomatic patient. f/u, follow-up; bx, biopsy.

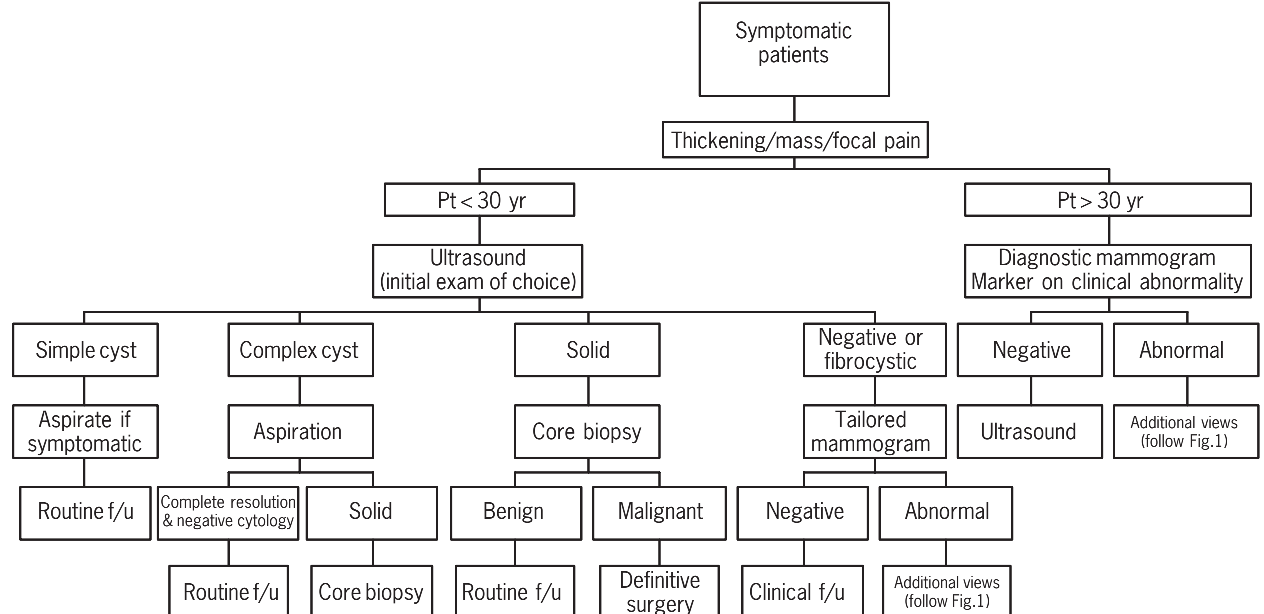


Figure 2 Breast imaging evaluation of the symptomatic patient.

question. Additional special mammographic views are performed following evaluation of the initial routine images to clarify mammographic or clinical findings. When used appropriately, additional views may reduce the number of biopsies and unnecessary short-term follow-up examinations by showing a characteristically benign mammographic sign. Alternatively, a biopsy may be indicated by suspicious features that may not be appreciated on routine films.

Mammographic findings that require further evaluation include microcalcifications, masses, architectural distortions and developing densities (particularly in post-menopausal women who are not on hormone replacement therapy).

Calcifications are often identified on both screening and diagnostic mammograms. The morphology, size and distribution of the microcalcifications are features that must be analysed to differentiate between benign and malignant calcifications. Magnification views in the craniocaudal and 90° lateral projections should be performed to characterize better calcifications that are seen on routine views. Calcifications that can be classified as definitely benign include dermal, vascular, secretory, eggshell and coarse calcifications related to involuting fibroadenomas or fat necrosis. Calcifications which occur in micro- or macrocysts can be confidently diagnosed as benign by their change in appearance between the craniocaudal and true lateral view forming a meniscus on the 90° view (milk of calcium). Features that favour a benign aetiology of microcalcifications include round, smooth calcifications that are scattered, diffuse and often bilateral (Monsees, 1995).

Malignant calcifications tend to appear pleomorphic, granular or linear in shape (**Figure 3**). Orientation in a linear or segmental distribution should raise the suspicion for malignancy. Calcifications that cannot be classified as benign or malignant are considered indeterminate and biopsy must be considered for definitive diagnosis.

Microcalcifications account for approximately 50% of biopsies performed for nonpalpable lesions and, of these, 35–70% are found to be ductal carcinoma *in situ* (Monsees, 1995). In the past, indeterminate or malignant calcifications were biopsied surgically following image-guided wire localization. Stereotactic large-core needle biopsy has been shown to be a safe, cost-effective alternative to open biopsy (Lee *et al*., 1997; Liberman, 2000). In one study, patients were spared an open surgical biopsy in 88% of indeterminate clustered calcifications by using stereotactic core biopsy as the initial diagnostic procedure (Lee *et al*., 1997). If the calcifications prove to be benign at histological review, the patient can be placed back into routine screening. If the biopsy shows atypical ductal hyperplasia (ADH), surgical biopsy must be performed since carcinoma can be found in a significant percentage of cases when more tissue is removed. Approximately 20–56% of lesions yielding ADH by 14-gauge automated core biopsy will prove to have carcinoma at surgery (Liberman, 2000). The directional vacuum-assisted device, especially when used with an 11-gauge probe, decreases but does not eliminate the problem of histological underestimates including the upgrading of DCIS to invasive carcinoma (Brem *et al*., 1999; Darling *et al*., 2000; Philpotts *et al*., 2000).

If malignancy is diagnosed by stereotactic core biopsy, treatment options can be discussed with the patient. In one study, a single surgical procedure was performed in 84% of patients who underwent core biopsy versus 29% of patients who underwent surgical biopsy initially for diagnosis (Liberman *et al*., 1997). Smith *et al*. found that the average number of surgeries performed was 1.25 in women whose carcinoma was diagnosed by percutaneous core biopsy versus 2.01 in women with surgically diagnosed carcinoma (Smith *et al*., 1997). There was no difference in the likelihood of obtaining tumour-free margins whether the patient underwent open surgical or core biopsy as the initial diagnostic procedure (Liberman *et al*., 1997; Gundry and Berg, 1998). These studies indicate that percutaneous core biopsy provides the information needed to plan surgery and can decrease the number of surgical procedures required, which reduces the overall cost of treatment (Liberman *et al*., 1997; Smith *et al*., 1997; Gundry and Bery, 1998).

A screening mammogram may also require further evaluation when a mass or density is identified. A lesion is described as a density when it is seen in only one projection while the term mass is applied to a space occupying lesion seen in two orthogonal projections (Evans, 1995). Spot compression views can show that an apparent 'mass' on a screening mammogram is due to overlapping glandular tissue by providing better compression to the area in question. Once a mass is identified, it should be characterized as to its size, shape, margins and density as well as its location. Features that favour a mass being benign include round or oval shape with well circumscribed margins. Fat-containing masses such as hamartomas, lymph nodes and lipid cysts can be accurately classified as benign requiring no further intervention. The likelihood of malignancy increases with irregularly shape and obscured, poorly defined or spiculated margins. These features are easier to appreciate on spot compression magnification views (**Figure 4**).

Ultrasound is the most important adjunct to the mammographic evaluation of masses and can help localize a lesion seen only on one view (Sadowsky *et al*., 1997). When strict criteria are applied, ultrasound is 98–100% accurate in distinguishing between cystic and solid lesions. A cyst should be anechoic, have a sharp posterior wall and be enhanced through transmission. A cyst does not need to be aspirated if these criteria are met unless it is causing symptoms or to prove that the cyst accounts for the mammographic abnormality (Sadowsky *et al*., 1997). Solid masses identified at ultrasound should be evaluated

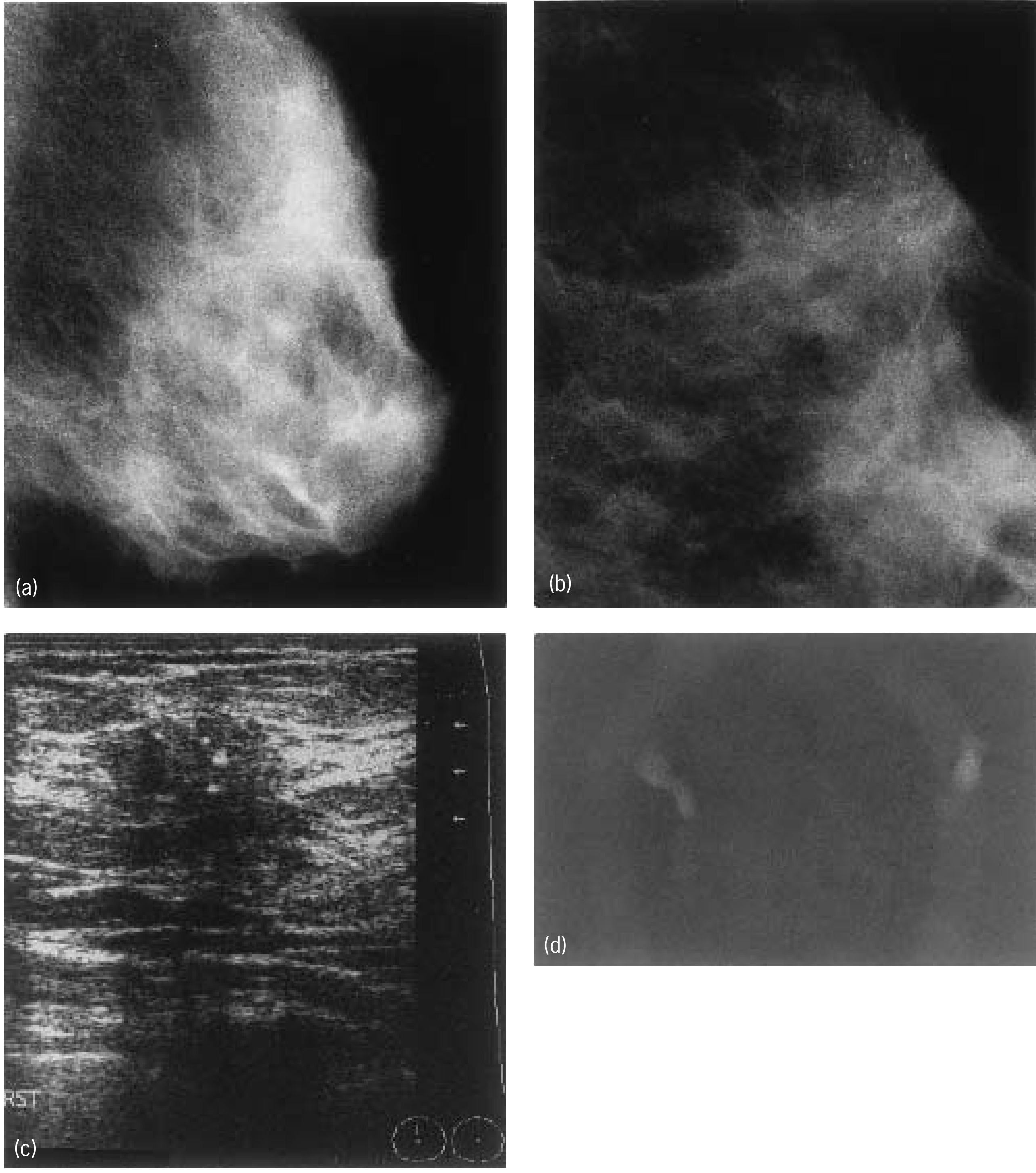


Figure 3 Baseline mammogram of a 39-year-old woman with dense breast tissue. Her mother was diagnosed with breast cancer at age 49 years. (a) The oblique view shows calcifications above the nipple which are difficult to appreciate. (b) A 90° lateral spot magnification view confirms the presence of suspicious linear, pleomorphic calcifications. (c) Ultrasound reveals an unsuspected, irregular mass containing calcifications. (d) The specimen X-ray obtained after ultrasound guided core biopsy was performed confirms that calcifications were removed. After consultation with the multidisciplinary team, the patient was treated by lumpectomy, radiation and chemotherapy for node-negative invasive ductal carcinoma.

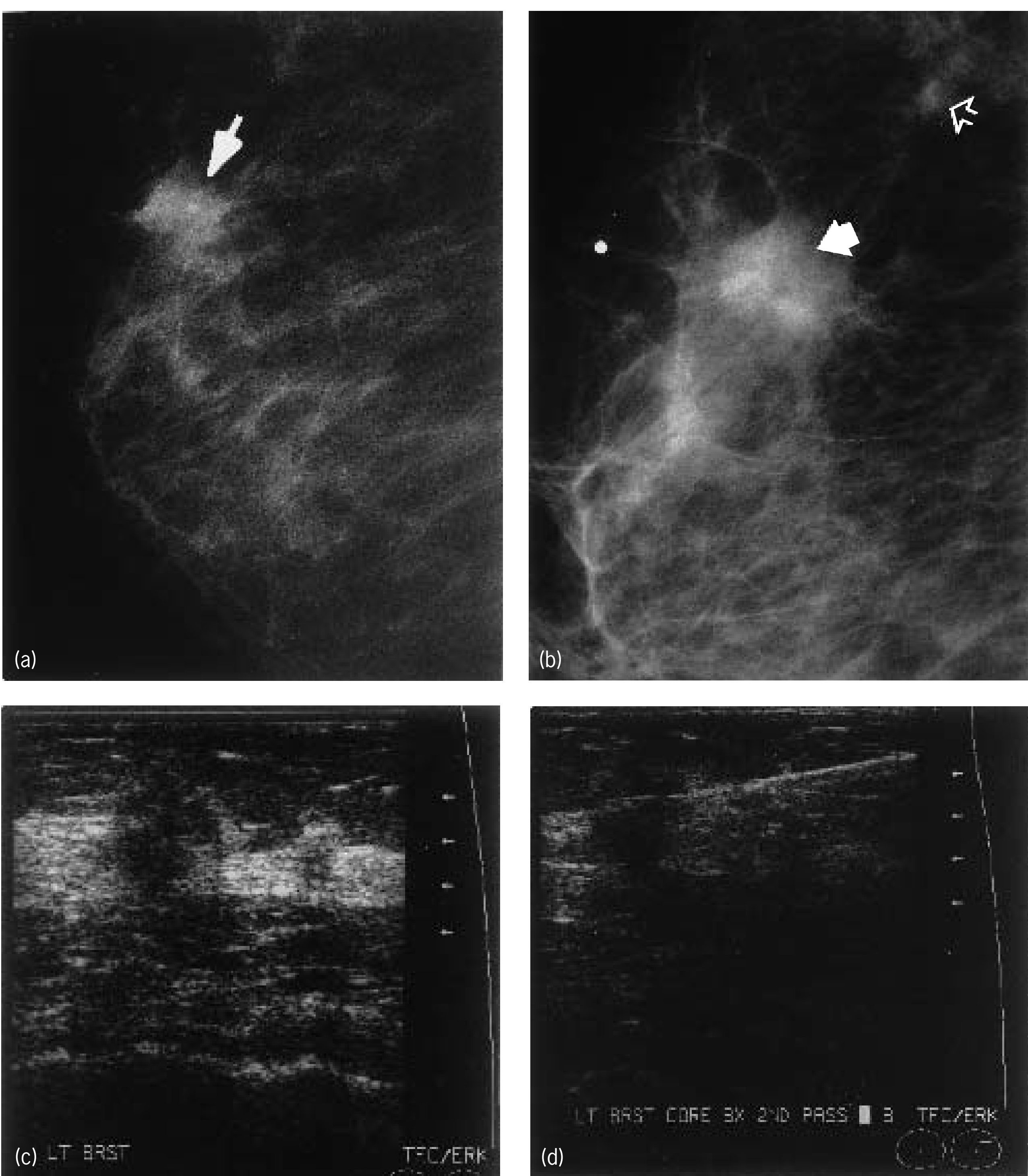


Figure 4 Screening mammogram of a 59-year-old women with fatty replaced breast tissue. (a) The oblique view showed a new mass in the superior left breast (arrow). (b) A 90° lateral spot magnification view demonstrates the irregular margins of the mass (solid arrow) and also shows a second smaller spiculated mass containing microcalcifications (open arrow). (c) Ultrasound features of both masses are also suspicious (irregular contour, taller than wide). (d) Ultrasound image of core biopsy needle passing through the larger mass. Both masses proved to be moderately differentiated invasive ductal carcinoma with an extensive intraductal component. A mastectomy with immediate reconstruction was performed.

for size, shape and margins to help differentiate between benign and malignant breast masses (Stavros *et al*., 1995; Skaane and Engedal, 1998; Rahbar *et al*., 1999). Breast carcinoma can be suspected when a solid mass has irregular margins, is taller than it is wide and is associated with acoustic shadowing. There can be overlap between the ultrasound features of benign and malignant breast masses, and for this reason breast ultrasound should be performed by the breast imaging radiologist who can correlate the mammographic, ultrasonographic and clinical findings (Sadowsky *et al*., 1997). Contrast-enhanced breast MRI is emerging as another useful imaging modality when mammography and ultrasound findings are indeterminate (Hrung *et al*., 1999; Lee *et al*., 1999).

Ultrasound is invaluable in the evaluation of palpable breast masses and is the initial examination of choice in patients aged under 30 years (Evans, 1995). If the ultrasound is negative or shows fibrocystic tissue, a tailored mammogram is performed primarily to look for calcifications which can be missed on ultrasound. In patients over 30 years old, a diagnostic mammogram is performed with a radiopaque marker on the palpable mass. If the mammogram shows a mass or calcifications, additional mammographic views are obtained as described previously.

In one study, the sensitivity of mammography alone for the detection of cancer was 83% but improved to 91% after mammography and ultrasound were combined. Improved sensitivity was particularly evident in patients aged under 50 years (Zonderland *et al*., 1999). It is important to remember that although a diagnostic mammogram is performed to answer a specific clinical question, the remainder of both breasts must be screened for nonpalpable breast cancers (Rosen *et al*., 1999).

If a cyst is diagnosed at ultrasound, aspiration can be performed for symptomatic relief. If the mammogram and ultrasound are both negative, the decision to biopsy a palpable abnormality must be based on the clinical examination. When a solid mass is identified by ultrasound, whether palpable or nonpalpable, a core biopsy can be performed to establish the histological diagnosis (Liberman *et al*., 1998). Ultrasound-guided core biopsy is easier on the patient than stereotactic core biopsy and can be performed faster with no ionizing radiation and at lower cost. In one study, percutaneous image guided biopsy of palpable breast masses spared 74% of patients an additional diagnostic tissue sampling procedure (Liberman *et al*., 2000). If the imaging and clinical findings suggest a benign lesion and the patient wants the lesion removed regardless of the tissue diagnosis, surgical excision may be preferable. The use of core biopsy on indeterminate masses increased the malignancy rate at open biopsy from 21 to 55% with a reduction in billed charges estimated at $368 000 per 100 biopsies (Rubin *et al*., 1995). If a mass proves to be malignant after ultrasound-guided core biopsy, the therapeutic options can be discussed with the patient and a single surgery can be planned.

Contrast-enhanced MRI is also being used to stage newly diagnosed breast cancer patients prior to surgery since it may reveal unsuspected multifocal, multicentric or contralateral breast cancer (Orel *et al*., 1995; Orel, 1998). In one study, the therapeutic procedure was changed in 14.3% of patients because the MRI findings were more extensive than appreciated on clinical, mammographic and ultrasonographic exams (Fischer *et al*., 1999).

GALACTOGRAPHY (DUCTOGRAPHY)

Galactography is a contrast examination of a major lactiferous duct and its branches which is performed after a significant nipple discharge has been identified (Tabar *et al*., 1983; Jones, 1992; Cardenosa *et al*., 1994). A patient who presents with nipple discharge should initially be evaluated clinically to determine whether the discharge is physiological or pathological (see **Figure 5**). A significant discharge is usually unilateral, from one duct orifice and most often occurs spontaneously. The discharge should be tested for occult blood and the colour and consistency of the discharge should be noted. Serous, serosanguineous, bloody and clear watery discharge should be investigated further. Bilateral, multiple duct milky discharge is usually related to hormonal factors or can be secondary to various medications. Thick green, yellow or brown discharge is usually related to duct ectasia and most often is a nonspontaneous discharge which comes from multiple ducts.

Once it is determined that a significant nipple discharge is present, a ductogram can be performed to identify the cause of the discharge and the location of the lesion. If no recent mammograms are available, preliminary films should be obtained. The study is usually performed with the patient in the oblique supine position with the nipple at the highest point of the chest, but some prefer to have the patient seated. Infection can be prevented by meticulous cleansing of the nipple and areola with Betadine and alcohol and observing aseptic technique. Following cleansing of the nipple, a tiny drop of discharge is produced to identify the nipple orifice, which is easier to see with use of a fluorescent lamp with magnifying lens. A special 30-g cannula bent at 90° is used to probe the nipple gently until the duct orifice is entered and then contrast material is slowly injected using a 1-mL syringe with attached tubing. Mammograms are then obtained in the craniocaudal and mediolateral projections followed by spot compression magnification views once the location of the duct system has been identified. Following the ductogram, fluid can be collected for cytological evaluation.

Galactography is a safe, well tolerated procedure with few complications, but it is contraindicated in patients with acute mastitis or breast abscess since it may exacerbate the inflammatory process. If there is a strong allergic history, nonionic contrast can be employed with or without

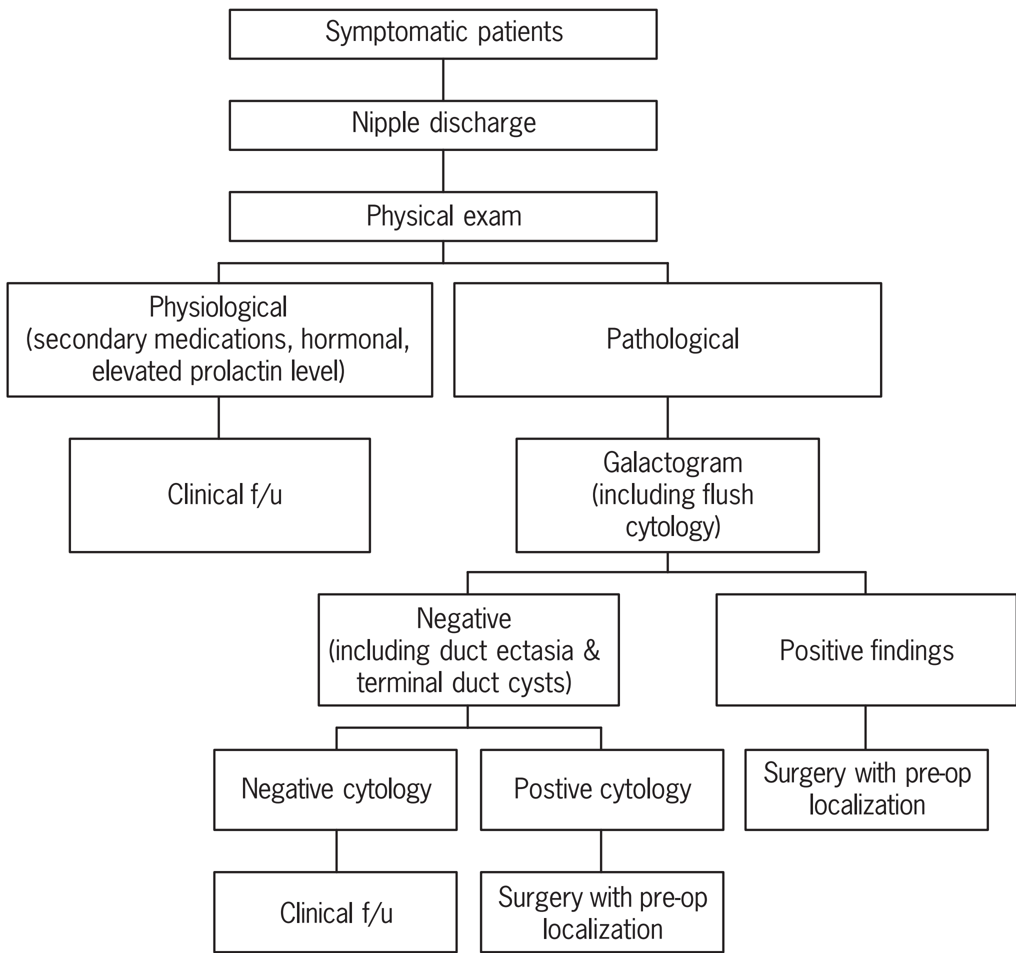


Figure 5 Evaluation of nipple discharge.

pretreatment with steroids and antihistamines (Levin and Sadowsky, 1996).

Galactographic findings suggestive of papillomas include solitary or multiple intraluminal filling defects or complete obstruction to the retrograde flow of contrast (Tabar *et al*., 1983; Woods *et al*., 1992). Most solitary papillomas occur centrally in the large subareolar ducts while multiple papillomas (papillomatosis) are usually peripheral in location, arising within the terminal duct lobular unit. It can be difficult to distinguish between a benign intraductal papilloma and intraductal carcinoma, especially since there can be microscopic involvement of a papilloma with DCIS. Findings suggestive of carcinoma include irregular filling defects, ductal irregularity, distortion, displacement or encasement of ducts (Cardenosa *et al*., 1994). Additional findings suggestive of carcinoma include parenchymal pooling of contrast or opacification of a duct which leads to a suspicious mass, area of increased density or microcalcifications.

Galactographic findings which do not require surgical intervention include duct ectasia and terminal duct cysts. If significant pathology is demonstrated, preoperative localization galactography can be performed using a combination of two-thirds contrast and one-third methylene blue to stain the involved duct. If the lesion is peripheral in location, or several branch points from the main duct, the preoperative galactogram can be supplemented by placement of a localizing hook-wire to guide the surgeon further. These preoperative localization procedures allow for a more conservative surgical excision while ensuring complete removal of the lesion and affected duct (Van Zee *et al*., 1998). In one study, when preoperative localization procedures were not employed, 20% of lesions (6/30) shown by galactography were not identified at histopathological examination (Baker *et al*., 1994).

Currently, galactography is the only diagnostic study available for the evaluation of nipple discharge. Preliminary reports indicate that contrast-enhanced MRI of the breast may be an alternative diagnostic tool for the evaluation of nipple discharge (Orel *et al*., 2000).

Even though galactography is a safe, well-tolerated procedure, it has not gained widespread acceptance by physicians (Jones, 1992; Woods *et al*., 1992). In postmenopausal women, some surgeons advocate excision of the major subareolar ducts while solitary duct excision is performed in premenopausal women without performing

preoperative diagnostic imaging (Woods *et al.*, 1992; Cardenosa *et al.*, 1994). This type of surgery will treat the symptom of nipple dischage but does not ensure that the lesion causing the discharge has been removed. Significant breast deformity can occur following excision of the major subareolar ducts.

CONCLUSION

Mammographic screening is the most reliable method to find early breast cancer in large populations. The detection of both benign and malignant breast disease has increased as more women participate in screening. The secondary costs resulting from screening mammography can account for 50% of the total cost of diagnosing breast cancer. The psychological and economic costs should not be underestimated and therefore when an abnormality is detected on a screening examination, a thorough diagnostic evaluation must be performed to characterize the lesion. Periodic mammographic follow-up of probably benign lesions has been established as a safe, reasonable alternative to surgical biopsy (Sickles, 1991). If a lesion remains indeterminate after a complete mammographic and ultrasonographic evaluation, percutaneous image guided biopsy can be performed. When used appropriately, these procedures are not only cost effective but also truly cost saving. In most institutions, the overall cost for a needle biopsy is 25–30% that of wire localization and surgical excision. Considering the large number of biopsies performed each year, the cost savings has been estimated to between $200 million and $1 billion per year. The reasons to biopsy a probably benign lesion include patient anxiety due to high risk or family history, planned pregnancy, difficulty with follow-up (unreliable or relocating patient) or ipsilateral or contralateral breast cancer to confirm benign diagnosis prior to definitive treatment.

Digital mammography and computer-aided detection are emerging technologies which may improve both screening and diagnostic mammography. High-resolution ultrasound and contrast-enhanced MRI are important adjuncts to mammography in the detection and diagnosis of benign and malignant breast diseases.

REFERENCES

Baker, K. S., *et al.* (1994). Ductal abnormalities detected with galactography: frequency of adequate excisional biopsy. *American Journal of Roentgenology*, **162**, 821–824.

Brem, R. F., *et al.* (1999). Atypical ductal hyperplasia: histologic underestimation of carcinoma in tissue harvested from implapable breast lesions using 11-gauge stereotactically guided direction vacuum assisted biopsy. *American Journal of Roentgenology*, **172**, 1405–1407.

Burhenne, L. J. W., *et al.* (2000). Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology*, **215**, 554–562.

Cardenosa, G., *et al.* (1994). Ductography of the breast: technique and findings. *American Journal of Roentgenology*, **162**, 1081–1087.

Carlson, K. L., *et al.* (1999). Relationship between mammographic screening intervals and size and histology of ductal carcinoma *in situ*. *American Journal of Roentgenology*, **172**, 313–317.

Darling, M. L. R., *et al.* (2000). Atypical ductal hyperplasia and ductal carcinoma *in situ* as revealed by large-core needle breast biopsy: results of surgical excision. *American Journal of Roentgenology*, **175**, 1341–1346.

Evans, W. P. (1995). Breast masses: appropriate evaluation. *Radiology Clinics of North America*, **33**, 1085–1108.

Feig, S. A. and Yaffe, M. J. (1995). Digital mammography, computer-aided diagnosis and telemammography. *Radiology Clinics of North America*, **33**, 1205–1230.

Feig, S. A., *et al.* (1998). American College of Radiology guidelines for breast cancer screening. *American Journal of Roentgenology*, **171**, 29–33.

Field, L. R., *et al.* (1998). Mammographic screening in women more than 64 years old: a comparison of 1 and 2 year intervals. *American Journal of Roentgenology*, **170**, 961–965.

Fischer, U., *et al.* (1999). Breast carcinoma: effect of preoperative contrast enhanced MR imaging or the therapeutic approach. *Radiology*, **213**, 881–888.

Gabriel, H., *et al.* (1997). Breast cancer in women 65–74 years old: earlier detection by mammographic screening. *American Journal of Roentgenology*, **168**, 23–27.

Gordon, P. B. and Goldenberg, S. L. (1995). Malignant breast masses detected only by ultrasound. *Cancer*, **76**, 626–630.

Gundry, K. R. and Berg, W. A. (1998). Treatment issues and core needle breast biopsy: clinical context. *American Journal of Roentgenology*, **171**, 41–49.

Harrison, R. A., *et al.* (1997). Breast cancer detection rates by screening mammography in elderly women. *The Breast Journal*, **3**, 331–336.

Hrung, J. M., *et al.* (1999). Cost-effectiveness of MR imaging and core-needle biopsy in the preoperative work-up of suspicious breast lesions. *Radiology*, **213**, 39–49.

Hunt, K. A., *et al.* (1999). Outcome analysis for women undergoing annual versus biennial screening mammography: a review of 24,211 examinations. *American Journal of Roentgenology*, **173**, 285–289.

Jones, M. K. (1992). Galactography: procedure of choice for evaluation of nipple discharge. *Seminars in Interventional Radiology*, **9**, 112–119.

Kolb, T. M., *et al.* (1995). Occult cancer in women with dense breasts: detection with screening US – diagnostic yield and tumor characteristics. *Radiology*, **207**, 191–199.

Kopans, D. B. and Feig, S. A. (1993). The Canadian National Breast Screening Study: a critical review. *American Journal of Roentgenology*, **161**, 755–760.

Kuhl, C. K., *et al*. (2000). Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: preliminary results. *Radiology*, **215**, 267–279.

Lee, C. H., *et al*. (1997). Cost-effectiveness of stereotactic core needle biopsy: analysis by means of mammographic findings. *Radiology*, **202**, 849–854.

Lee, C. H., *et al*. (1999). Clinical usefulness of MR imaging of the breast in the evaluation of the problematic mammogram. *American Journal of Roentgenology*, **173**, 1323–1329.

Levin, E. and Sadowsky, N. L. (1996). Complications of interventional procedures in the breast. In: Ansell, G. *et al*. (eds), *Complications in Diagnostic Imaging and Interventional Radiology*, 3rd edn. 603–609 (Blackwell Science, Oxford).

Liberman, L. (2000). Percutaneous imaging-guided core breast biopsy: state of the art at the millenium. *American Journal of Roentgenology*, **174**, 1191–1199.

Liberman, L., *et al*. (1997). Impact of core biopsy on the surgical management of impalpable breast cancer. *American Journal of Roentgenology*, **168**, 495–499.

Liberman, L., *et al*. (1998). US-guided core breast biopsy: use and cost effectiveness. *Radiology*, **208**, 717.

Liberman, L., *et al*. (2000). Palpable breast masses: is there a role for percutaneous imaging-guided core biopsy? *American Journal of Roentgenology*, **175**, 779–787.

Monsees, B. S. (1995). Evaluation of breast microcalcifications. *Radiology Clinics of North America*, **33**, 1109–1121.

Orel, S. (1998). High-resolution MR imaging for the detection, diagnosis and staging of breast cancer. *Radiographics*, **18**, 903–912.

Orel, S. G., *et al*. (1995). Staging of suspected breast cancer: effect of MR imaging and MR-guided biopsy. *Radiology*, **196**, 115–122.

Orel, S. G., *et al*. (2000). MR imaging in patients with nipple discharge: initial experience. *Radiology*, **216**, 248–254.

Philpotts, L. E., *et al*. (2000). Underestimation of breast cancer with 11-gauge vacuum suction biopsy. *American Journal of Roentgenology*, **175**, 1047–1050.

Rahbar, G., *et al*. (1999). Benign versus malignant solid breast masses: US differentiation. *Radiology*, **213**, 889–894.

Rimer, B. K. (1997). Current use and how to increase mammography screening in women. *Surgical Oncology Clinics of North America*, **6**, 203–221.

Rosen, E. L., *et al*. (1999). Ability of mammography to reveal nonpalpable breast cancer in women with palpable breast masses. *American Journal of Roentgenology*, **172**, 309–312.

Rubin, E., *et al*. (1995). Needle-localization biopsy of the breast: impact of a selective core needle biopsy program on yield. *Radiology*, **195**, 627–631.

Sadowsky, N. L. and Levin, E. (1994). Radiologic diagnosis of breast cancer. In: Putman, C. E. and Ravin, C. E. (eds), *Textbook of Diagnostic Imaging*, 2nd edn. 2117–2131, (WB Saunders, Philadelphia).

Sadowsky, N. L., *et al*. (1997). Ultrasonographic guidance for needle biopsy of breast lesions. *Surgical Oncology Clinics of North America*, **6**, 265–284.

Sickles, E. A. (1991). Periodic mammographic follow-up of probably benign lesions: results in 3,184 consecutive cases. *Radiology*, **179**, 463–468.

Skaane, P. and Engedal, K. (1998). Analysis of sonographic features in the differentiation of fibroadenoma and invasive ductal carcinoma. *American Journal of Roentgenology*, **170**, 109–114.

Smith, R. A. and D'Orsi, C. J. (2000). Screening for breast cancer. In: Harris, J. R. and Lippman, M. E. (eds), *Diseases of the Breast*, 2nd edn. 101–121, (Lippincott Williams & Wilkins, Philadelphia).

Smith, D. N., *et al*. (1997). Large-core needle biopsy of nonpalpable breast cancers: the impact on subsequent surgical excision. *Archives of Surgery*, **132**, 256–259.

Stavros, A. T., *et al*. (1995). Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology*, **196**, 123–134.

Tabar, L., *et al*. (1983). Galactography: the diagnostic procedure of choice for nipple discharge. *American Journal of Roentgenology*, **149**, 31–38.

Van Zee, K. J. *et al*. (1998). Preoperative galactography increases the diagnostic yield of major duct excision for nipple discharge. *Cancer*, **82**, 1874–1880.

Woods, E. R., *et al*. (1992). Solitary breast papilloma: comparison of mammographic, galactographic and pathologic findings. *American Journal of Roentgenology*, **159**, 487–491.

Zonderland, H. M., *et al*. (1997). Diagnosis of breast cancer: contribution of US as an adjunct to mammography. *Radiology*, **213**, 413–422.

FURTHER READING

Bassett, L., *et al*. (1997). Stereotactic core-needle biopsy of the Breast: a report of the joint task force of the American College of Radiology, American College of Surgeons, and College of American Pathologists. *CA Cancer Journal for Clinicians*, **47**, 171–190.

Feig, S. A. (1996). Methods to identify benefit from mammographic screening of women aged 40–49 years. *Radiology*, **201**, 309–316.

Liberman, L., *et al*. (1993). Screening mammography: value in women 35–39 years old. *American Journal of Roentgenology*, **161**, 53–56.

Roubidoux, M. A., *et al*. (1998). Breast cancer in women who undergo screening mammography: relationship of hormone replacement therapy to stage and detection method. *Radiology*, **208**, 725–728.

Slanetz, P. J., *et al*. (1997). Screening mammographic effect of national guidelines on current physician practice. *Radiology*, **203**, 335–338.

Tabar, L., *et al*. (1999). The natural history of breast carcinoma. What have we learned from screening? *Cancer*, **86**, 449–462.

Thurfjell, E. L. and Lindgren, J. A. (1994). Population-based mammography screening in Swedish clinical practice: prevalence and incidence screening in Uppsala County. *Radiology*, **193**, 351–357.

Thurfjell, E. L. and Lindgren, J. A. (1996). Breast cancer survivial rates with mammographic screening: similar favorable survival rates for women younger and those older than 50 years. *Radiology*, **201**, 421–426.

Thurfjell, E. L., *et al*. (1997). Screening mammography: sensitivity and specificity in relation to hormone replacement therapy. *Radiology*, **203**, 339–341.

Chapter 81

Percutaneous Biopsy

Brian C. Lucey, Michael M. Maher, Debra A. Gervais and Peter R. Mueller
Massachusetts General Hospital, Boston, MA, USA

CONTENTS

- Introduction
- Technique
- Lung and Mediastinum
- Liver, Spleen and Kidney
- Musculoskeletal System
- Miscellaneous Sites
- Conclusion

INTRODUCTION

With advances in treatment regimes for many neoplastic processes, it is becoming increasingly important to make an accurate diagnosis in each patient. Although there are characteristic radiological appearances of many tumours, a histological diagnosis is preferable in most cases. Traditionally this has meant an open surgical biopsy; percutaneous biopsy, however, has become an increasingly used diagnostic tool. All but the most superficial biopsies are now performed by interventional radiologists using image guidance. The purpose of this chapter is to outline the indications, techniques and success rates of percutaneous biopsy and explain its many uses.

TECHNIQUE

When presented with a diagnostic dilemma, the usual scenario involves a patient with a suspected diagnosis and a mass seen on imaging. The aim of the biopsy is to obtain sufficient cells or a tiny core of tissue from the mass to enable the pathologist to arrive at a diagnosis. Obtaining cells through a tiny needle is termed a fine-needle aspiration biopsy (FNAB). By definition, these needles are 20, 22, 23 or greater in gauge. The other percutaneous approach involves obtaining a core of tissue, usually with a 16–18-gauge needle. There are few contraindications to performing such a procedure. In most cases a large-gauge needle (16–18) can be inserted into a lesion and both large cores and FNA can be obtained through this needle. A mass located deep within the abdomen or pelvis or adjacent to vital structures may increase the technical difficulty of the procedure but these are not contraindications. Even a coagulopathy is only a relative contraindication to FNAB and then only if the mass is deeply located. In these cases, blood products can be given to correct the coagulopathy. Superficial masses may be aspirated despite an abnormal coagulation status. There is no organ or area in the body apart from the skull which is not amenable to percutaneous biopsy. Lung masses, liver, lymph nodes, renal, adrenal, pancreatic and prostate masses are the most commonly biopsied lesions. Musculoskeletal, breast and thyroid lesions are also amenable to a percutaneous approach. Several imaging modalities are available to guide the procedure: fluoroscopy, ultrasound, computed tomography (CT) and, in some institutions, magnetic resonance imaging (MRI).

Fluoroscopy is used primarily in aiding lung mass biopsies and today this is becoming uncommon. The initial lesion is usually identified on a chest radiograph. If the lesion can be seen on fluoroscopy then the patient is placed in the appropriate position on the fluoroscopy table with the lesion uppermost. The use of a 'C' arm aids needle localization in the 90° plane. A 20–22-gauge needle is directed into the lesion under fluoroscopic guidance. Once the needle is confirmed to be within the lesion, the aspirates are taken. Fluoroscopy has little value in assisting percutaneous biopsy in other organs.

Ultrasound is more frequently used in assisting percutaneous biopsy. Many organs are accessible to ultrasound guidance, particularly the liver, kidneys, spleen, thyroid, prostate and pancreas. Ultrasound is easy to perform, inexpensive and free from ionizing radiation. It allows real-time imaging of the biopsy needle, ensuring that the correct area is sampled. More recently, ultrasound has been used to guide biopsy of chest wall lesions and even hilar and mediastinal masses through areas of collapsed lung.

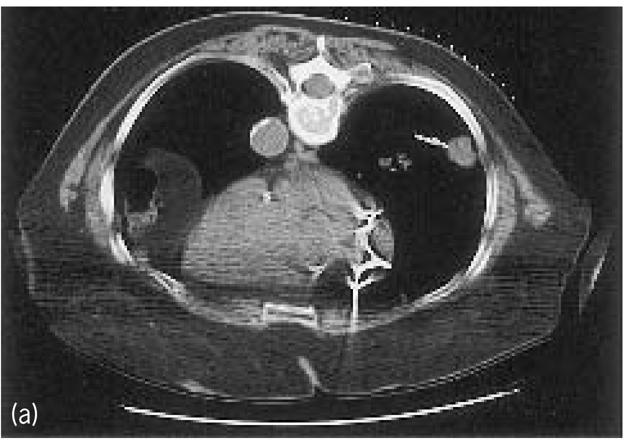
CT is also frequently used to assist percutaneous biopsy. It is particularly valuable for localizing lesions which are not clearly visualized by ultrasound, such as adrenal lesions and retroperitoneal structures. These tend to lie deep within the body cavity and CT gives a better overview of the target area. CT can also be used for solid organ biopsy. It enables smaller lesions to be located more accurately. A useful addition to the CT technique is to create a window through which the needle can reach the target. This has been reported as an approach for the adrenal glands (Karampekios *et al*., 1998).

Normal saline is injected into the paravertebral space to create a wider pathway for the needle insertion. The artificial window that is formed by displacing the pleura laterally allows a direct and safer access route to the retroperitoneum, thereby decreasing the risk of damaging adjacent structures, especially the pleura. A similar approach has been reported in the thorax for biopsying a posterior mediastinal mass through an extrapleural window created by dilute contrast (Bhagat *et al*., 1999).

The technique for biopsy remains similar irrespective of the imaging modality used. The lesion must first be located on a preliminary scan. Following this, an access route which is safe, effective and convenient must be chosen to approach the lesion. The patient should then be positioned accordingly. For all but the most superficial lesions, conscious sedation is recommended in the form of midazolam and fentanyl. The skin is then prepared and draped in a sterile manner. The skin is infiltrated with 1% lidocaine as a local anaesthetic. In a typical case, a 17-gauge Temno (Allegiance, McGaw Park, IL, USA.) coaxial needle is positioned within the lesion to be biopsied. This is done under real-time imaging in the case of ultrasound. Having positioned the 17-gauge needle correctly, fine-needle aspirates may be obtained using 20–22-gauge needles while aspirating with a 20-mL syringe. These samples are spread on to slides and immediately placed in formalin before they can air dry. An 18-gauge biopsy needle can then be passed through the coaxial needle and samples taken as required. Following the procedure, the patient should be observed for a period of 2–4 h. The procedure is usually very well tolerated by the patient. In other cases of superficial lesions such as thyroid or breast, conscious sedation is not routinely required. In thyroid lesions, the biopsy is performed using a 25-gauge needle. Between five and six passes are usually made. This usually provides sufficient cells for diagnosis. In cases near vital organs, 20-gauge core biopsies can be obtained without using a coaxial system.

LUNG AND MEDIASTINUM

There have been many studies looking at the accuracy of FNABs and core biopsies relating to nearly every organ in the body. One large study reviewing the results and complications of FNAB of lung lesions reviewed 651 pulmonary FNABs (Swischuk *et al*., 1998). The number of needle passes and needle size were correlated with the rates of pneumothorax and chest tube placement; 18–22-gauge needles were used. There was a diagnostic accuracy of 94% with a sensitivity for malignancy of 95%. The major complication when performing lung biopsies is pneumothorax. This study reported a pneumothorax rate of 26.9% with 9.2% requiring a chest tube placement. The pneumothorax rate has been shown to be independent of the depth of lung tissue traversed, number of passes and size of needle used, but does correlate with the number of pleural surfaces crossed. This study demonstrates excellent results for the technique. Similar results for FNAB have been reported by other investigators. A sensitivity of 92% for diagnosing malignancy was reported in another study involving 150 cases, with a pneumothorax rate of 2% (Subhannachart *et al*., 1999) (**Figure 1**). These results show that the technique is accurate, effective and safe for lung lesions.



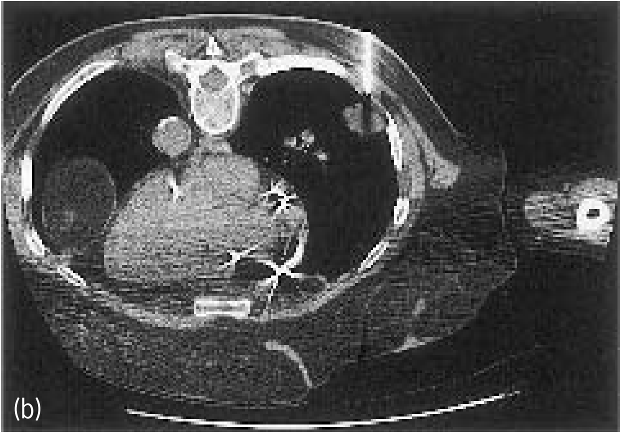
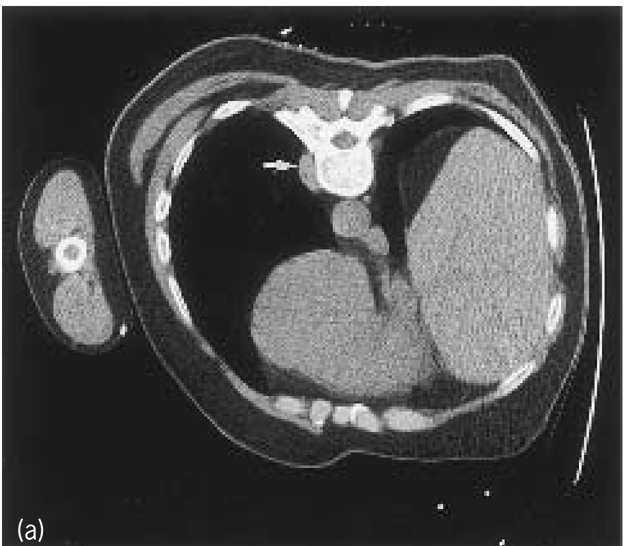


Figure 1 An example of a percutaneous lung biopsy. With the patient prone, (a) shows a right lung mass (arrow) with a localizing grid on the patient's back; (b) shows the same lesion with the biopsy needle placed at the posterior surface of the mass.



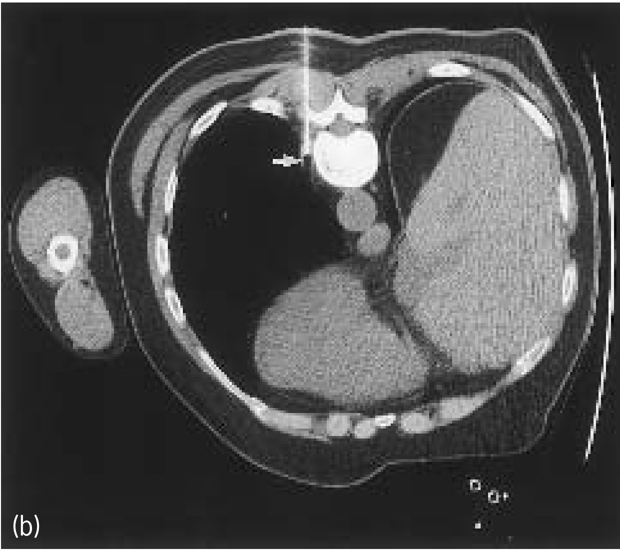


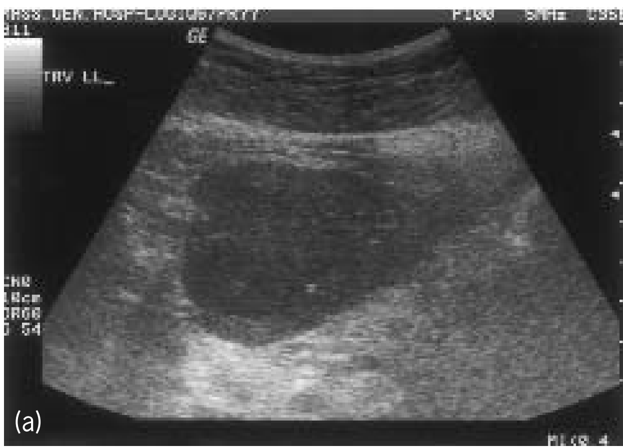
Figure 2 (a) A paravertebral soft tissue mass (arrow) in a 33-year-old male patient. The patient is in a left lateral decubitus position to create an intercostal window for biopsy. (b) The same lesion with the needle tip within it. This proved to be non-Hodgkin lymphoma. Note the black streak at the tip of the needle (arrow) indicating that the exact location of the needle tip has been imaged.

It can often be difficult to decide whether a small lesion is amenable to biopsy and to this end the results of biopsying lesions smaller than 2 cm was addressed by investigators in France (Laurant *et al*., 2000). They biopsied 202 lung lesions and recorded the sensitivity and specificity of the biopsy by lesion size. Lesions smaller than 2 cm had a sensitivity of 89.5% and specificity of 100% for malignancy. The corresponding figures for lesions larger than 2 cm were 95.5% and 100%. The pneumothorax rate was 15% and 16%, respectively. Thus the accuracy and complication rate for CT-guided biopsy using a 20-gauge core was similar for large and small lesions.

A percutaneous approach may be used for mediastinal lesions also (**Figure 2**). Lymphoma is the cause of many mediastinal masses. If tissue can be obtained without the risks of even minimal thoracic surgery then this is of great benefit to the patient. In the case of lymphoma, it is also useful to know the subhistiotype for treatment. In a study group of 83 cases of mediastinal masses, core needle biopsy had an overall sensitivity of 81% for diagnosing lymphoma and the subhistiotype was accurately assessed in 93% of these (Zinzani *et al*., 1999). This study concluded that core needle biopsy is an effective and safe method of diagnosing mediastinal lymphoma, with a high likelihood of determining the subtype and thus defining subsequent specific treatment. A similar study reported a diagnostic success rate of 71.5% for core biopsy in diagnosing mediastinal lymphoma, with no complications reported (Sklair-Levy *et al*., 2000). As a result, the authors recommended the technique as the initial diagnostic procedure of choice. In a review of all types of mediastinal tumours, an overall sensitivity of 91.9% has been reported (Greif *et al*., 1999) with sufficient tissue obtained for diagnosis in 88.6%. A pneumothorax rate of 11% was encountered during the study. These and many other studies prove the value and efficacy of the percutaneous approach in the thorax.

LIVER, SPLEEN AND KIDNEY

Excellent results have been reported for abdominal biopsies also. The liver is probably the organ most commonly biopsied when malignancy is suspected (**Figure 3**). Large studies have repeatedly shown that percutaneous biopsy is both safe and effective; 244 liver lesions underwent a FNAB as reported by Herszenyi *et al*. (1995), with a reported diagnostic sensitivity of 93% and specificity of 100% with only one complication, that of tumour seeding along the biopsy tract. This compares favourably with the sensitivity and specificity of 18-gauge needle biopsy reported by Yu *et al*. (1998). They found a sensitivity and specificity of 96.4% and 100% respectively for the detection of both benign and malignant liver lesions. In a direct comparison between FNAB and core needle biopsy on the diagnosis of focal liver lesions suspected to be malignant, 55 patients had both a FNAB and core needle biopsy. In 41 cases of 48 with proven malignancy, both biopsies were positive. In the remaining seven cases, only FNAB yielded malignant cells. The tumour grading of six adenocarcinomas was more accurate with the histological (core needle) sample. This study performed by Jacobsen



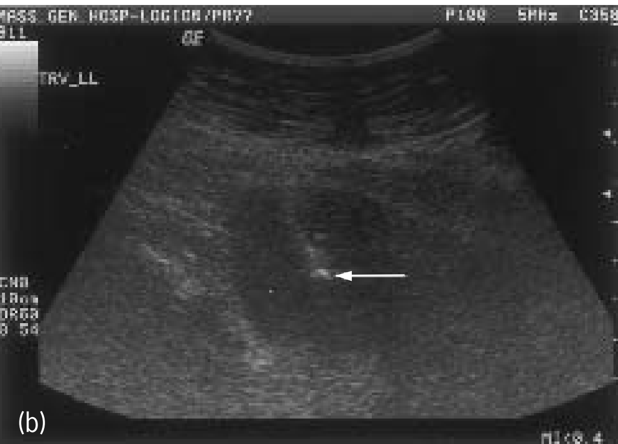


Figure 3 Large hypoechoic mass in the liver (a) in a 48-year-old female with the cirrhosis. In (b) the biopsy needle can be seen as an echogenic line lying within the mass. This was hepatocellular carcinoma.

et al. (1983) concluded that given that FNAB requires fewer prebiopsy precautions, core needle biopsy could be restricted to those cases in which FNAB does not yield sufficient information.

The ability of FNAB to detect malignancy in small liver lesions was examined by Middleton *et al*. (1997). They examined 30 liver lesions 1.5 cm or smaller which underwent an ultrasound-guided FNAB. The reported diagnostic accuracy rate for detecting malignancy was 93% using a 22-gauge needle. All these series, and many more, suggest that percutaneous biopsy is both safe and effective for liver lesions.

The spleen is amenable to biopsy when the need arises, although it is preferable to avoid splenic biopsies where possible given the highly vascular nature of the organ. The major indication for splenic biopsy is for a lesion which cannot be characterized on imaging, in patients in whom management will be altered significantly based on the result of the biopsy. Lymphoma, metastases and infection are the most frequently biopsied lesions. Twenty splenic lesions biopsied over a 2-year period were reported by Keogan *et al*. (1999). These were not all malignant lesions. A specific diagnosis was obtained in 16 cases (88.9%). No complications occurred. Diagnostic accuracy rates for other abdominal organs are similar (**Figures 5 and 6**). An 89% success rate has been reported for percutaneous biopsy of renal masses which proved to be renal cell carcinoma (Lechevallier *et al*., 2000) (**Figure 4**).

MUSCULOSKELETAL SYSTEM

The effectiveness of percutaneous biopsy of the musculoskeletal system has also been well established. In one review of 77 cases over a 24-month period, the diagnostic accuracy was found to be 83.8% overall. The accuracy for bone lesions was 85.5%, and for soft tissue lesions 76.9%. Accuracy was slightly better for lytic than sclerotic lesions (Hodge, 1999). A larger study involving 144 patients with primary bone tumours reported a diagnostic accuracy of 98.4% (Saifuddin *et al*., 2000) when needle biopsy was compared with surgical findings. FNAB has also been found to be accurate in the diagnosis of metastatic bone tumours (Collins *et al*., 1998).

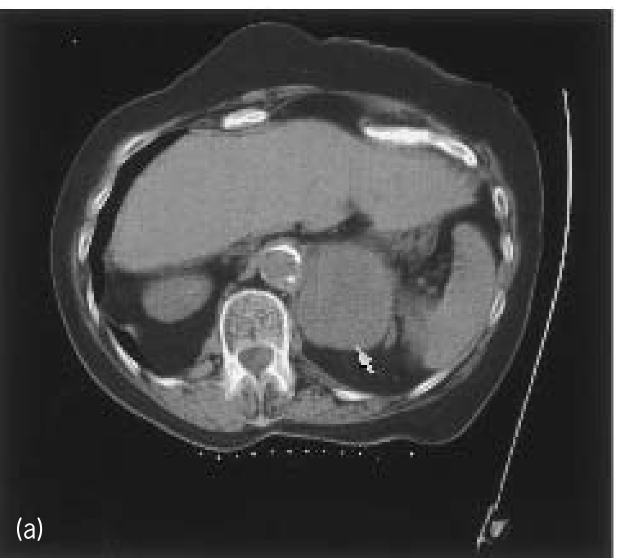
MISCELLANEOUS SITES

A percutaneous approach may even be used in the diagnosis of colonic carcinoma. The technique may be used following a negative colonoscopic biopsy or in cases of an inaccessible lesion on colonoscopy. Ultrasound is used for image guidance. One series reported a sensitivity for FNAB of 91.8% with a specificity of 100% in the diagnosis of colonic lesions when compared with surgical findings (Javid *et al*., 1999). In this series of 50 patients, only two minor complications were reported, supporting the utility, in certain cases, of ultrasound-guided percutaneous biopsy.

Percutaneous biopsy is also used with increasing frequency in diagnosing breast carcinoma. One difficulty with this lies in deciding how many cores should be obtained at the time of the procedure to make an accurate diagnosis without resorting to a second procedure. Breast biopsies are performed either under ultrasound guidance or stereotactically with radiographic control. The sensitivity for detecting malignancy has been reported as 84.3% when two cores are obtained. This rises to 97.7% when six cores are taken (Rich *et al*., 1999).

CONCLUSION

Percutaneous biopsy using either fine-needle aspiration or core biopsy is a straightforward technique which



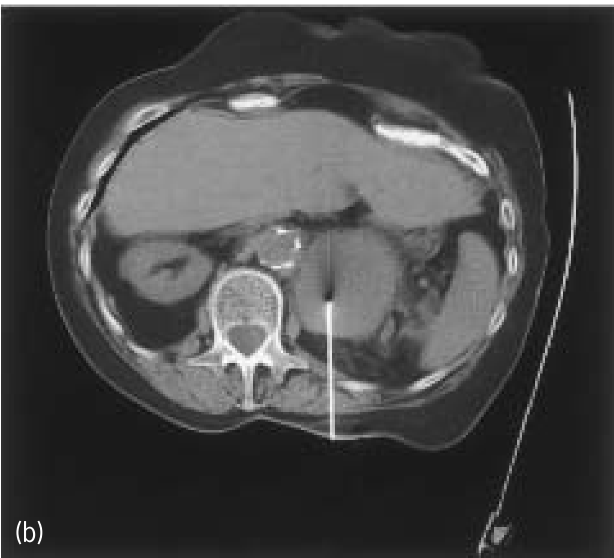
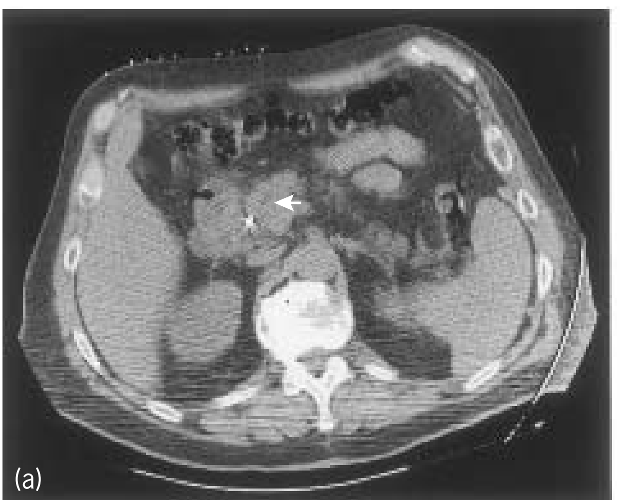


Figure 4 (a) A large, well-defined, soft tissue mass in the region of the upper pole of the left kidney in a 55-year-old female (arrow). Note the marking grid on the patient's skin. (b) An ideally located needle tip for good core biopsies.



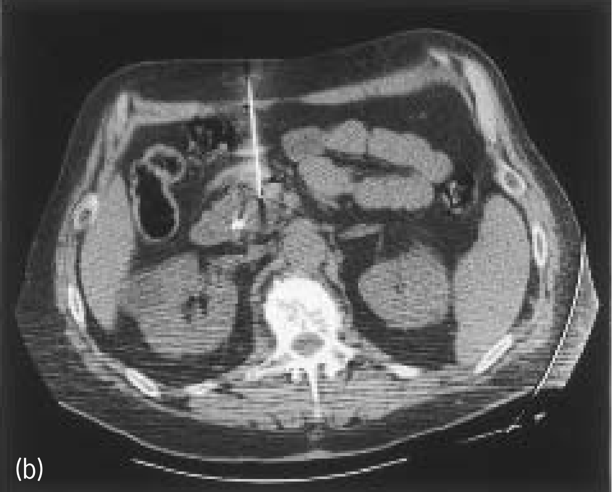


Figure 5 (a) A CT scan showing an irregular low attenuation mass in the head of the pancreas (arrow) surrounding a previously placed biliary stent in a 68-year-old male. (b) The needle within the mass. Puncturing the bowel with a 20-gauge needle is not associated with an increased complication rate.

may be used for almost any organ. The image guidance modalities are chosen according to local availability. The success rates are excellent and serious complication rates are low. There are few contraindications to the procedure, as demonstrated by the widespread use of percutaneous biopsy. Allowing for slight variations and tricks, the procedure is similar for most organs and modalities, which makes mastering the technique easier. As a tool for aiding both the initial diagnosis of malignancy and for assessing evidence of either recurrence or metastases, percutaneous biopsy has become invaluable.

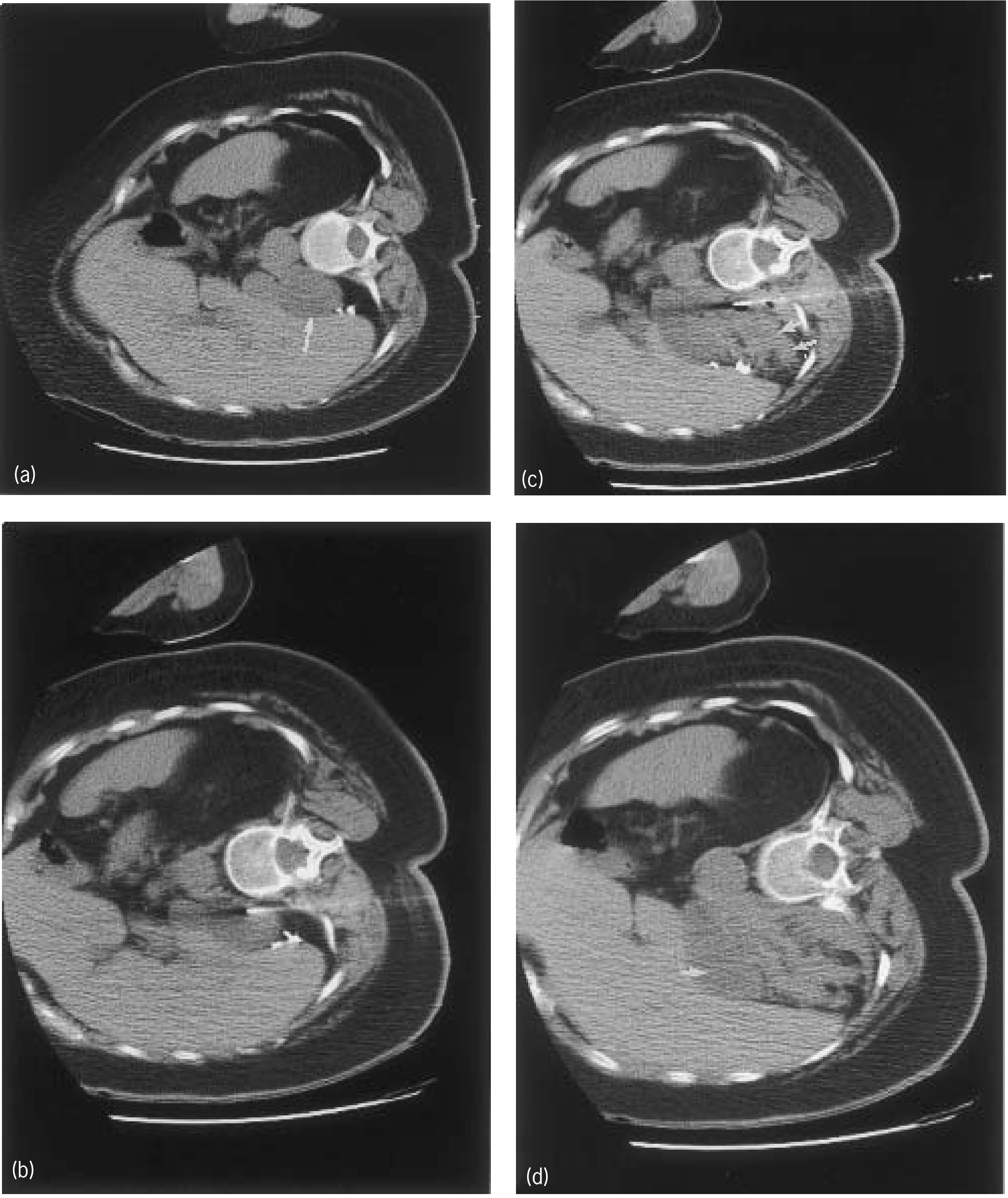


Figure 6 (a) Soft tissue mass in the right renal fossa (arrow) found on a surveillance CT following right nephrectomy for RCC. The decubitus position (right side down) decreases the risk of crossing the pleura. (b) Peri-procedural CT image demonstrating excellent needle location. This lesion proved to be an adrenal metastasis from the RCC. (c) CT image during the same procedure showing retroperitoneal haemorrhage postbiopsy. (d) Postprocedural CT scan demonstrating a large retroperitoneal haemorrhage (arrow) displacing the adrenal mass anteriorly (arrow). The patient was observed overnight and no intervention was required.

REFERENCES

Bhagat, V. J., *et al*. (1999). Percutaneous biopsy of a posterior mediastinal mass through an extrapleural window created with dilute contrast. *Journal of Thoracic Imaging*, **14**, 99–100.

Collins, B. T., *et al*. (1998). Fine needle aspiration biopsy of recurrent and metastatic osteosarcoma. *Acta Cytologica*, **42**, 357–361.

Greif, J., *et al*. (1999). Percutaneous core needle biopsy in the diagnosis of mediastinal tumors. *Lung Cancer*, **25**, 169–173.

Herszenyi, L., *et al*. (1995). Ultrasound guided fine-needle aspiration biopsy in the diagnosis of hepatocellular carcinoma. *Orvosi Hetilap*, **136**, 1545–1549.

Hodge, J. C., (1999). Percutaneous biopsy of the musculoskeletal system: a review of 77 cases. *Canadian Association of Radiologists Journal*, **50**, 121–125.

Jacobsen, G. K., *et al*. (1983). Coarse needle biopsy versus fine needle aspiration biopsy in the diagnosis of focal lesions of the liver. Ultrasonically guided needle biopsy in suspected hepatic malignancy. *Acta Cytologica*, **27**, 152–156.

Javid, G., *et al*. (1999). Percutaneous sonography-guided fine needle aspiration biopsy of colonoscopic biopsy-negative colonic lesions. *Indian Journal of Gastroenterology*, **18**, 146–148.

Karampekios, S., *et al*. (1998). Artificial paravertebral widening for percutaneous CT-guided adrenal biopsy. *Journal of Computer Assisted Tomography*, **22**, 308–310.

Keogan, M. T., *et al*. (1999). Imaging-guided percutaneous biopsy of focal splenic lesions: update on safety and effectiveness. *American Journal of Roentgenology*, **172**, 933–937.

Laurant, F., *et al*. (2000). CT-guided transthoracic needle biopsy of pulmonary nodules smaller than 20 mm: results with an automated 20-gauge coaxial cutting needle. *Clinical Radiology*, **55**, 281–287.

Lechevallier, E., *et al*. (2000). Fine-needle percutaneous biopsy of renal masses with helical CT guidance. *Radiology*, **216**, 506–510.

Middleton, W. D., *et al*. (1997). Small (1.5 cm or less) liver metastases: US-guided biopsy. *Radiology*, **205**, 729–732.

Rich, P. M., *et al*. (1999). Stereotactic 14G core biopsy of non-palpable breast cancer: what is the relationship between the number of core samples taken and the sensitivity for detection of malignancy? *Clinical Radiology*, **54**, 384–389.

Saifuddin, A., *et al*. (2000). Ultrasound-guided needle biopsy of primary bone tumors. *Journal of Bone and Joint Surgery, British Volume*, **82**, 50–54.

Sklair-Levy, M., *et al*. (2000). CT-guided core needle biopsy in the diagnosis of mediastinal lymphoma. *European Radiology*, **10**, 714–718.

Subhannachart, P., *et al*. (1999). Percutaneous transthoracic needle aspiration biopsy of localized lung lesions under fluoroscopic or ultrasound guidance. *Journal of the Medical Association of Thailand*, **82**, 268–274.

Swischuk, J. L., *et al*. (1998). Percutaneous transthoracic needle biopsy of the lung: review of 651 lesions. *Journal of Vascular Intervention and Radiology*, **9**, 347–352.

Yu, S. C., *et al*. (1998). Percutaneous biopsy of small hepatic lesions using an 18 gauge automated needle. *British Journal of Radiology*, **71**, 621–624.

Zinzani, P. L., *et al*. (1999). Core needle biopsy is effective in the initial diagnosis of mediastinal lymphoma. *Haematologica*, **84**, 600–603.

FURTHER READING

Castaneda-Zuniga, W. R. and Murthy Tadavarthy, S. (1992). *Interventional Radiology*, 2nd edn, Vol. 2. Chap. 20 (Williams and Wilkins, Baltimore).

Dondelinger, R. F., *et al*. (1990). *Interventional Radiology* (Georg Thieme, Stuttgart).

Ring, E. J. and McLean, G. K. (1981). *Interventional Radiology: Principles and Techniques.* Chap. 5 (Little-Browne, Boston).

Chapter 82

Transcatheter Therapy

Michael A. Bettmann
Dartmouth Medical School, Hanover, NH, USA

CONTENTS

- Introduction
- Basic Principles
- Transcatheter Therapy of Liver Malignancies
- Indications for Liver Chemoembolization
- Contraindications and Complications
- Other Uses and Techniques
- Conclusion

INTRODUCTION

Transcatheter therapy has an innate logical appeal. The basic aim of such therapy is to deliver one or more chemotherapeutic agents directly to a tumour through the vascular tree. If possible, the therapy is also delivered in such a manner as to make it remain in the treated tumour for relatively long periods of time. This general aim has been approached in various ways, for many years. For example, one approach to the treatment of colorectal carcinoma metastatic to the liver has been to place a catheter in the hepatic artery surgically, with a port implanted subcutaneously (Kemeny *et al*., 1999). This port can then be utilized for continuous or intermittent infusion of chemotherapy directly into the liver. Others have proposed placement of indwelling arterial catheters radiographically, via the upper or lower extremity arteries, for the same purpose. This approach has the advantage of being less invasive than surgical placement, with fluoroscopic guidance for accurate placement. The focus of this chapter, however, will be on transcatheter, short-term delivery of chemotherapeutic agents, a method which has been widely used throughout the world. This approach is usually referred to as 'chemoembolization.'

BASIC PRINCIPLES

Chemoembolization is performed starting with placement through the skin of a catheter, a small plastic tube, into a peripheral artery. Using the technique developed by Seldinger, a Swedish radiologist, in 1954, access to the artery is gained by advancing a needle through the skin. Through the centre of the needle, a guidewire is placed. The needle is then removed and the catheter is placed directly into the artery over the wire. The usual access to the artery is gained via the femoral vein in the groin, but any artery which is relatively superficial, such as an artery in the axilla or upper arm, can also be utilized. The catheter is then directed selectively into the vessels supplying the tumour to be treated (**Figure 1**). This is accomplished by utilizing catheters and guidewires of various shapes, combined with fluoroscopic, real-time monitoring. After the appropriate arterial position has been selected, the chemotherapeutic agent is delivered. In general, this consists of infusion over a short period, e.g. 5–10 min. Longer infusions, with the use of rescue medications, have also been used (Kemeny *et al*., 1999). After treatment, the catheter is removed and haemostasis is achieved. Ancillary therapy is generally employed to prevent or to treat complications of the chemotherapeutic agent and the procedure, and most often, chemoembolization is done with hospitalization of the

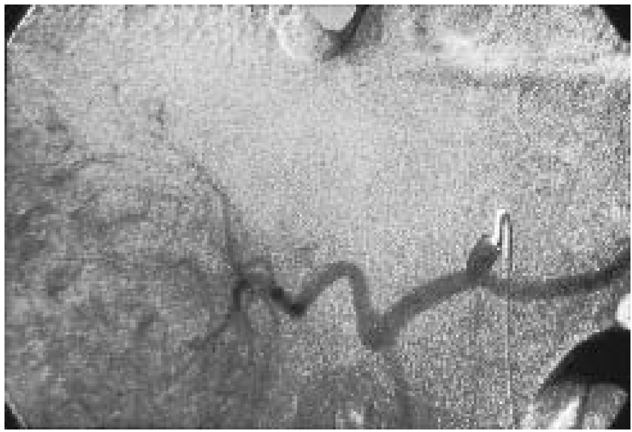


Figure 1 Catheter in place in coeliac axis. Contrast injection shows filling of the splenic artery and the hepatic artery. There is increased vascularity and draping of vessels around a large hepatocellular carcinoma in the right lobe of the liver.

patient. Additional treatment focuses, first, on pain control, since infusing material into an artery can lead to occlusion of that artery and resultant ischaemia. In part, the aim of this approach is to cause ischaemia and tumour necrosis. Opiates are often required for pain control, usually only for 24 h or less after the procedure. Other medications that are frequently used include corticosteroids, to help limit the inflammatory reaction to the treatment, and antiemetics. As with chemotherapy, nausea and vomiting are relatively frequent complications, but as with pain, tend to be limited to the first 24 h.

Chemoembolization in this fashion is often repeated one or more times, depending on the response to the treatment and the condition of the patient. Chemoembolization can be performed in essentially any area of the body, depending on the blood supply to the tumour and the organ that it involves. Transcatheter therapy has been widely used to treat head and neck tumours, renal malignancies, malignancies of the central nervous system and musculoskeletal lesions. The greatest experience, however, has been in the treatment of primary or metastatic liver malignancies, and that will be the major focus of this chapter.

The types of liver tumours treated by chemoembolization can be divided into three categories: (1) primary hepatocellular carcinoma, (2) metastatic carcinoma, primarily adenocarcinomas, and (3) neuroendocrine tumours (primary or metastatic).

TRANSCATHETER THERAPY OF LIVER MALIGNANCIES

Chemoembolization as a concept implies the use of both a chemotherapeutic agent and an embolic agent. That is, a specific anticancer drug is injected, in combination with a material that will block arterial inflow into the lesion, allowing direct delivery of the agent to the site at which it is needed, and then prolonged retention. In theory, as noted, there is also a beneficial effect in decreasing or eliminating blood flow to all or part of the tumour. To achieve these aims, various combinations have been utilized (Charusangavej, 1993). Among the chemotherapeutic agents that have been used are adriamycin, mitomycin, cisplatin, FUdR and 5-fluorouracil, and among the embolic agents used are poly(vinyl alcohol) (PVA) beads, absolute alcohol, Gelfoam and Ethiodol (Lipiodol). PVA beads are minute plastic particles available in various sizes, ranging from <100 μm to >1 mm. Several different types of calibrated plastic beads or particles are commercially available. PVA particles do not dissolve, so they cause permanent occlusion of whatever vessels into which they are instilled. In reality, some recurrence of blood supply generally occurs, either because of the recanalization through or around these beads or because of tumour neovascularity formation. Absolute alcohol is another agent which leads to permanent occlusion. It has been used widely as a primary treatment for hepatomas in Asia, both by infusion through a catheter and by percutaneous injection directly into tumour nodules (see the chapter on *Percutaneous Biopsy*). Absolute alcohol leads to occlusion of small vessels (capillaries and arterioles) and to direct tissue death by causing rapid and complete dehydration of cells into which it comes in contact, including blood cells, cells of the vessel wall and tumour. Gelfoam, or surgical gelatin sponge, is a material which dissolves and is resorbed from the vessels it occludes over the course of 2–4 weeks. This is a disadvantage, in that reflow to the treated area occurs, but an advantage in that the use of Gelfoam encourages prolonged retention of the chemotherapeutic agent in the tumour and the ability to re-treat the area fairly easily subsequently. Ethiodol is a radiographic contrast agent, one of the first, developed in the 1920s. It is a poppy seed oil derivative, a nonsoluble lipid. When infused into the hepatic artery or branches, it is thought to pass through the portal triad and be taken up into and then obstruct the portal vein capillaries. Over time, it is taken up by the reticuloendothelial system of the liver, and is then largely broken down and excreted. Empirically, it seems also to block hepatic arterial branches. It has the advantage, then, of relatively long-lasting obstruction of blood flow at a small vessel level, but without associated damage to larger vessels (**Figure 2**).

A major reason for both the interest in and the success of transcatheter therapy in the liver is that the liver has two blood supplies: the hepatic artery and the portal vein. Normally, the bulk of the blood flow to the liver is from the portal vein (about 70%) with the remainder from the hepatic artery. If the hepatic artery is acutely occluded, as by surgical ligation, portal flow increases to a sufficient level to maintain normal liver function. Further, liver malignancies receive the bulk of their blood supply from the hepatic artery. It is thought that some of the blood supply to the periphery of the tumour may be from the portal vein, which is not as readily accessible percutaneously as is the hepatic artery. This portal vein supply, in theory, is obstructed by Ethiodol. In most other organs, if the supplying arterial supply is obstructed, cell death results. In these organs, then, it is very important that only arterial flow to the tumour, and not to normal tissue, be embolized. This is often technically very difficult to achieve. In normal liver, however, more effective transcatheter therapy can be given, since liver adjacent to the tumour can be embolized without complications.

INDICATIONS FOR LIVER CHEMOEMBOLIZATION

Lesions which are most amenable to chemoembolization are those with a high degree of macroscopic neovascularity. Not only are such lesions technically fairly

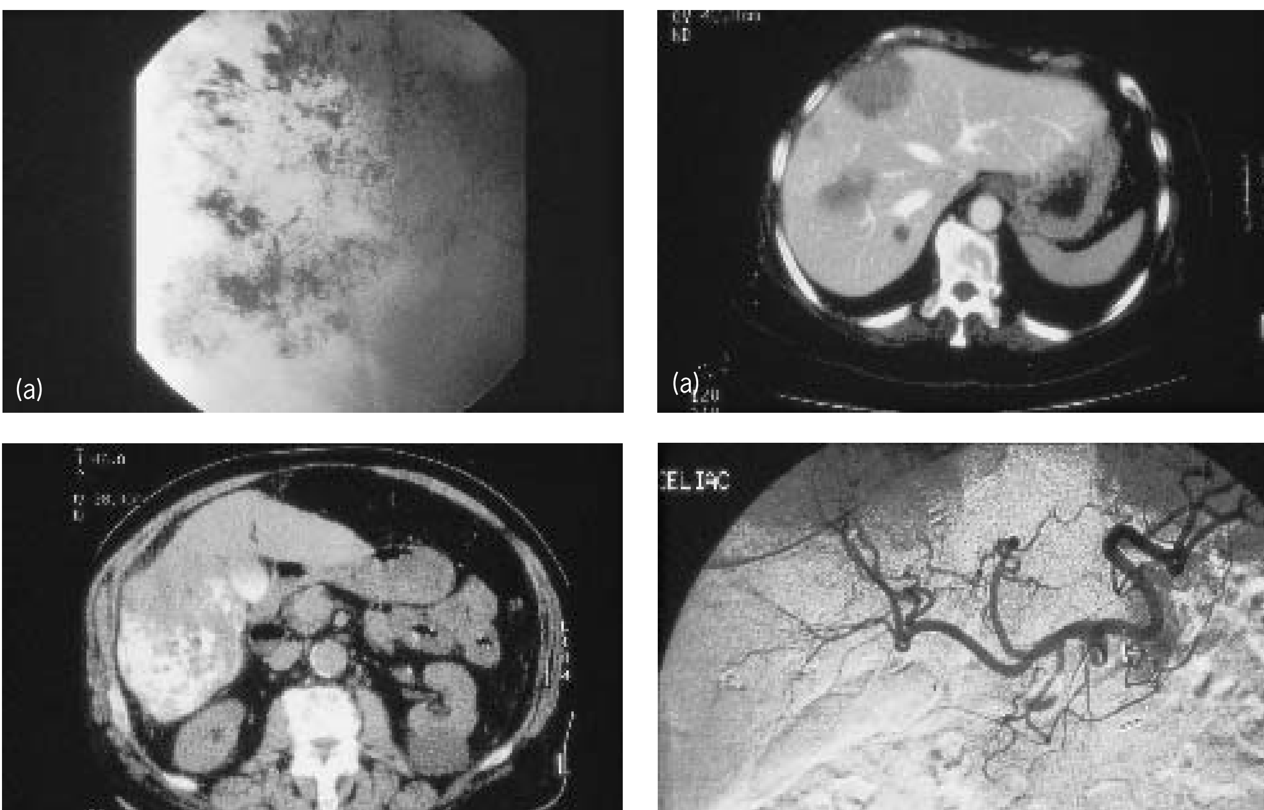


Figure 2 Same patient as in **Figure 1**. (a) Angiographic image immediately following chemoembolization, showing retention of Ethiodol in the tumour; (b) image from computed tomography performed 1 day later demonstrates retention of the oily contrast in much (but not all) of the tumour.

Figure 3 Patient with colon adenocarcinoma metastatic to the liver. (a) Image from contrast-enhanced CT shows four distinct lesions which have low density (relative to normal liver parenchyma, which shows contrast enhancement); (b) image from a coeliac axis arteriogram shows a faint blush (increased density) in the liver lesions, without visible neovascularity.

accessible to chemoembolization, but also treatment tends to be more effective than in lesions with minimal neovascularity. Clearly, all malignancies have a vascular supply, but in many, e.g. colorectal adenocarcinoma, lesions are relatively avascular on angiography (**Figure 3**). Primary hepatocellular carcinoma tends to be hypervascular angiographically (**Figures 1** and **4**). Similarly, most neuroendocrine tumours which are metastatic to the liver tend to be hypervascular (**Figure 5**).

The results of chemoembolization for hepatocellular carcinoma have been documented better than those for other lesions. One difficulty in interpreting results of the various studies, however, is that so many variables exist. These include the chemotherapeutic agent utilized, the embolic agent utilized, the number of treatments, the use of ancillary treatments such as radiation or systemic chemotherapy, the patient's liver function and the stage of the hepatocellular carcinoma (HCC). Most often, HCC that is treated by chemoembolization is considered non-operable. This is either because there are multiple lesions, with the presence of lesions in multiple lobes, because of the large size of the primary tumour, or because the patient's liver function is considered so poor that the operative risk, balanced against the likelihood of cure, is unacceptably high. At best, the cure rate for lesions considered surgically resectable is of the order of 30%, and mean survival is <18 months (Dusheiko *et al*., 1992; Colombo, 1993). Further, there is currently no effective systemic chemotherapeutic agent or regimen. For these reasons, in addition to its high incidence as a result of the high prevalence of hepatitis B or C, chemoembolization has been used widely for the treatment of HCC. Treatment is often repeated, at intervals ranging from 1 month to several years. One study demonstrated no statistically significant increase in survival when chemoembolization was compared with standard chemotherapy, although there was a trend toward prolonged survival (Groupe d'Etude et de Traitement du Carcinome Hepatocellulaire, 1995).

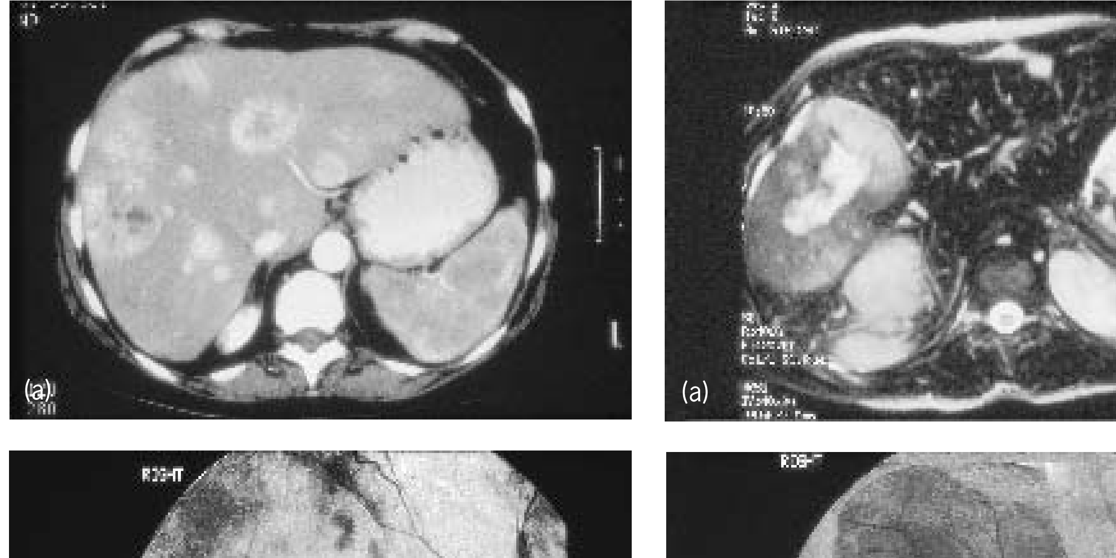


Figure 4 Elderly woman with hepatocellular carcinoma. (a) Image from a contrast-enhanced CT study shows numerous lesions with dense peripheral enhancement, greater than surrounding normal liver; (b) corresponding image from selective hepatic angiography shows dense lesions with marked neovascularity (tumour-induced, somewhat disordered appearing new vessels).

Other series have shown median survival of up to 2 years in patients treated with chemoembolization, a result which certainly rivals the efficacy of surgery in the relatively small proportion of lesions that are considered operable (Colombo, 1993; Simonetti *et al*., 1997). Another small study examined the effectiveness of chemoembolization with and without Ethiodol and gelatin sponge particles and demonstrated improved efficacy with the combination therapy, consisting of a chemotherapeutic agent, gelfoam and Ethiodol (Hatanaka *et al*., 1995). Another alternative is to combine embolization with percutaneous ablative techniques (Pacella *et al*., 2001). Over the last few years, with the emergence of radiofrequency ablation (see the chapter on *Molecular Mechanisms of Radiotherapy*), a combination of chemoembolization and percutaneous radiofrequency has been employed for discrete lesions, generally those ≤ 5 cm in overall size, even if multiple.

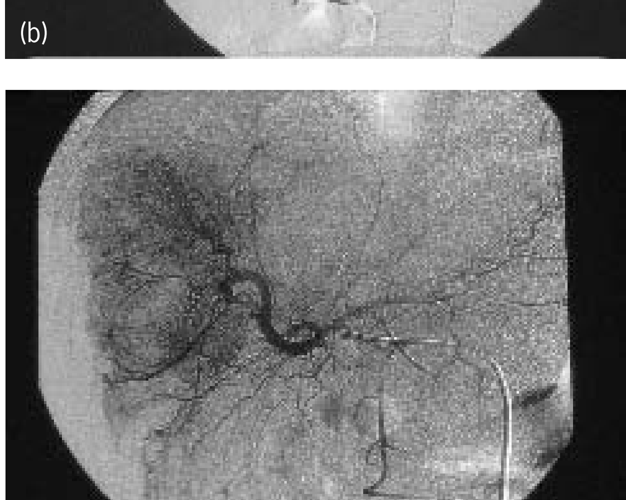


Figure 5 26-year-old woman with neuroendocrine tumour involving the liver, with unknown primary. (a) Contrast-enhanced MRI scan shows signal enhancement in a large lesion in the right lobe of the liver, with markedly increased signal enhancement centrally, suggestive of necrosis; (b) Selective arteriogram in right hepatic artery shows dense hypervascularity of this lesion; (c) repeat arteriogram 4 months after initial chemoembolization shows that this lesion is substantially smaller and less hypervascular. Repeat chemoembolization at this time led to a further marked decrease in the size of this and other liver lesions.

As HCC is often well circumscribed and hypervascular angiographically, it is reasonable to hypothesize that chemoembolization, particularly in combination with a percutaneous ablative technique such as radiofrequency, will be curative for single lesions. To date, however, although such an approach has been used, its true efficacy in terms of cure or improved survival has not been definitively evaluated.

Chemoembolization for the treatment of neuroendocrine tumours, perhaps because of the lesser frequency of such lesions, is less well documented. In general, all neuroendocrine tumours, including carcinoid, gastrinoma, insulinoma and pheochromocytoma, can be treated. It is relatively common for neuroendocrine tumours to present with widespread liver involvement, with multiple small lesions. Again, there is no widely effective systemic chemotherapy, although octreotide acetate may be useful, particularly in controlling symptoms. In general, patients who are treated appear to be more likely to respond favourably if the primary tumour is identified and removed. Survival for 2–5 years or more is not uncommon in such patients, even with widespread liver metastases, but again, chemoembolization cannot be considered curative. In one series, mean survival was ~4 years from the last chemoembolization (Stokes *et al*., 1993). It is relatively common for such lesions to shrink markedly after chemoembolization (**Figure 5**) but the occurrence of additional lesions and damage to the normal liver, as well as decreased vascular accessibility due to occlusion by the treatment, may limit the option of retreatment.

Many types of metastatic lesions have been treated by chemoembolization, either with one or more chemoembolizations or in combination with other therapy, such as percutaneous radiofrequency ablation, surgical resection or systemic chemotherapy. Some success has been reported in the treatment of highly lethal metastatic ocular melanoma, and there has also been some success in the treatment of GI stromal cell tumours (Soulen, 1997). As noted above, treatment is realistically directed at prolonging life rather than achieving a cure. The greatest experience to date with lesions other than HCC has been in the treatment of metastatic colorectal carcinoma. Unlike with more vascular lesions, such tumours tend not to shrink following chemoembolization. It is common, however, for such lesions to stabilize. Since most patients who undergo chemoembolization for metastatic colorectal carcinoma have multiple liver metastases which make them unresectable surgically or have metastatic disease elsewhere in the body, it is likely that survival is not prolonged by chemoembolization as currently practised. It is likely, however, that such treatment does prolong good quality of life. This approach is most appropriate, therefore, in patients who are unresponsive to standard systemic chemotherapy but remain in fairly good general condition, and often is considered after failure of first-line systemic chemotherapy.

CONTRAINDICATIONS AND COMPLICATIONS

As indicated above, chemoembolization should not be performed in certain patients, primarily those who are unlikely to gain substantial benefit from the procedure, in terms of prolongation of life or improvement or stabilization of quality of life. Thus patients who have far advanced tumours with failing liver function are not good candidates. At this point, it is not clear whether or not there are certain types of tumours which are so unlikely to benefit that the discomfort and other risks outweigh potential benefits in quality of life or prolongation of life. Chemoembolization of liver lesions can be considered in any otherwise suitable patients who are in sufficiently good general health (activity level, liver and renal function) and do not have other good options for treatment.

There are, however, several risk factors which are strong relative contraindications (Chung *et al*., 1996; Sakamoto *et al*., 1998). Patients with markedly compromised liver function are far more likely than those with normal liver function to develop progressive and even fatal liver failure. As a general rule, patients with a significant elevation in bilirubin should not undergo chemoembolization, although some of these patients may benefit from embolization with Ethiodol alone. This treatment is unlikely to damage liver function further if Ethiodol is injected selectively into the tumour. This is probably not true if a chemotherapy agent is used; even if injected selectively, some of the drug is likely to leak out of the tumour and be delivered systemically and to the liver as a whole (Raoul *et al*., 1992). Similarly, any patient with an obstructed portal vein should be treated with great caution. As noted, the principle of chemoembolization in the liver is that the tumour will be deprived of its blood supply, the hepatic artery, but normal liver will be preserved since it is also supplied by the portal vein. In the absence of portal vein blood flow, embolic occlusions of portions of the hepatic artery may lead to focal hepatic infarcts or to hepatic failure. Although in theory chemoembolization is performed solely to the arteries which supply the tumour, in practice such precise treatment is not usually feasible. This is because of multiple feeding arteries to the tumour(s) as well as technical limitations in placing the catheter sufficiently selectively.

The usual precautions with chemotherapy agents and with X-ray contrast agents must be observed. The dose of the particular agent used must be modified based on prior treatment and the individual's cardiac, renal and hepatic function. This procedure relies on localization of vessels and lesions by the injection of contrast agents. In patients with markedly compromised renal function, these contrast agents may lead to renal failure, even requiring dialysis (Tepel *et al*., 2000). Again, the risk/benefit ratio must be carefully assessed.

OTHER USES AND TECHNIQUES

As noted, selective high-dose arterial chemoembolization with systemic rescue has been used with some success in the treatment of head and neck cancer. Embolization without the addition of chemotherapy has been shown to be valuable in the treatment of renal cell carcinoma, primarily to control bleeding into the urinary tract, to diminish pain caused by a large vascular tumour or to decrease bleeding interoperatively and thus facilitate surgery. This is usually done using a permanent occlusive agent, such as PVA or absolute alcohol, infused as selectively as possible into the tumour. The reason why no chemotherapeutic agent is given is that none are thought to be effective, and they may add toxicity. The use of indwelling arterial catheters, with ports, is likely to prove increasingly useful. Such arterial catheters allow long-term percutaneous access and, if indicated, continuous infusions, and they can be safely placed on an outpatient basis.

CONCLUSION

A variety of transcatheter methods have been in use for many years for the selective treatment of tumours and organs. These techniques are most widely utilized currently for the treatment of primary and secondary lesions of the liver. Their use is also promising for other areas. Many questions remain regarding the specific efficacy of chemoembolization, because of the wide variation in the drugs utilized, the stage of tumour treated and the embolic agents employed. Over the next few years, there is likely to be increased use of transcatheter therapies with specifically targeted agents, such as genetic vectors and specific antitumour antibodies. As the ability to treat cancers selectively improves, the use of transcatheter approaches to deliver such therapy selectively to tumours is likely to increase also.

REFERENCES

Charusangavej, C. (1993). Chemoembolization of liver tumors. *Seminars in Intervention Radiology*, **10**, 150–160.

Chung, J. W., *et al*. (1996). Hepatic tumors: predisposing factors for complications of transcatheter oily chemoembolization. *Radiology*, **198**, 33–40.

Colombo, M. (1993). Hepatocellular carcinoma in cirrhotics. *Seminars in Liver Disease*, **13**, 374–383.

Dusheiko, G. M., *et al*. (1992). Treatment of small hepatocellular carcinomas. *Lancet*, **340**, 285–288.

Groupe d'Etude et de Traitement du Carcinome Hepatocellulaire (1995). A comparison of Lipodol chemoembolization and conservative treatment for unresectable hepatocellular carcinoma. *New England Journal of Medicine*, **332**, 1256–1261.

Hatanaka, Y., *et al*. (1995). Unresectable hepatocellular carcinoma: analysis of prognostic factors in transcatheter management. *Radiology*, **195**, 747–752.

Kemeny, N., *et al*. (1999). Hepatic artery infusion of chemotherapy after resection of hepatic metastases from colorectal cancer. *New England Journal of Medicine*, **341**, 2039–2048.

Pacella, C., *et al*. (2000). Hepatocellular carcinoma: long-term results of combined treatment with laser thermal ablation and transcatheter arterial chemoembolization. *Radiology*, **219**, 669–678.

Raoul, J. L., *et al*. (1992). Chemoembolization of hepatocellular carcinomas. A study of the biodistribution and pharmacokinetics of doxorubicin. *Cancer*, **70**, 585–590.

Sakamoto, I., *et al*. (1998). Complications associated with transcatheter arterial embolization for hepatic tumors. *Radiographics*, **18**, 605–619.

Simonetti, R. G., *et al*. (1997). Treatment of hepatocellular carcinoma: a systematic review of randomized controlled trials. *Annual Review of Oncology*, **8**, 117–136.

Soulen, M. C. (1997). Chemoembolization of hepatic malignancies. *Seminars in Intervention Radiology*, **14**, 305–311.

Stokes, K. R., *et al*. (1993). Hepatic artery chemoembolization for metastatic endocrine tumors. *Journal of Vascular Intervention Radiology*, **4**, 341–345.

Tepel, M., *et al*. (2000). Prevention of radiographic-contrast-agent-induced reductions in renal function by acetylcysteine. *New England Journal of Medicine*, **343**, 180–184.

Chapter 83

Direct Percutaneous Tumour Therapy

Gregory E. Blackman, Thomas Casciani and Michael A. Bettmann
Dartmouth–Hitchcock Medical Center, Lebanon, NH, USA

CONTENTS

- Introduction
- Target Sites for PAT
- Goals of PAT
- Imaging Related to PAT
- Specific Forms of PAT
- Conclusion

INTRODUCTION

With the introduction, wide acceptance and refinement of medical imaging techniques over the past 40 years, the approach to diagnosis and treatment of malignancy has significantly changed. Computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound in particular have tremendously improved the ability to detect and characterize many malignancies, allowing for earlier diagnosis of primary and metastatic tumours. In addition, less invasive forms of therapy are made possible by imaging guidance.

Percutaneous treatment of malignancy utilizes imaging for the delivery of ablative agents. Percutaneous ablative therapy (PAT) includes thermal ablation and chemical ablation. Thermal ablation currently includes radiofrequency ablation (RFA), interstitial laser photocoagulation, microwave coagulation, cryotherapy and instillation of heated fluids (Scudamore *et al*., 1997; Yoon *et al*., 1999; Dodd *et al*., 2000). Chemical ablation currently includes instillation of absolute alcohol (95–99.5% ethanol) and acetic acid (Shiina *et al*., 1987; Livraghi *et al*., 1995; Liang *et al*., 2000).

TARGET SITES FOR PAT

The bulk of published literature on PAT is devoted to interventions within the liver. The liver is an attractive target not only because of its frequent involvement by metastatic disease, but also because of its large size, which provides organ reserve and frequently provides access to the tumour through normal tissue. Multiple tumour foci, large tumour size, and location next to sensitive areas such as the diaphragm, porta hepatis or liver capsule provide technical challenges.

PAT has also been used to treat malignancies within the kidney, lung, breast, adrenal glands, prostate, pancreas, head and neck and brain (Anzai *et al*., 1995; Zlotta *et al*., 1998; Jeffrey *et al*., 1999; Mack and Vogl, 1999; Merkle *et al*., 1999a,b; Polascik *et al*., 1999; Dupuy *et al*., 2000; Shibata *et al*., 2000). Successful treatment of benign conditions, such as autonomously functioning thyroid nodules and symptomatic vertebral body haemangiomas, has also been reported (Doppman *et al*., 2000; Tarantino *et al*., 2000).

Results of interventions within one organ cannot be generalized to interventions elsewhere, given differences in tissue characteristics. A comparison of *in vivo* RFA within calf liver versus calf muscle showed a 95% greater average diameter of ablation within muscle than within liver (Goldberg, 1999). It will be important in the future to verify separately the effectiveness of various forms of PAT within specific target tissues.

GOALS OF PAT

Currently, the primary goal of PAT is palliation for patients who are not candidates for curative surgery. These patients are usually categorized as unresectable because of unfavourable location and extent of tumour, or general debility compounding the risks of major surgery. Local ablative therapy is employed to shrink tumours, or to slow or halt tumour growth, usually in the liver. Given that hepatocellular carcinoma (HCC) is the most common malignancy in the world and that 20–25% of all cancer deaths can be related to hepatic metastases, there is a large population that may benefit from PAT (Stone and Cady, 1994; D'Agostino and Solinas, 1995).

The early results for PAT reported in the literature suggest that curative therapy is possible, especially with

further refinements in technology and technique. While there is no clear consensus on the largest tumour amenable to PAT, tumour size is a critical factor in treatment planning. If tumour diameter is larger than the diameter of the zone of ablation, multiple overlapping treatments are necessary and the possibility of persistently viable tumour cells increases. Precise probe, or needle placement, reliable delivery of ablative agent and rigorous imaging follow-up, with retreatment as indicated, are necessary to effect a cure (see **Figure 1**). It has been suggested that tumours within cirrhotic livers, typically seen in patients with HCC, may have enhanced tumour kill with thermal ablation due to an 'oven effect' caused by the cirrhotic tissue limiting heat diffusion away from the tumour (Livraghi *et al*., 1999). Complete necrosis in up to 71% of HCC 3.1–5.0 cm in diameter and 25% of those 5.1–9.5 cm in diameter has been reported with radiofrequency ablation (Livraghi *et al*., 2000). A series including microscopic evaluation of liver metastases, 2.0–4.0 cm in diameter, treated with RFA reported complete necrosis in eight of nine (89%) resected specimens (Scudamore *et al*., 1999).

Differences in equipment, technique, indications and the relatively small size of early series make the application of published data regarding PAT to individual patients somewhat difficult. However, it seems that effective local control is a reasonable expectation for patients with small- to medium-sized tumours in the liver. Patients with localized disease may deserve aggressive treatment, given the possibility of cure. Multi-modality treatment including surgery, chemotherapy, chemoembolization and radiation, in addition to local ablative therapy, should be considered.

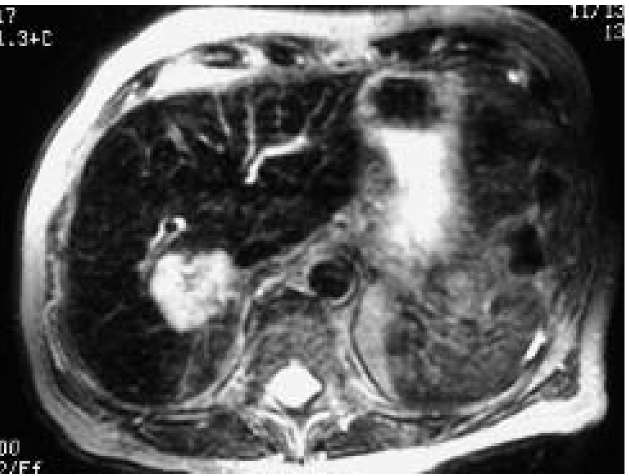


Figure 1 Axial MRI enhanced with intravenous ferumoxides showing a high signal lesion in the right lobe of the liver, which represents a metastasis from colon carcinoma.

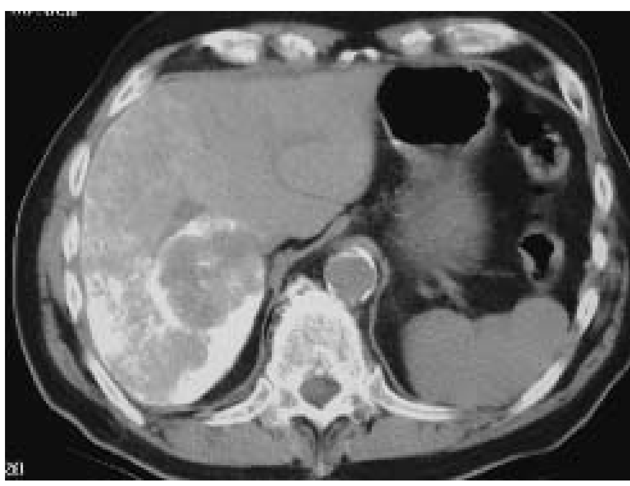


Figure 2 CT image showing the same lesion as in Figure 1. This image was obtained after ethiodol embolization of the right hepatic artery. The high-attenuation ethiodol is seen throughout the right lobe of the liver and outlines the relatively low-attenuation tumour.

IMAGING RELATED TO PAT

The main roles for imaging in PAT are initial tumour detection and characterization, PAT guidance, and follow-up. Contrast-enhanced CT and MRI are most often used for lesion detection and characterization, and follow-up (**Figure 1**). Thorough preprocedure imaging is imperative in order to select the appropriate therapy and to ensure that all areas of tumour are treated if possible. CT scanning following hepatic artery infusion of lipiodol (iodized oil) has been described as the most sensitive method of detecting HCC (90–97% sensitivity) (**Figure 2**) (Palma, 1998). This technique is also helpful in improving the conspicuity lesions treated under CT guidance (**Figure 3**).

The imaging modality used for PAT guidance is generally selected according to operator preference. Ultrasound is popular because of availability, ease of use and its ability to provide real-time monitoring. CT guidance is also common and MR guidance is possible with specialized equipment (particularly for injection therapy and laser ablation). The goals of imaging guidance are to ensure precise delivery of the ablative agent, to detect potential complications early and accurately, and to ensure as complete treatment as possible.

Postprocedure follow-up generally consists of contrast-enhanced CT or MR. Ultrasound is generally considered to be less reliable owing to the heterogeneous appearance of treated lesions and the need to demonstrate absence of lesion enhancement. Ultrasound contrast agents and Doppler sonography, however, may make this modality more useful for follow-up (Koito *et al*., 2000). Imaging is generally performed 1–2 months after treatment and then at 3–6 month intervals for 1 year, followed by more widely spaced intervals if the patient and lesions remain stable.

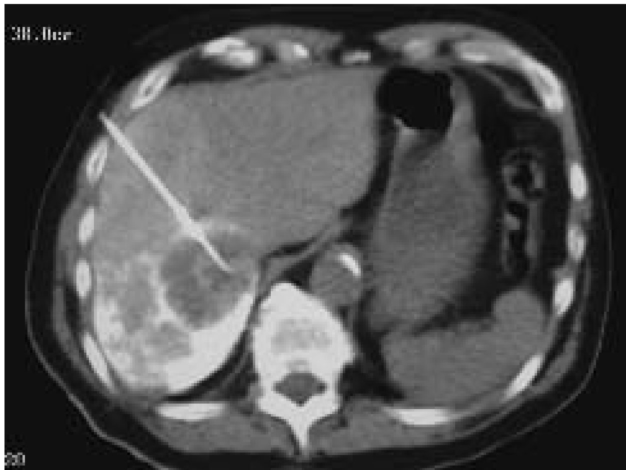


Figure 3 CT image obtained during RFA showing the RF probe within the tumour.

Completely necrotic lesions are of low signal intensity on T2 weighted MR and often have thin enhancing rims. Nodular areas of enhancement on MR or CT indicate residual tumour. Successfully treated lesions remain the same size or become smaller.

SPECIFIC FORMS OF PAT

Radiofrequency Ablation (RFA)

Radiofrequency electrocautery devices have been used throughout the last century to directly ablate exposed tissue, particularly to achieve haemostasis. Advances in technology have allowed for percutaneous, image-guided placement of electrode containing probes. These probes deliver an alternating current in the RF wave range (460 kHz), resulting in ionic agitation that produces heat and at temperatures above 60 °C, cell death occurs rapidly. The distance the generated heat is conducted with the tissue determines the zone of ablation. The power of the RF current, time of application, the impedance of the tissue and the size, number and spacing of the electrodes affect the zone of ablation. As tissue is ablated, the impedance of the tissue increases. Generators with increased power and pulsed current have been developed. Probes now may include multiple electrodes that, in some designs, have variable diameters of expansion, and in other designs are saline cooled. Saline cooling directly improves conduction and can limit charring of tissue, which increases impedance. Techniques to enhance the efficacy of PAT include preprocedure chemoembolization, concomitant injection of chemotherapeutic agents and intraoperative RFA combined with clamping the porta hepatis (Pringle manoeuvre). The Pringle manoeuvre and preprocedure chemoembolization are believed to increase the area of ablation by decreasing heat loss via flowing blood (heat sumping). The direct injection of chemotherapeutic agent as well as chemoembolization have the added benefit of targeted tumour toxicity (Goldberg *et al*., 2000).

The three currently available RFA devices are marketed by RITA Medical Systems (Mountain View, CA, USA), Radionics (Burlington, MA, USA) and Radiotherapeutics (Mountain View, CA, USA). All function on the basic principle of delivery of an alternating current via one or multiple electrodes inserted into the target area with repositioning or retreatment as needed based on lesion size, location and success of initial placement (Scudamore *et al*., 1997; Dodd *et al*., 2000). These three devices are designed to create zones of ablation 3–4.5 cm in diameter in the liver. In practice, however, ablation diameters as small as 1.2 cm, with ranges of 1.2–3.5 and 1.8–3.6 cm, have been reported with use of a single type of RF electrode (Radionics; Goldberg *et al*., 1998; Scudamore *et al*., 1999). These differences in ablation diameter are likely related to local factors such as blood flow and lesion density and underscore the need for precise imaging to determine if viable tumour remains. The small ablation diameters previously reported are also likely related to older electrode and generator designs, which have been improved upon.

The RITA device consists of a 15-gauge probe with seven or nine retractable electrodes at the needle tip that expand to 3–5 cm diameter in a ‘star burst’ array. It uses temperature as an endpoint, while the other two devices use tissue impedance. Its generator creates a 50–150 W alternating current and monitors the temperature at the tip of five electrodes and automatically diverts current when an electrode exceeds a preset temperature (e.g. 100 °C). The electrodes are retracted into the probe needle prior to placement, and are then deployed into the target tissue once ideal position is confirmed. Recently developed probes offer the option of increasing the diameter of the electrode array in a sequential fashion to allow for consecutive ablations at increasing diameters. The manufacturer recommends ablating large tumours in a stepwise fashion (e.g. first at 3 cm, then at 4 cm and finally at 5 cm) to ensure complete treatment.

The Radionics device utilizes 1–3 saline-cooled straight 17 F probes with 1–3-cm lengths of active electrode and a 100–200-W alternating current generator. A multiprobe array in a triangular configuration (cluster probe) offers the possibility of larger zones of ablation than the single probe model. In situations where the cluster probe is not feasible (e.g. intercostal approach), two parallel probes may be placed. Power from the generator is delivered in pulses that automatically adjust to recorded tissue impedance. Theoretically, current is directed away from ablated tissue in the target area that has high impedance toward areas of lower impedance resulting in uniform ablation. The generator shuts off when there is a sharp rise in tissue impedance, which is thought to indicate tissue necrosis. Saline

cooling of the electrodes represents a technical improvement from the original design. The Radionics and RITA devices may be used for electrocautery of the tract as the needle is withdrawn.

The Radiotherapeutics device has a multi-electrode 14-gauge probe that is similar to the RITA device in appearance and deployment characteristics and is available in three diameters. Ten electrodes expand into a 2.0-, 3.0- or 3.5-cm spherical umbrella-like configuration once the probe is in the target area. A probe with electrodes that expand to 5.0 cm is being developed (Dodd *et al*., 2000). The generator monitors both electrode temperature and tissue impedance. A pulsed alternating current is delivered to the target area. The generator is shut off when there is a sharp rise in tissue impedance.

Reported results with RFA have indicated major complications ranging from 0 to 10% and local control ranging from 48 to 98.2% (Scudamore *et al*., 1997, 1999; Goldberg *et al*., 1998, 2000; Curley *et al*., 1999; Livraghi *et al*., 1999, 2000; Rhim and Dodd, 1999; Dodd *et al*., 2000). The vast differences in local control are probably due to differences in reporting standards, length of follow-up, equipment differences and operator experience. Each of the three devices has compelling features or data to support their use, and recent design updates make it difficult to compare previous data with present performance. No single device has yet demonstrated clear superiority.

Percutaneous ethanol injection therapy (PEIT)

PEIT has been extensively studied, largely in Japan and Italy, with promising results and low complication rates (Shiina *et al*., 1990; Livraghi *et al*., 1995, 1999; Shibata *et al*., 2000). The relative simplicity of technique and availability of tools are attractive features. Needles are generally 20–22-gauge Chiba and spinal needles or 21-gauge needles with a closed conical tip and multiple terminal side holes (PEIT needle, Hakko, Tokyo, Japan, or Cliny needle, Yokohama, Japan). Absolute alcohol (95–99.5% ethanol) is injected in small aliquots (fractions of a millilitre) with real-time ultrasound guidance both at the time of needle placement and during injection. Usually, 1–8 mL are injected per treatment session with the volume dictated by the size of the lesion, imaging findings during injection and the patient's tolerance to the procedure. Injection is typically halted once pooling of ethanol is seen within the lesion (the ethanol no longer disappears rapidly from the needle tip) or if leakage of ethanol outside the target area is detected. Treatment sessions can be performed 2–3 times per week until the entire lesion has been treated. Therapeutic efficacy is assessed by follow-up imaging and clinical assessment (AFP and CEA assay).

Large-dose, 'single-session' treatment has been described with doses up to 165 mL (mean, 75 mL). However, general anaesthesia was required and the patients were medicated with 1000–1500 mg of i.v. fructose 1,6-diphosphate (Esafosfina; Biomedica Foscama, Ferentino, Italy) and 1200 mg of i.v. glutathione (TAD 600; Biomedica Foscama) to neutralize the systemic effects of the alcohol (Livraghi *et al*., 1995). Furthermore, two cases of prolonged hypotension, one of which resulted in death, were reported with injection of 35 and 80 mL respectively (Gelczer *et al*., 1998).

PEIT causes cell death by two mechanisms. The first is diffusion into cells, which results in dehydration and coagulation necrosis. The second is vascular occlusion caused by direct endothelial damage and platelet aggregation. PEIT is felt to be less effective in the treatment of liver metastases than for HCC. It has been suggested that HCC respond favourably due to hypervascularity and to the contrast in tissue density between hard cirrhotic tissue and softer HCC (Dodd *et al*., 2000).

In comparing PEIT with RFA in the treatment of HCC 3.0 cm or less in diameter, RFA was found to require fewer treatment sessions, 1.2 vs 4.8 sessions per tumour, and to have a higher rate of complete tumour necrosis, 90% (47/52) vs 80% (48/60) (D'Agostino and Solinas, 1995). Complications with PEIT are reportedly slightly less frequent than with RFA. However, most PEIT series report treatment of smaller lesions than those treated with RFA. A large series of PEIT treatment of sizeable HCC ($\geq$3–5 cm in diameter) reported a 1.7% rate of severe complications (10 patients) and one death (0.13%). There was a 27% rate of residual tumour in resected specimens (Livraghi *et al*., 1995).

Percutaneous Acetic Acid Injection

Percutaneous instillation of acetic acid is similar in principle and technique to PEIT. Small quantities of acetic acid are injected through a multihole infusion needle, placed into the target tissue with image guidance. Acetic acid appears to have stronger cytotoxic effects than ethanol, which may allow for larger ablation zones. Excellent local control of small lesions (1.5–3.0 cm in diameter) has been reported, with severe pain (11%), high fever (4%) and segmental wedge infarction (4%) noted as complications (Liang *et al*., 2000). Additional experience is needed to determine safety, efficacy and general treatment guidelines.

Interstitial Laser Photocoagulation (ILP)

ILP utilizes ND:YAG laser energy (approximate wavelength 1064 nm) delivered via a thin optical fibre placed through needles or guiding sheaths positioned within the target tissue, usually at 1–1.5-cm intervals. An MRI-compatible applicator kit (Somatex, Berlin, Germany) is available in 'conventional' and 'power' laser forms. The kits consist of a cannulation needle, guidewire, 7–9F

dilator or sheath and a 4–7F, 40-cm protective Teflon catheter. The conventional applicator delivers a maximum power of 6 W for up to 22 min and creates a diameter of necrosis of <2 cm. The power applicator is liquid cooled, can deliver up to 30 W for 14–23 minutes and creates a mean diameter of necrosis of 3.3 cm (Mack and Vogl, 1999; Vogl *et al*., 1999; Dodd *et al*., 2000).

Limited reports are available to determine the efficacy of ILP. One series employed contrast-enhanced CT, from 6 weeks to 3 months following treatment, to determine the degree of tumour necrosis. Complete necrosis of 21/35 lesions <4 cm was reported, but none of 20 lesions >4 cm showed complete necrosis (Amin *et al*., 1993). A larger series (659 lesions in 230 patients) did not report degree of necrosis, but instead reported median survival, which was significantly increased in treated patients compared with historical controls of untreated patients: median survival of 36.4 months compared with 10–21 months expected median survival (Vogl *et al*., 1999).

Percutaneous Microwave Coagulation Therapy (PMCT)

Several groups in Japan have reported on PMCT in the treatment of HCC (Seki *et al*., 1994; Murakami *et al*., 1995; Ohmoto *et al*., 1999). A 1.6-mm, 30-cm polytetrafluorethylene-coated bipolar electrode is passed through a needle guide into the target area and a 1.5- or 3.0-cm inner conductor delivers 2450-MHz microwaves at 60–150 W, usually for 60–120 s. Generally, ultrasound guidance is utilized and multiple overlapping treatments are administered per treatment session. Each application results in a relatively small diameter of ablation, approximately 2 cm. Follow-up contrast-enhanced CT or MRI is obtained and residual areas of viable tumour are retreated. Cited complications have been infrequent with one of nine patients in one series developing a pleural effusion and ascites (Murakami *et al*., 1995). A local sensation of heat during the ablation and postprocedure fever are common. Local lesion control has been reported in 56–100%. The need for multiple overlapping treatments due to the small diameter of ablation makes completeness of ablation less certain. Despite this theoretical concern, successful induction of complete necrosis has been reported in lesions up to 6.0 cm in diameter, using intraoperative microwave coagulation therapy (Hamazoe *et al*., 1995).

The efficacy of PMCT in hepatic metastases is unknown. However, a study that compared intraoperative microwave coagulation therapy with hepatic resection for the treatment of multiple (2–9) hepatic metastases found them to be of equal efficacy (Shibata *et al*., 2000). Thirty patients with multiple liver metastases, felt to be amenable to hepatic resection, went to laparotomy and were randomized to hepatectomy (16 patients) or sonographically guided intraoperative microwave coagulation (14 patients). The 1-, 2- and 3-year survival rates and mean survival times were 71%, 57%, 14% and 27 months, respectively, in the microwave group and 69%, 56%, 23% and 25 months in the hepatectomy group.

Cryosurgical Ablation (CSA)

It is difficult to compare CSA with percutaneous ablative techniques, because CSA is typically employed intraoperatively. CSA produces tissue destruction by creating focal subzero temperatures adjacent to the tip of a probe, resulting in cellular dehydration and protein denaturation. Probes of 9–12 mm are generally used, although smaller probes have been developed for laparoscopic surgery. Ultrasound is used to guide probe placement and monitor 'ice-ball' formation. Several freeze–thaw cycles, lasting 7–30 min, are applied with each application. Lesions up to 17 cm have been treated, but complications increase with increasing treatment volume. A variety of serious complications have been reported, including myoglobinuria, surface liver cracking, haemorrhage, biliary fistula, abscess, portal vein thrombosis and thrombocytopenia (Mahvi and Lee, 1999; Wallace *et al*., 1999).

CONCLUSION

Recent advances in imaging and percutaneous ablative techniques have expanded cancer therapy. PAT offers hope for improved palliation and possibly improved long-term survival. As with other forms of cancer therapy, PAT is most effective when employed in a setting of multidisciplinary collaboration. Experience with PAT continues to increase. However, controlled trials are needed to confirm the safety and efficacy of the different forms of PAT and to explore the role of PAT in various organs.

REFERENCES

Amin, Z., *et al*. (1993). Hepatic metastases: interstitial laser photocoagulation with real-time US monitoring and dynamic CT evaluation of treatment. *Radiology*, **187**, 339–347.

Anzai, Y., *et al*. (1995). Preliminary experience with MR-guided thermal ablation of brain tumors. *American Journal of Neuroradiology*, **16**, 39–52.

Curley, S. A., *et al*. (1999). Radiofrequency ablation of unresectable primary and metastatic hepatic malignancies: results in 123 patients. *Annals of Surgery*, **230**, 1–8.

D'Agostino, H. B. and Solinas, A. (1995). Percutaneous ablation therapy for hepatocellular carcinomas. *American Journal of Roentgenology*, **164**, 1165–1167.

Dodd, G. D., III, *et al*. (2000). Minimally invasive treatment of malignant hepatic tumors: at the threshold of a major breakthrough. *RadioGraphics*, **20**, 9–27.

Doppman, J. L., *et al*. (2000). Symptomatic vertebral hemangiomas: treatment by means of direct intralesional injection of ethanol. *Radiology*, **214**, 341–348.

Dupuy, D. E., *et al*. (2000). Percutaneous radiofrequency ablation of the malignancies in the lung. *American Journal of Roentgenology*, **174**, 57–59.

Gelczer, R. K., *et al*. (1998). Complications of percutaneous ethanol ablation. *Journal of Ultrasound Medicine*, **17**, 531–533.

Goldberg, S. N., *et al*. (1998). Ablation of liver tumors using percutaneous RF therapy. *American Journal of Roentgenology*, **170**, 1023–1028.

Goldberg, S. N. (1999). Percutaneous radiofrequency tissue ablation: optimization of pulsed-radiofrequency technique to increase coagulation necrosis. *Journal of Vascular Intervention Radiology*, **10**, 907.

Goldberg, S. N., *et al*. (2000). Thermal ablation therapy for focal malignancy: a unified approach to underlying principles, techniques, and diagnostic imaging guidance. *American Journal of Roentgenology*, **174**, 323–331.

Hamazoe, R., *et al*. (1995). Intraoperative microwave tissue coagulation as treatment for patients with nonresectable hepatocellular carcinoma. *Cancer*, **75**, 794–800.

Jeffrey, S. S., *et al*. (1999). Radiofrequency ablation of breast cancer. *Archives of Surgery*, **134**, 1064–1068.

Koito, K., *et al*. (2000). Power Doppler sonography: evaluation of hepatocellular carcinoma after treatment with transarterial embolization or percutaneous ethanol injection therapy. *American Journal of Roentgenology*, **174**, 337–341.

Liang, H. L., *et al*. (2000). Small heptocellular carcinoma: safety and efficacy of single high-dose percutaneous acetic acid injection for treatment. *Radiology*, **214**, 769–774.

Livraghi, T., *et al*. (1995). Hepatocellular carcinoma and cirrhosis in 746 patients: long-term results of percutaneous ethanol injection. *Radiology*, **197**, 101–108.

Livraghi, T., *et al*. (1999). Small hepatocellular carcinoma: treatment with radio-frequency ablation versus ethanol injection. *Radiology*, **210**, 655–661.

Livraghi, T., *et al*. (2000). Hepatocellular carcinoma: radiofrequency ablation of medium and large lesions. *Radiology*, **214**, 761–768.

Mack, M. G. and Vogl, T. J. (1999). Magnetic resonance-guided laser-induced thermotherapy of recurrent head and neck tumors. *Seminars in Intervention Radiology*, **16**, 63–70.

Mahvi, D. M. and Lee, F. T. (1999). Radiofrequency ablation of hepatic malignancies: is heat better than cold? *Annals of Surgery*, **230**, 9–11.

Merkle, E. M., *et al*. (1999a). MR-guided RF thermal ablation of the kidney in a porcine model. *American Journal of Roentgenology*, **173**, 645–651.

Merkle, E. M., *et al*. (1999b). MR imaging-guided radiofrequency thermal ablation in the pancreas in a porcine model with a modified clinical C-arm system. *Radiology*, **213**, 461–467.

Murakami, R., *et al*. (1995). Treatment of hepatocellular carcinoma: value of percutaneous microwave coagulation. *American Journal of Roentgenology*, **164**, 1159–1164.

Ohmoto, K., *et al*. (1999). Percutaneous microwave coagulation therapy for hepatocellular carcinoma located on the surface of the liver. *American Journal of Roentgenology*, **173**, 1231–1233.

Palma, L. D. (1998). Diagnostic imaging and interventional therapy of hepatocellular carcinoma. *British Journal of Radiology*, **71**, 808–818.

Polascik, T. J., *et al*. (1999). Ablation of renal tumor in a rabbit model with interstitial saline-augmented radiofrequency energy: preliminary report of a new technology. *Urology*, **53**, 465–470.

Rhim, H. and Dodd, G. D., III. (1999). Radiofrequency thermal ablation of liver tumors. *Journal of Clinical Ultrasound*, **27**, 221–229.

Scudamore, C. H., *et al*. (1997). Liver tumor ablation techniques. *Journal of Investigative Surgery*, **10**, 157–164.

Scudamore, C. H., *et al*. (1999). Radiofrequency ablation followed by resection of malignant liver tumors. *American Journal of Surgery*, **177**, 411–417.

Seki, T., *et al*. (1994). Ultrasonically guided percutaneous microwave coagulation therapy for small hepatocellular carcinoma. *Cancer*, **74**, 817–825.

Shiina, S., *et al*. (1987). Percutaneous ethanol injection in the treatment of liver neoplasms. *American Journal of Roentgenology*, **149**, 949–952.

Shiina, S., *et al*. (1990). Percutaneous ethanol injection therapy for neoplasms located on the surface of the liver. *American Journal of Roentgenology*, **155**, 507–509.

Shibata, T., *et al*. (2000). Percutaneous ethanol injection for treatment of adrenal metastasis from hepatocellular carcinoma. *American Journal of Roentgenology*, **174**, 333–335.

Shibata, T., *et al*. (2000). Microwave coagulation therapy for multiple hepatic metastases from colorectal carcinoma. *Cancer*, **89**, 276–284.

Stone, M. S. and Cady, B. (1994). Cancer of the liver, gallbladder and bile duct. In: McKenna, R. J. and Muphy, G. P. (eds), *Cancer Surgery*. 87–93 (J.B. Lippincott, Philadelphia).

Tarantino, L., *et al*. (2000). Percutaneous ethanol injection of large autonomous hyperfunctioning thyroid nodules. *Radiology*, **214**, 143–148.

Vogl, T. J., *et al*. (1999). Percutaneous laser-induced thermotherapy of malignant liver tumors. *Seminars in Intervention Radiology*, **16**, 3–12.

Wallace, J. R., *et al*. (1999). Cryotherapy extends the indications for treatment of colorectal liver metastases. *Surgery*, **126**, 766–771.

Yoon, H. K., *et al*. (1999). Percutaneous hot saline injection therapy: effectiveness in large hepatocellular carcinoma. *Journal of Vascular Intervention Radiology*, **10**, 477.

Zlotta, A. R., *et al*. (1998). Percutaneous transperineal radiofrequency ablation of prostate tumour: safety, feasibility and pathological effects on human prostate cancer. *British Journal of Urology*, **81**, 265–275.

FURTHER READING

Dodd, G. D., III, *et al*. (2000). Minimally invasive treatment of malignant hepatic tumors: at the threshold of a major breakthrough. *RadioGraphics*, **20**, 9–27.

Goldberg, S. N., *et al*. (2000). Thermal ablation therapy for focal malignancy: a unified approach to underlying principles, techniques, and diagnostic imaging guidance. *American Journal of Roentgenology*, **174**, 323–331.

Moradpou, D. and Wands, J. R. (1996). Hepatic oncogenesis. In: Zakim, D. and Boyer, T. D. (eds), *Hepatology: A Textbook of Liver Disease*, 3rd edn, **Vol. 2**. 1490–1512 (W.B. Saunders, Philadelphia).

Palma, L. D. (1998). Diagnostic imaging and interventional therapy of hepatocellular carcinoma. *British Journal of Radiology*, **71**, 808–818.

Sigurdson, E. R. and Cohen, A. M. (1994). Surgery for liver metastases. In: McKenna, R. J. and Murphy, G. P. (eds), *Cancer Surgery*. 95–104 (J.B. Lippincott, Philadelphia).

Stone, M. D. and Cady, B. (1994). Cancer of the liver, gallbladder and bile duct. In: McKenna, R. J. and Muphy, G. P. (eds), *Cancer Surgery*. 87–93 (J.B. Lippincott, Philadelphia).

Chapter 85

Drug Resistance and Reversal

Andrew Harbottle
University of Newcastle Upon Tyne, Newcastle Upon Tyne, UK

CONTENTS

- Introduction
- Transport Proteins
- Reversal of Multidrug Resistance

INTRODUCTION

One of the major obstacles to the ultimate success of cancer chemotherapy is the ability of malignant tumours to develop resistance to cytotoxic compounds and this can affect patients with a variety of blood cancers and solid tumours, including breast, ovarian, lung, and lower gastrointestinal tract cancers. Resistance may be to a single drug, e.g. resistance to methotrexate due to gene amplification of dihydrofolate reductase, or to a number of drugs. Clinical oncologists were the first to observe that cancers treated with a number of different anticancer drugs were able to develop cross-resistance to many other cytotoxic agents to which they had never been exposed, greatly decreasing the chance of curing these tumours with chemotherapy. In many cases, cells grown in culture from such multidrug-resistant tumours demonstrate similar patterns of resistance *in vitro* to those seen *in situ*. This observation suggests that multidrug resistance (MDR) is, in many cases, the result of heritable changes in cancer cells causing altered levels of specific proteins, or mutant proteins, which allow cancer cells to survive in the presence of many different cytotoxic agents. These genetic changes within the tumour cell population that confer drug resistance may affect cell cycle dynamics, uptake and efflux of drugs, cellular drug metabolism, susceptibility of cells to apoptosis, intracellular compartmentalization of drugs or repair of drug-induced damage (usually to DNA). Tumours usually consist of a mixed population of malignant cells, and as such multidrug resistance in tumours may develop owing to a selecting out of drug-resistant cells during the course of treatment, or indeed may be due to direct genetic changes within the population caused by exposure to the chemotherapeutic agents.

Multidrug resistance of tumours has been correlated to the expression of at least two molecular efflux pumps in tumour cell membranes, which actively expel chemotherapeutic drugs from the cell. The active efflux of drugs from the cell allows the cell to avoid the cytotoxic effects of the compounds. The two efflux pumps most commonly found in multidrug-resistant tumours are P-glycoprotein and the multidrug resistance-associated protein (MRP). Although most research has concentrated on efflux transport mechanisms of resistance, other possible mechanisms of resistance have been suggested including the expression of the metabolizing enzyme glutathione *S*-transferase P1, which is often overexpressed in multidrug-resistant tumours.

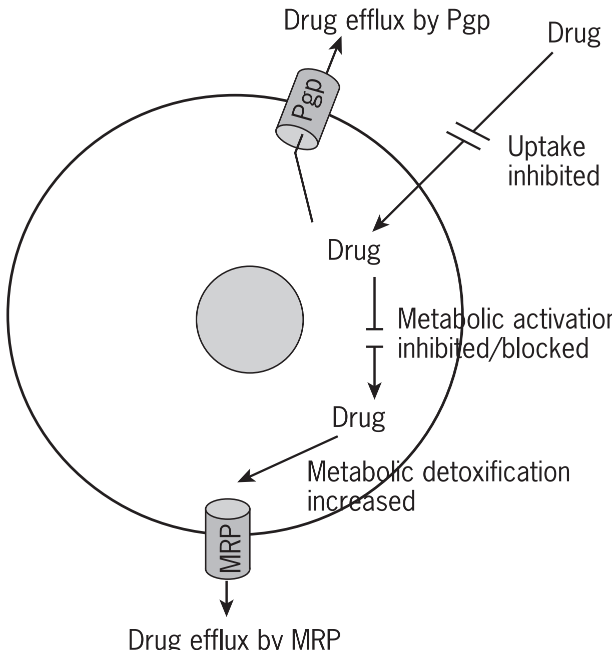


Figure 1 Summary of the major mechanisms of multidrug resistance.

A summary of the major mechanisms of MDR is shown in **Figure 1**.

TRANSPORT PROTEINS

The efflux transport of chemotherapeutic drugs and its role in multidrug resistance is well established, and this section will focus on the two major transport proteins involved in

multidrug resistance, those members of the ATP-binding cassette superfamily, P-glycoprotein and MRP.

The ATP-binding Cassette Superfamily

The ATP-binding cassette ((ABC)-ATPase) proteins constitute a very large and highly conserved superfamily found in large numbers in all organisms. This particular class of ATPase is usually found linked to transport protein, with all ABC systems made up of a highly conserved ABC-ATPase, and at least one cognate, but much less conserved membrane domain (MD). In these systems the ABC-ATPase acts as an energy generator and the membrane domain as a transport pathway.

One of the outstanding features about members of the ABC transporters is the huge range of compounds transported by different members of the family, from ions to large polypeptides, polysaccharides and xenobiotics. Another feature of this family of transporters is the ability of the transporters to act as either export or import pumps.

The binding and hydrolysis of ATP is the key to the energization of movement of substrates across the membrane by the ABC transporters. Many studies have shown that the two ABC domains of Pgp function cooperatively and that inactivation of one catalytic site completely abolishes all ATPase activity (Holland and Blight, 1999). Studies of vanadate trapping of ADP in the specific catalytic sites, ABC1 or ABC2, have led to the development of a model of ATP hydrolysis alternating equally between the two ABC domains, implying that ATP bound at one ABC site prevents hydrolysis at the second site. This model of the ATP catalytic cycle is still widely held and can be considered likely to hold for many ABC transporters.

The two transmembrane domains of ABC transporters are highly hydrophobic. Each is predicted, from its sequence, to consist of multiple α-helical segments that could span the membrane. The majority of transporters are predicted to have six membrane-spanning segments per domain, with the N- and C-termini on the cytoplasmic side of the membrane and three extracellular and two intracellular loops per domain. An example of this classical structure of ABC transporters is that of P-glycoprotein, which is discussed in greater detail in the next section and is shown in **Figure 3**. Not all ABC transporters have this classical structure; MRP1 (discussed later) is believed to have an unusual arrangement of the membrane domains with five additional transmembrane domains at the N-terminus, making a total of 11 for MD1 and six for MD2 (**Figure 2**) (Borst *et al*., 1999).

P-glycoprotein

Introduction

The best-studied mechanism of multidrug resistance is that due to the overexpression of P-glycoprotein (Pgp). Pgp was first detected as a surface phosphoglycoprotein overexpressed in cultured cells selected for MDR (Juliano and Ling, 1976), and was subsequently cloned from human cells based on amplification of the *MDR* locus (Chen *et al*., 1986). Since its discovery, the overexpression of human *MDR1* cDNA has been shown to confer resistance to a number of cytotoxic anticancer drugs, and also many other hydrophobic pharmacological agents.

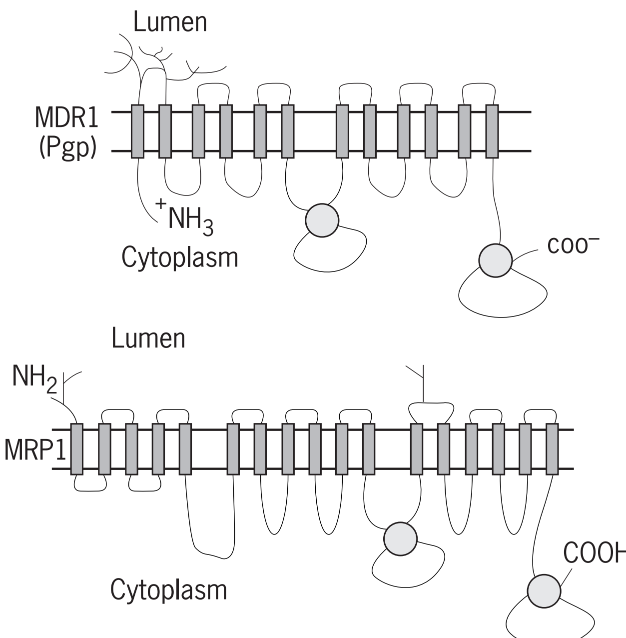


Figure 2 Proposed structures of P-glycoprotein and multidrug resistance-associated protein.

Table 1 Classes of *MDR* genes found in a number of species

	Class I	Class II	Class III
Human	*MDR1*		*MDR3*
Mouse	*mdr3*	*mdr1*	*mdr2*
Hamster	*Pgp1*	*Pgp2*	*Pgp3*

MDR1 Gene

The human *MDR1* gene, which encodes for Pgp, has been localized to chromosome 7, band q21.1. Pgp is composed of two homologous halves, each containing six transmembrane domains and an ATP-binding domain, separated by a flexible linker polypeptide (**Figure 3**). The protein-coding portion of the gene contains 27 exons: 14 coding for the left and 13 coding for the right half of the protein. There is no clear correlation between individual exons and specific structural units of the protein. Genetic analysis in a number of species has revealed the existence of a number of *MDR* genes (**Table 1**) (Juranka *et al*., 1989).

Expression in Normal Human Tissues

Expression of *MDR1* mRNA has been observed in a number of normal human tissues, with the highest levels of expression being found in adrenal gland, kidney, liver and colon (Gottesman *et al*., 1991), while lower levels were found in lung, jejunum, rectum, brain and prostate. Pgp shows specific localization within the plasma membrane of cells and has been immunolocalized on the apical surface of the brush border of the proximal tubule of the kidney, the apical mucosal surface of the colon, the luminal biliary canilicular surface of hepatocytes and ductular epithelium in the liver (Thiebaut *et al*., 1987).

Possible Normal Functions of P-glycoprotein

The localization of Pgp on the luminal surface of secretory organs such as the liver and kidney may have evolved to protect animals against toxic hydrophobic natural products, many of which are found in foodstuffs. Pgp can be viewed as a transmembrane energy-dependent drug efflux pump which uses energy derived from nucleoside triphosphates to pump hydrophobic amphipathic drugs, which have defused into the proximal tubule cells of the kidney, or the pericanalicular hepatocytes, into the urine or the bile. Another possible function for Pgp in the liver would be to act as a pump for normal metabolites that are excreted in the bile. The liver is an active modifier of circulating metabolites and also produces bile salts (which are hydrophobic, amphipathic molecules), some of which may be substrates for Pgp. Expression of Pgp in the mucosal cells of the jejunum and colon may be involved in the prevention of reabsorption of compounds excreted in bile, and indeed the prevention of absorption of dietary toxins. (See chapter on *Dietary Genotoxins and Cancer*.)

The function of expression of Pgp within the adrenal gland is more uncertain. The high level of expression of Pgp observed in the adrenal cortex may suggest a role for Pgp in steroid transport. Progesterone has been shown to interact with Pgp in multidrug-resistant cells and in the endometrium gravid uterus (Yang *et al*., 1989).

Expression of Pgp in the placenta could act as a barrier protecting the foetus from xenobiotics. Experiments using mice, which had no placental Mdr1a (Pgp), showed increased foetal exposure to [^{3}H]digoxin, [^{14}C]saquiniavar and paclitaxel, indicating that placental expression of Pgp is important in limiting the foetal penetration of various potentially harmful or therapeutic compounds (Smit *et al*., 1999).

Similarly expression of Pgp on the luminal surface of specialized endothelial capillary cells of the brain and testis indicates that Pgp may have an important function as part of the blood–brain or blood–testis barrier.

Structure and Function of P-glycoprotein

One of the first and most consistent alterations observed in MDR tumour cells was a decrease in cellular chemotherapeutic drug accumulation. There is considerable evidence that Pgp is involved in removal of chemotherapeutic drugs from the cell; however, the exact mechanism by which this occurs is still open to debate.

The most prevailing hypothesis is that Pgp functions as a transmembrane pore-forming protein, which acts in an energy-dependent manner to export compounds. The identification of ATPase activity associated with P-glycoprotein provides a mechanism by which energy may be transduced for active drug efflux. This hypothesis of drug efflux through a single pore formed by Pgp leaves a number of aspects not fully explained, such as the broad substrate specificity observed for Pgp (see the next section), and the inability to correlate the initial rate of transport with the Pgp concentration. Gottesman and Pastan hypothesized that the MDR1 transporter might act as a 'hydrophobic vacuum cleaner' by removal of compounds directly from the plasma membrane before they reached the cytosol (Gottesman and Pastan, 1993). Under the model proposed by Gottesman and Pastan the primary determinant of substrate specificity would be the ability of a compound to interact with the lipid bilayer and the secondary determinant would be its ability to interact with the binding site of the transporter. This is consistent with the observation that substrates for Pgp are lipophilic compounds and that the major determinant of a particular substrate to be transported by Pgp is its relative hydrophobicity (Zamora *et al*., 1988).

Photoaffinity labelling, mutational analysis and inhibitor studies indicate that transport of compounds by Pgp occurs through a single barrel of the transporter as illustrated in **Figure 3**. This model does not imply that different compounds enter the transporter in the same way, as the initial point of contact may vary from compound to compound, but it limits the passage of compounds through the transporter to a single transport channel.

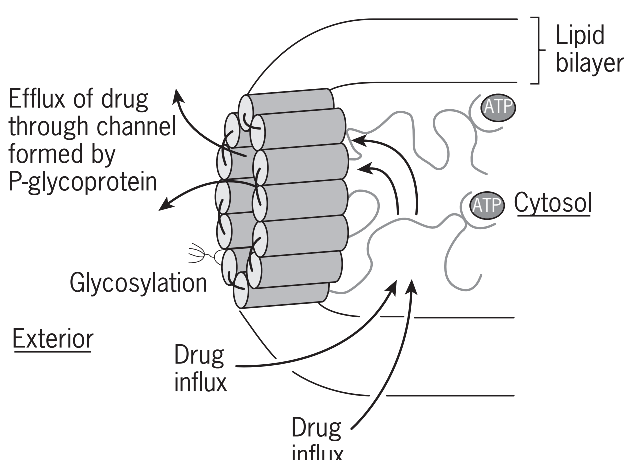


Figure 3 Schematic diagram of the proposed structure of P-glycoprotein, which functions as an energy-dependent drug efflux pump.

Very little is known about how the energy of ATP is harnessed by Pgp in order to facilitate drug transport, although it is know that binding of ATP to both binding sites is required for transport to occur efficiently (Azzaria *et al*., 1989). A number of hypothetical mechanisms of action for Pgp have been put forward, including possible action as a flippase (Higgins and Gottesman, 1992), or action as a proton pump and chloride channel providing motive force (Gottesman and Pastan, 1993). So far none of these mechanisms of action have fully explained all aspects of Pgp structure and function.

Substrates of P-glycoprotein

Pgp confers resistance against a wide spectrum of compounds that are hydrophobic, amphipathic natural product drugs (**Table 2**). These compounds include not only anticancer drugs, but also therapeutic agents. These compounds are chemically diverse, some of them may carry a positive charge at physiological pH and, because all of them are hydrophobic, they enter cells by passive diffusion. Substrates and antagonists for Pgp have been shown to interact with Pgp in different ways (Scala *et al*., 1997). In this study the interactions of a number of substrates and antagonists were examined, 35 compounds were identified as substrates only, while 42 were identified as antagonists only, with just seven compounds showing features of both substrates and antagonists. This suggests that antagonists bind to Pgp, but failing to be transported they block the binding and transport of other compounds, while substrates, in being transported, do not affect the transport of other compounds. Interestingly, this study noted that compounds in the antagonist group were relatively less toxic than compounds in the substrate group, with a median IC_{50} value of 22.5 μmol L^{-1} for the antagonist group, compared with 0.157 μmol L^{-1} for the substrate group.

Table 2 Common examples of substrates for P-glycoprotein

Compound type	Examples
Anticancer agents	*Vinca* alkaloids (vincristine, vinblastine)
	Anthracyclines (doxorubicin, epirubicin)
	Epipodophyllotoxins (etoposide, teniposide)
	Paclitaxel (Taxol)
	Actinomycin D
	Topotecan
	Mithramycin
	Mitomycin C
Other cytotoxic agents	Colchicine
	Emetine
	Ethidium bromide
	Puromycin
Cyclic and linear peptides	Granicidin D
	Valinomycin
	N-Acetylleucylleucylnorleucine
	Yeast a-factor pheromone
HIV protease inhibitors	Ritonavir
	Indiavir
	Saquinavir
Other compounds	Hoechst 33342
	Rhodamine 123
	Calcein-AM

Modulation of P-glycoprotein Activity

Not surprisingly, owing to the diverse nature of substrates for Pgp, the number of compounds that have been shown to be capable of modulating Pgp is also large and diverse, with the list including detergents, anthracycline analogues, progestational and antioestrogenic agents, antibiotics, antihypertensives and immunosuppressives. Some of the more commonly used modulators of Pgp are listed in **Table 3**. Perhaps the most important observation in the field of Pgp modulation came when Tsuruo *et al*. noted that verapamil, a membrane-active drug used in cardiology, could sensitize vincristine-resistant cells to the cytotoxic effects of vincristine and vinblastine (Tsuruo *et al*., 1981). Since verapamil blocks voltage-gated Ca^{2+} channels, other calcium channel blockers were studied and many such as nifedipine were also found to inhibit Pgp.

Many modulators, however, are not specific to Pgp, and may inhibit other transporters such as the multidrug resistance-related protein MRP (discussed later). For this reason, specific inhibitors have been sought. One such inhibitor is LY335979, which has been shown specifically to inhibit Pgp (Dantzig *et al*., 1999).

Expression of P-glycoprotein in Tumours

High intrinsic expression of MDR1/Pgp has been observed in a number of human tumours derived from tissues that normally express Pgp, including the kidney, the colon, the adrenal gland, and liver (Fojo *et al*., 1987; Gottesman *et al*., 1991). Examination of the level of expression of Pgp during various stages of colorectal carcinogenesis has shown that despite the initial high level of Pgp expression found in

Table 3 Commonly used modulators of P-glycoprotein activity

Type of modulator	Compound
Calcium channel inhibitors	Verapamil
	Nifedapine
Immunosuppressants	Cyclosporin A
	PSC 833
Other inhibitors	LY335979

normal colorectal tissue, the progression of normal tissue to adenomas and then to carcinomas shows a concordant stepwise increase in Pgp expression (Meijer *et al*., 1999).

Some other untreated cancers often show high levels of *MDR1* RNA, despite the tissue of origin not expressing MDR1. These cancers can include acute and chronic leukaemias of children and adults, non-Hodgkin lymphoma, neuroblastoma and sarcomas. The expression of the *MDR1* gene in tumours derived from tissues which do not normally express Pgp suggests that the process of malignant transformation in carcinogenesis can activate expression of the *MDR1* gene. This seems to have been confirmed in a study showing the *MDR1* promoter can be stimulated by Ras and mutant p53, two proteins commonly associated with tumour progression (Chin *et al*., 1992).

A group of tumours exist which show low or undetectable levels of Pgp, including breast cancers, non-small-cell lung cancer (NSCLC), small-cell lung cancer (SCLC), bladder cancer, oesophageal carcinoma, melanoma, ovarian cancer and prostate cancer. Despite low or undetectable expression of Pgp, some of these tumours often show a degree of multidrug resistance, indicating the possibility of the involvement of other mechanisms of drug resistance.

Increased or acquired expression of Pgp is commonly seen in tumours treated with chemotherapeutic drugs that have relapsed either during or after treatment (Fojo *et al*., 1987). The expression of Pgp in these tumours is thought be due to either selection of a small number of MDR1-expressing cells, present at the onset of treatment, or a direct induction of *MDR1* gene expression. There has been only limited study in this area owing to the difficulty in obtaining samples both before and after treatment, and for this reason most research has concentrated on the development of acquired resistance *in vitro*.

Expression of P-glycoprotein In Vitro

The intrinsic and acquired expression of Pgp *in vitro* has been studied extensively, through the use of both *in vivo* and *in vitro* selection procedures. (See chapters on *Basic Tissue Culture in Cancer Research*; *Models for Drug Development and Drug Resistance*.)

Intrinsic Multidrug-resistant Cell Lines

Many cell lines have been isolated from multidrug-resistant tumours from both humans and rodents. These cell lines often show classical characteristics of multidrug-resistant tumours, with increased expression of Pgp.

In Vivo Selection of MDR Cell Lines

Several MDR rodent tumour cell lines have been established, generally with transplantable tumour cell lines/grafts being inserted into an animal that is then exposed to a chemotherapeutic drug. For example, the murine leukaemia cell line P388 has been used to establish a doxorubicin-resistant subclone, by treating P388-inoculated mice with doxorubicin over repeated transplantation periods (Johnson *et al*., 1976). Generally, *in vivo* resistant cell lines show 10–100-fold resistance to various anticancer agents.

In vitro Selection of Cell Lines

Another procedure for the isolation of MDR cell lines is the exposure of cultured cells to a selective drug *in vitro*. There are two types of drug selection, first continuous exposure to a sublethal drug concentration, followed by gradual increases in the drug concentration. The second method of selection is a short exposure to a relatively high dose of the drug, followed by subsequent propagation of the cells in drug-free medium. The selection pressures under these circumstances are not the same as those experienced during the development of MDR during tumour chemotherapy, but a number of similarities are seen between cells selected in this manner and drug-resistant tumours.

An example of the *in vitro* establishment of a multidrug-resistant cell line is the selection of the doxorubicin-resistant human laryngeal carcinoma cell line HEp2A, from HEp2 cells (Redmond *et al*., 1990). In this case, human laryngeal carcinoma cells were exposed to increasing doses of doxorubicin over a 14-month period. HEp2A cells were found be 100-fold more resistant to doxorubicin compared with HEp2 cells, and they were 10–20-fold more resistant to the *Vinca* alkaloids vinblastine and vincristine. Pgp is usually found to be overexpressed in the *in vitro*-selected multidrug-resistant cell lines compared with the relevant drug-sensitive parental cell line.

Multidrug Resistance-associated Protein

Introduction

The multidrug resistance-associated proteins (MRPs) are members of the ATP-binding cassette superfamily of transport proteins. Although Pgp overexpression has been associated with multidrug resistance in many tumours and cell lines, there are a number of examples of multidrug-resistant cell lines and tumours where Pgp is not involved. Probably the most extensively characterized of these cell lines is H69AR, a small-cell lung cancer cell line derived from the H69 parental line, through stepwise selection in doxorubicin (Mirski *et al*., 1987). Through the use of differential hybridization, a 6.5-kb mRNA sequence was identified as being approximately 100-fold higher than in parental cells. Sequencing of cDNA clones derived from this mRNA revealed that it had the potential to encode a 1531 amino acid protein that was predicted to be a member of the ABC transporter superfamily and was subsequently named the multidrug-resistance protein (MRP) (Cole *et al*., 1992). The *MRP* gene has been mapped to chromosome 16 at band p13.13–13.12, and is amplified relatively frequently in drug-selected human cell lines that overexpress *MRP* mRNA.

The Mammalian MRP Family

Since the discovery and cloning of *MRP* in 1992 by Cole and Deeley (Cole *et al*., 1992), at least seven members of the MRP family have been discovered in mammals (Borst *et al*., 1999). The original *MRP* cloned by Cole is now known as MRP1, complicating early expression data for tissues, tumours and cells, with the specificity of early studies on MRP towards MRP1 not being clear. The next MRP transporter to be cloned was *MRP2* in 1996 (Buchler *et al*., 1996), although it had been characterized as the canalicular multispecific organic anion transporter (cMOAT) a number of years earlier. Within the MRP family the greatest sequence homology is between MRP1, -2, -3 and -6 (Borst *et al*., 1999).

Expression of MRPs in Normal Tissues

The expression of MRPs in normal tissues has been examined using both mRNA and protein analysis (Borst *et al*., 1999). MRP1 and MRP5 have been found to be ubiquitously expressed in normal tissues, with only relatively low levels of MRP1 in the liver. MRP2 is found in the liver, the kidney and the gut, while MRP3 has been shown to be expressed in liver, adrenals, pancreas, kidney and gut. MRP4 is expressed in prostate, lung, muscle, pancreas, testis, ovary, bladder and gall bladder, while MRP6 is only found in liver and kidney.

The localization of some MRPs in transfected polarized kidney epithelial cells has shown MRP1, MRP3 and MRP5 to localize to the basolateral membrane, while MRP2 localizes to the apical membrane (Borst *et al*., 1999; Konig *et al*., 1999).

Physiological Functions of MRPs

MRPs are thought to be like Pgp, having both physiological roles and a protective role against xenobiotics.

MRP1 is the major high-affinity transporter of leukotriene C_4 (LTC_4), as shown by biochemical experiments and the knockout (KO) mice (Wijnholds *et al*., 1997). Interestingly, in KO mice the absence of this major LTC_4 pathway does not have any observed negative effect on the mice within the confines of an animal house; on the contrary, the KO mice are more resistant than wild-type mice against inflammatory stimuli. This suggests that there may be a possible overlap in function between MRPs, with another MRP being able to compensate for a lack of MRP1; evidence for such a hypothosis may lie with MRP2, which has also been shown to be capable of transporting LTC_4 (Cui *et al*., 1999).

As MRP1 is located at the basolateral side of epithelial cells, it tends to pump drugs into the body, rather than away from it via the bile, urine or the gut in the way that Pgp does. However, MRP1 has been shown to protect vital cell layers from destruction, e.g. the basal stem cell layer in the oral mucosa (Wijnholds *et al*., 1998). There are also other important organs in the body that require a basolateral transporter for protection, e.g. the testicular tubules, where a ring of Sertoli cells protects the male germ cells, with their apical surface towards the lumen. The high concentration of MRP1 in the basal membrane of the Sertoli cells helps to protect the germ cells by pumping out drugs from the testicular tubule. Another position in the body where a basolateral transporter is required for protection is the choroid plexus, which protrudes into the cerebrospinal fluid (CSF) and is essential for exchange of metabolites between blood and CSF. The epithelial layers of the plexus contain high levels of MRP1 and the absence of MRP1 in KO mice results in the accumulation of etoposide in the CSF, as well as in brown adipose tissue, colon, salivary gland, heart and the female urogenital system (Wijnholds *et al*., 2000).

In the liver MRP2 is localized in the biliary canalicular membrane, and there is a lot of evidence that both MRP1 and MRP2 provide a major route for the secretion of organic anions such as bilirubin glucuronides from the liver. Rats and humans lacking MRP2 develop the liver disease Dubin–Johnson syndrome, mainly owing to the inability of the liver to excrete bilirubin glucuronides (Konig *et al*., 1999).

The expression of MRP3 on the basolateral membrane in the intestine and liver suggests a role for MRP3 in the intestinal uptake of organic anions and the removal of organic acids from bile and liver cells under conditions of cholestasis (Borst *et al*., 1999).

Substrates of MRPs

Following the demonstration in 1994 that MRP1 functions as a unidirectional ATP-dependent efflux pump for the endogenous glutathione *S*-conjugate LTC_4 (Jedlitschky *et al*., 1994), the substrate specificity of MRPs has been intensively studied. The substrate spectrum comprises amphiphilic anions, particularly conjugates of lipophilic compounds with glutathione, glucuronate or sulfate, including cysteinyl leukotrienes, bilirubin glucuronosides, 17β-glucuronosyloestradiol and sulfatolithocholyltaurine as endogenous compounds.

So far the glutathione *S*-conjugate LTC_4 is the substrate with the highest affinity for MRP1 ($K_m = 0.1\ \mu mol\ L^{-1}$) (Konig *et al*., 1999). The range of substrates for MRP1 and MRP2 is very similar, although MRP1 exhibits a 10-fold greater affinity for the model substrate LTC_4, while MRP2 has a higher affinity for the bilirubin glucuronides (Jedlitschky *et al*., 1997).

Modulators of MRPs

The identification of potent and specific inhibitors of MRP-mediated transport function has been sought since the characterization of the family. The most potent inhibitors reported so far, and tested against MRP1-mediated

transport in isolated vesicles, are analogues of cysteinyl leukotrienes, such as the LTD_4 analogue MK571. This analogue is a monoanionic quinoline derivative developed as a LTD_4 receptor antagonist, and has so far been shown to inhibit MRP1 transport and to a lesser extent MRP2 transport (Buchler *et al*., 1996).

MRPs and Multidrug Resistance

To date MRP1, MRP2 and MRP3 have been shown to be expressed in multidrug-resistant tumours and cell lines.

MRP1

Since its discovery in the highly resistant H69AR lung cancer cell line, MRP1 has been found overexpressed in multidrug-resistant cell lines derived from many different tissue and tumour types, including small- and large-cell lung cancers, carcinomas of the colon, breast, bladder, prostate, thyroid and cervix, glioma, neuroblastoma, fibrosarcoma, and various forms of leukaemia. Many of these cell lines have been established using selection to various chemotherapeutic agents such as doxorubicin and vincristine, although intrinsic expression of MRP1 in cell lines derived from tumours has also been observed.

In order to circumvent many of the problems associated with the use of drug-selected cell lines, namely the induction of multiple proteins such as MRPs, Pgp and xenobiotic-metabolizing enzymes, MRP1-mediated multidrug resistance has also been investigated through transfection with MRP1 expression vectors. The resistance profiles of MRP1 transfected cells are similar to those of cell lines expressing Pgp, and typically include resistance to anthracyclines and *Vinca* alkaloids. In contrast to Pgp, human MRP1 expression also confers low-level resistance to arsenical and antimonial oxyanions. Like Pgp, MRP1-mediated resistance is usually linked to a decrease in cellular accumulation of xenobiotics, e.g. MRP1-transfected cells accumulate less vincristine and daunorubicin than control cells (Cole *et al*., 1994). Again like Pgp, active transport of compounds corresponds to ATPase activity, and this has also been shown to be stimulated in isolated membrane vesicles from MRP1-transfected cells, via the addition of the anthracyclines daunomycin and doxorubicin and the *Vinca* alkaloids vincristine and vinblastine (Chang *et al*., 1997).

MRP2

Like MRP1, stable transfection of MRP2 has been shown to confer resistance to the cytotoxic drugs etoposide (4-fold), cisplatin (10-fold), doxorubicin (7.8-fold), and epirubicin (5-fold) in human embryonic kidney cells (HEK-293) (Cui *et al*., 1999). Further evidence of a role for MRP2 in multidrug resistance was provided by the transfection of antisense cDNA in the human hepatoma cell line HEpG2 in order to reduce MRP2 expression (Koike *et al*., 1997). Upon transfection of the antisense *MRP2* cDNA HEpG2 cells showed increases in intracellular glutathione levels and enhanced sensitivity to cisplatin, doxorubicin and vincristine.

MRP3

Only a limited amount of data exists on the relationship between MRP3 and multidrug resistance. A survey of lung cancer cell lines has shown a strong correlation between MRP3 expression and doxorubicin resistance (Young *et al*., 1999). In the same study a weaker, yet significant, correlation between MRP3 expression and resistance to vincristine, etoposide and cisplatin was also found.

Role of Glutathione in MRP-mediated Drug Resistance

Despite the link observed between MRP1 expression and resistance to various chemotherapeutic agents, the *in vitro* transport of these compounds using MRP1-enriched membrane vesicles had not been observed. Evidence for the reasons behind the lack of detectable transport in these models first came from studies using buthionine sulfoxamine (BSO). BSO is a potent irreversible inhibitor of γ-glutamylcysteine synthesis, the enzyme that catalyses the first, rate-limiting step in the synthesis of glutathione (GSH). Treatment of cells with BSO results in reduction of intracellular GSH levels by up to 90% within 24 h, depending on the cell line examined. A number of studies have shown that BSO-treatment is capable of enhancing drug accumulation and toxicity in MRP1-overexpressing cell lines (Lautier *et al*., 1996). The mechanism by which BSO inhibits drug transport by MRP is believed to be as a direct result of depletion of GSH, and not an interaction between BSO and MRP. Evidence supporting this includes the finding that acute exposure to BSO had no effect on drug accumulation (Schneider *et al*., 1995), and treatment with GSH ethyl ester increased cytoplasmic GSH and decreased daunorubicin accumulation in two BSO-treated MRP1-overexpressing cell lines (Versantvoort *et al*., 1995).

Observations that the glutathione-conjugated chemotherapeutic drugs doxorubicin and daunorubicin but not the unconjugated drugs were capable of competitively inhibiting the MRP-mediated transport of LTC_4 *in vitro* (Priebe *et al*., 1998) have led to speculation that MRP-mediated drug resistance occurs by the transport of drug conjugates from cells, and in fact MRPs had previously been called MRP/GS-X pumps (Ishikawa, 1992). This, however, has been shown to be inappropriate since, as previously discussed, MRPs, particularly MRP1 and MRP2, have been shown to be capable of transporting compounds not conjugated to glutathione such as bilirubin mono- and diglucuronides. Although many substrates for MRPs are conjugated compounds, the conjugation is not thought to be critical for transport; there has been some evidence suggesting that reduced GSH can also act as

either a co-substrate or possess the ability to activate MRPs. Evidence for this came from the observation that unconjugated chemotherapeutic agents, such as doxorubicin, daunorubicin, vincristine and vinblastine, fail to inhibit LTC_4 transport in isolated vesicles even at very high doses (Loe *et al*., 1996). However, on addition of reduced GSH, the inhibitory potency of some of these drugs is markedly increased, with the most dramatic effect occurring with vincristine and vinblastine. Further to this, vincristine uptake into membrane vesicles from MRP1-transfected HeLa cells has been shown to be both ATP and GSH dependent and could be inhibited by the MRP1-specific mouse antibody QCRL-3 (Loe *et al*., 1998). In this study, vincristine transport in the absence of GSH was extremely low and showed no ATP dependence, whereas active vincristine transport increased with increasing concentrations of GSH. As GSH appears not to be a substrate for MRP1, there is strong evidence for the co-substrate theory, provided by a study showing that the ATP-dependent uptake of [^{3}H]GSH into membrane vesicles from MRP1-transfected cells could be stimulated by vincristine in a dose-dependent manner (Loe *et al*., 1996).

Taken as a whole, the evidence so far collected on the transport of compounds by MRPs suggests that GSH is required and may act as a co-substrate during transport of compounds. The mechanism by which this occurs is not clear, but it may be that GSH acts as a co-substrate with unconjugated compounds, or if the compound is already conjugated, the conjugate attached to the compound acts as a co-substrate, negating the need for a second GSH molecule.

Expression of MRPs in Tumours

To date, most studies on the expression of MRPs in human tissues and tumours have examined mRNA levels through the use of Northern blot analysis and reverse transcriptase polymerase chain reaction (RT-PCR). There are inherent problems involved in the use of these techniques, owing to the lack of quantification of protein expression and function by these techniques. Bearing this in mind, mRNA of MRPs has been detected in various tumours including many leukaemias, non-small-cell lung cancers, squamous cell carcinomas of the lung, neuroblastomas, oesophageal squamous cell carcinomas and colon carcinomas.

Expression of MRP1 in a number of human cancers shows that MRP1 expression in tumours falls into three groups. The first are tumours with frequently high MRP1 expression above those of normal tissues, including chronic lymphocytic leukaemia and polymorphocytic leukaemia. The second group of tumours described were those that often exhibit low expression but on occasions can express high levels; these included oesophageal squamous cell carcinoma, non-small-cell lung carcinoma and acute myelocytic leukaemia. The third group were those tumours with predominantly low levels of MRP1 expression, comparable to those found in normal tissues, and these included soft tissue tumours, melanomas and cancers of the prostate, kidney, bladder, testis, ovary and colon. Interestingly, this study found no significant difference in the expression of MRP1 in chemotherapy-treated versus untreated tumours, suggesting that MRP1 does not play a significant role in acquired multidrug resistance, and is more important in the intrinsic resistance of particular types of tumours.

The expression of MRPs in lung tumours has been extensively studied, primarily due to the high levels of expression found in normal lung tissues. Analysis of the expression of *MRP1* to *MRP5* mRNA by RT-PCR showed expression of *MRP1* in 88% of small cell lung cancer samples and 100% of non-small-cell lung cancer samples (Young *et al*., 1999). *MRP2* was detected in 75% of small-cell lung cancer samples and 57% of non-small-cell lung cancer samples, while *MRP3* was found in only 12% of small-cell lung cancer samples and 86% of non-small-cell lung cancer samples. Expression of *MRP4* and *MRP5* was found in 100% of small-cell lung cancer samples and 71% and 100% of non-small-cell lung cancer samples, respectively. Interestingly, analysis of expression of MRPs in unselected cell lines established from patients at various stages of treatment in the same study showed a strong correlation between *MRP1* and *MRP3* mRNA levels in these cells and doxorubicin and vincristine resistance, as determined by a modified MTT method.

Analysis of the expression of MRPs has been carried out in normal colorectal tissue and tumours, showing MRP1 and MRP3 to be expressed at high levels in both normal and tumour tissues. In contrast, MRP2 expression is generally lower in normal tissue than in tumour samples.

Other Possible Mediators of Multidrug Resistance

Despite most research into the field of multidrug resistance focusing on mechanisms of drug transport, other factors should also be considered. One such factor is that of drug-metabolizing enzymes, and in particular glutathione *S*-transferase isoforms. In a variety of experimental models, increased tolerance to toxic xenobiotics and in particular chemotherapeutic compounds has been associated with increased GST expression. Examples of models include the multidrug-resistant phenotype of preneoplastic hepatocyte nodules in the rat (Farber, 1990), the analyses of tumour cells from patients before and after the onset of clinical drug resistance (Lewis *et al*., 1988), selection of resistance to chemotherapeutic agents in tumour cell lines (Wang *et al*., 1999), transfection of GST into mammalian cell lines (Puchalski and Fahl, 1990) and transfection of antisense genes causing increased drug sensitivity (Niitsu *et al*., 1998). Despite the large body of evidence supporting the involvement of glutathione *S*-transferase isoforms in

resistance to chemotherapeutic drugs, the area remains controversial, with many groups failing to show the involvement of GSTs in drug resistance (Townsend *et al*., 1992).

Expression of GST Isoforms in Tumours

The overexpression of GSTs in tumours has been observed in tumours derived from a number of tissues. Howie *et al*. showed increased total GST activity as measured by the CDNB assay in tumours of the lung, colon, stomach and breast (Howie *et al*., 1990). Expression of GSTs is not always increased in tumours, however, with the same study reporting a decrease in total GST activity in tumours of the kidney and liver compared with normal tissue. The increase observed in total GST activity seems to be due to an increase in the level of expression of the P1 isoform, which showed approximately a 2-fold increase in protein expression in lung, colon and stomach tumours. The decrease in total GST activity observed in kidney and liver tumours coincided with a decrease in the expression of the α-isoforms A1 and A2. Total GST activity measured in normal, peritumoral and tumoral colonic tissue supports the observations of Howie *et al*. showing a significant increase in activity in tumoral tissue compared with normal and peritumoral tissue, with a concurrent increase in GSTP1 expression in tumoral tissue (de Waziers *et al*., 1991).

In Vitro Expression of GSTs in Resistant Cells

Many groups have found that GSTs are overexpressed in a variety of human and rodent cell lines selected for resistance *in vitro* to doxorubicin, etoposide, cyclophosphamide, vincristine and vinblastine (Hayes and Pulford, 1995). In addition to changes in GST expression in acquired drug resistance, studies in cell lines showing differential intrinsic drug resistance have shown a correlation between resistance and increased expression of GST. The majority of resistant cell lines show an increase in the hGSTP1-1 isoform (**Table 4**).

Unfortunately, in most of the resistant cell lines studied there is overexpression of a number of detoxification enzymes, making it difficult to be certain that any increased resistance observed is due to the overexpression of GST isoforms and not a coinduced protein. In order to distinguish GST-mediated resistance from other detoxification mechanisms a number of studies have been carried out in GSH-depleted cells through the use of buthionine sulfoximine (BSO). In general, these studies have shown an increase in cytotoxicity of a variety of compounds, including doxorubicin, vincristine, etoposide, melphalan, cisplatin and daunorubicin, following GSH depletion with BSO. It must be noted, however, that in addition to the activity of GSTs, the depletion of GSH by BSO will affect the activity of glutathione peroxidase, which may also affect the resistance of cells to cytotoxic agents.

Limited studies have been carried out on the effect of inhibitors of GST activity and drug resistance. Ethacrynic acid has been shown to produce a 2–4-fold increase in the cytotoxicity of doxorubicin in primary cultures of lymphatic malignancies and normal peripheral blood lymphocytes (Nagourney *et al*., 1990). *S*-Alkyl GSH derivatives have also been shown to increase the

Table 4 Altered GST expression observed in drug-resistant cell lines

Drug	Resistance (-fold)	Cell line	Cross-resistance	Changes in GST
Doxorubicin	32	H69AR human small lung	Colchicine, daunomycin	10-fold increase in hGSTP1
Doxorubicin	75	SW620-ADR human colon	Actinomycin D, puromycin	2-fold increase in hGSTP1
Doxorubicin	84	P388 murine leukaemia	Not stated	2.2-fold increase in hGSTP1
Doxorubicin	100	HEp2A human laryngeal carcinoma	Vincristine, vinblastine, VP-16, VM-26	2.9-fold increase in hGSTP1
Chlorambucil	>10	Mouse fibroblast 3T3 (N50-4)	Not stated	>10-fold increase in α class GST
Cisplatin	>18	HeLa human cervix (HeLa-CPR)	Not stated	6-fold increase in hGSTA1 and A2
Etoposide (VP-16)	14	MCF-7 human breast (VP6E)	Vincristine	>10-fold increase in hGSTP1
Mitoxantrone	200	Caco-2 human colon	Not stated	>6-fold increases in hGSTP1, A1 and A2
Vincristine	3	MCF-7 human breast (VCR6E)	VP-16	>10-fold increase in hGSTP1
Vincristine	11	MCF-7 human breast (VCREMS)	Adriamycin/ doxorubicin VP-16	>10-fold increase hGSTP1
Doxorubicin	30–65	MCF7 human breast	Not stated	Overexpression of hGSTP1

(Adapted from Hayes and Pulford, 1995).

cytotoxicity of doxorubicin in human lung adenocarcinoma cell lines by the inhibition of GST activity (Nakanishi *et al*., 1997).

Transfection Studies

As a result of the unpredictable nature of intrinsic and acquired resistant cell lines, a number of groups have attempted to confer drug resistance by the transfection of GST isoforms. Although variable results have been obtained, some compelling evidence for a role for GSTs in drug resistance has been obtained. It should be noted, however, that the transfection of cells with GSTs may also lead to changes in other genes within the cell, such as changes in heat shock proteins, or enzymes involved in oxidative stress responses, which may modulate the intrinsic resistance of the cells.

The first demonstration that GST can confer resistance against toxic xenobiotics was reported by Manoharan *et al*. in 1987, who showed that transfection with rGSTA2-2 in COS cells provided resistance to the alkylating agent benzo[*a*]pyrene (±)-*anti*-diol epoxide (Manoharan *et al*., 1987). The same group later showed, through the use of cytotoxicity assays, that transfection with rGSTA2-2 conferred a 1.3–2.9-fold resistance to chlorambucil and melphalan, while expression of rGSTM1-1 conferred a 1.5-fold resistance to cisplatin and hGSTP1-1 conferred a 1.3-fold resistance to doxorubicin. Similar results have been found using NIH-3T3 and HEp2 cells, where expression of hGSTP1-1 conferred a 3-fold resistance to doxorubicin (Nakagawa *et al*., 1990; Harbottle *et al*., 2001).

A number of studies have reported no increased resistance to a number of chemotherapeutic compounds following the transfection of GSTs into MCF-7 cells (Townsend *et al*., 1992). It seems that mechanisms to induce the expression of GSTs, are very dependent on the type of cell into which transfection is carried out. This may indicate that other factors are required before GST-mediated resistance to cytotoxic compounds can occur. These factors may include the level of GSH present in cells or the expression of transport proteins from the MRP superfamily, which are expressed at very low levels in MCF-7 cells.

REVERSAL OF MULTIDRUG RESISTANCE

In an attempt to increase the success of chemotherapy, a great deal of research is taking place into methods of reversing the resistance of tumours to chemotherapeutic agents. These methods range from the development of compounds intended to be administered in combination with cancer chemotherapeutic drugs, that inhibit the transport proteins responsible for the efflux of drugs from target cells, to the use of antisense technology to develop drugs to block the synthesis of these transport proteins within tumour cells. As a result, a large number of companies currently have research programmes in the field of multidrug resistance, with a number of compounds being in phase I/II trials (Persidis, 2000).

As understanding of the structure and function of drug transport proteins and multidrug resistance increases, it is becoming clear that although a large number of chemotherapeutic agents are substrates for transport proteins, some are not. For example, S16020-2, a novel derivative of olivacine, has shown significant antitumour activity both *in vitro* and *in vivo* against multidrug-resistant cells (Pierre *et al*., 1998). The COMPARE programme of the US National Cancer Institute (NCI) for the identification of agents in the NCI database that would be predicted to be good substrates and/or inhibitors of P-glycoprotein may lead to the discovery of more potent inhibitors or indeed cytotoxic compounds that are not substrates for P-glycoprotein.

REFERENCES

Azzaria, M., *et al*. (1989). Discrete mutations introduced in the predicted nucleotide-binding sites of the mdr1 gene abolish its ability to confer multidrug resistance. *Molecular Cell Biology*, **9**, 5289–5297.

Borst, P., *et al*. (1999). The multidrug resistance protein family. *Biochimica Biophysica Acta*, **1461**, 347–357.

Buchler, M., *et al*. (1996). cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance protein, cMrp, reveals a novel conjugate export pump deficient in hyperbilirubinemic mutant rats. *Journal of Biological Chemistry*, **271**, 15091–15098 .

Chang, X. B., *et al*. (1997). ATPase activity of purified multidrug resistance-associated protein. *Journal of Biological Chemistry*, **272**, 30962–30968; Erratum, *Molecular Pharmacology*, 1998, **273**, 7782.

Chen, C. J., *et al*. (1986). Internal duplication and homology with bacterial transport proteins in the mdr1 (P-glycoprotein) gene from multidrug-resistant human cells. *Cell*, **47**, 381–389.

Chin, K. V., *et al*. (1992). Modulation of activity of the promoter of the human MDR1 gene by Ras and p53. *Science*, **255**, 459–462.

Cole, S. P., *et al*. (1992). Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science*, **258**, 1650–1654.

Cole, S. P., *et al*. (1994). Pharmacological characterization of multidrug resistant MRP-transfected human tumor cells. *Cancer Research*, **54**, 5902–5910.

Cui, Y., *et al*. (1999). Drug resistance and ATP-dependent conjugate transport mediated by the apical multidrug resistance protein, MRP2, permanently expressed in human and canine cells. *Molecular Pharmacology*, **55**, 929–937.

Dantzig, A. H., *et al*. (1999). Selectivity of the multidrug resistance modulator, LY335979, for P-glycoprotein and effect

on cytochrome P-450 activities. *Journal of Pharmacology and Experimental Therapeutics*, **290**, 854–862.

de Waziers, I., *et al*. (1991). Drug-metabolizing enzyme expression in human normal, peritumoral and tumoral colorectal tissue samples. *Carcinogenesis*, **12**, 905–909.

Farber, E. (1990). Clonal adaptation during carcinogenesis. *Biochemical Pharmacology*, **39**, 1837–1846.

Fojo, A. T., *et al*. (1987). Expression of a multidrug-resistance gene in human tumors and tissues. *Proceedings of the National Academy of Sciences of the USA*, **84**, 265–269.

Gottesman, M. M. and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual Review of Biochemistry*, **62**, 385–427.

Gottesman, M., *et al*. (1991). Expression of MDR1 gene in normal human tissues. In: *Molecular and Cellular Biology of Multidrug Resistance in Tumor Cells*. Roninson, Igor B. New York: Plenum Press 279–288.

Harbottle, A., *et al*. (2001). Role of glutathione-S-transferase P1, P-glycoprotein and multidrug resistance-associated protein 1 in acquired doxorubicin resistance. *International Journal of Cancer*, **92**, 777–783.

Hayes, J. D. and Pulford, D. J. (1995). The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Critical Reviews in Biochemistry and Molecular Biology*, **30**, 445–600.

Higgins, C. F. and Gottesman, M. M. (1992). Is the multidrug transporter a flippase? *Trends in Biochemical Science*, **17**, 18–21.

Holland, I. B. and Blight, M. A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *Journal of Molecular Biology*, **293**, 381–399.

Howie, A. F., *et al*. (1990). Glutathione S-transferase and glutathione peroxidase expression in normal and tumour human tissues. *Carcinogenesis*, **11**, 451–458.

Ishikawa, T. (1992). The ATP-dependent glutathione S-conjugate export pump. *Trends in Biochemical Science*, **17**, 463–468.

Jedlitschky, G., *et al*. (1994). ATP-dependent transport of glutathione S-conjugates by the multidrug resistance-associated protein. *Cancer Research*, **54**, 4833–4836.

Jedlitschky, G., *et al*. (1997). ATP-dependent transport of bilirubin glucuronides by the multidrug resistance protein MRP1 and its hepatocyte canalicular isoform MRP2. *Biochemical Journal*, **327**, 305–310.

Johnson, R. K., *et al*. (1976). Activity of anthracyclines against an adriamycin (NSC-123127)-resistant subline of P388 leukemia with special emphasis on cinerubin A (NSC-18334). *Cancer Treatment Reports*, **60**, 99–102.

Juliano, R. L. and Ling, V. (1976). A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochimica Biophysica Acta*, **455**, 152–162.

Juranka, P. F., *et al*. (1989). P-glycoprotein – multidrug-resistance and a superfamily of membrane-associated transport proteins. *Faseb Journal*, **3**, 2583–2592.

Koike, K., *et al*. (1997). A canalicular multispecific organic anion transporter (cMOAT) antisense cDNA enhances drug sensitivity in human hepatic cancer cells. *Cancer Research*, **57**, 5475–5479.

Konig, J., *et al*. (1999). Conjugate export pumps of the multidrug resistance protein (MRP) family: localization, substrate specificity, and MRP2-mediated drug resistance. *Biochimica Biophysica Acta*, **1461**, 377–394.

Lautier, D., *et al*. (1996). Multidrug resistance mediated by the multidrug resistance protein (MRP) gene. *Biochemical Pharmacology*, **52**, 967–977.

Lewis, A. D., *et al*. (1988). Glutathione and glutathione-dependent enzymes in ovarian adenocarcinoma cell lines derived from a patient before and after the onset of drug resistance: intrinsic differences and cell cycle effects. *Carcinogenesis*, **9**, 1283–1287.

Loe, D. W., *et al*. (1996). Multidrug resistance protein (MRP)-mediated transport of leukotriene C(4) and chemotherapeutic agents in membrane vesicles. *Journal of Biological Chemistry*, **271**, 9675–9682.

Loe, D. W., *et al*. (1998). Characterization of vincristine transport by the M(r) 190,000 multidrug resistance protein (MRP): evidence for cotransport with reduced glutathione. *Cancer Research*, **58**, 5130–5136.

Manoharan, T. H., *et al*. (1987). Promoter-glutathione S-transferase Ya cDNA hybrid genes. Expression and conferred resistance to an alkylating molecule in mammalian cells. *Journal of Biological Chemistry*, **262**, 3739–3745.

Meijer, G. A., *et al*. (1999). Increased expression of multidrug resistance related proteins Pgp, MRP1, and LRP/MVP occurs early in colorectal carcinogenesis. *Journal of Clinical Pathology*, **52**, 450–454.

Mirski, S. E., *et al*. (1987). Multidrug resistance in a human small cell lung cancer cell line selected in adriamycin. *Cancer Research*, **47**, 2594–2598

Nagourney, R. A., *et al*. (1990). Enhancement of anthracycline and alkylator cytotoxicity by ethacrynic acid in primary cultures of human tissues. *Cancer Chemotherapy and Pharmacology*, **26**, 318–322.

Nakagawa, K., *et al*. (1990). Glutathione-S-transferase pi as a determinant of drug resistance in transfectant cell lines. *Journal of Biological Chemistry*, **265**, 4296–4301.

Nakanishi, Y., *et al*. (1997). Glutathione derivatives enhance adriamycin cytotoxicity in a human lung adenocarcinoma cell line. *Anticancer Research*, **17**, 2129–2134.

Niitsu, Y., *et al*. (1998). A proof of glutathione S-transferase-pi-related multidrug resistance by transfer of antisense gene to cancer cells and sense gene to bone marrow stem cell. *Chemical-Biological Interactions*, **111–112**, 325–332.

Persidis, A. (2000). Cancer multidrug resistance. *Nature Biotechnology*, **18**(Suppl.), IT18–IT20.

Pierre, A., *et al*. (1998). Circumvention of P-glycoprotein-mediated multidrug resistance by S16020-2: kinetics of uptake and efflux in sensitive and resistant cell lines. *Cancer Chemotherapy and Pharmacology*, **42**, 454–460.

Priebe, W., *et al*. (1998). Doxorubicin– and daunorubicin–glutathione conjugates, but not unconjugated drugs, competitively inhibit leukotriene C4 transport mediated by

MRP/GS-X pump. *Biochemical and Biophysical Research Communications*, **247**, 859–863.

Puchalski, R. B. and Fahl, W. E. (1990). Expression of recombinant glutathione S-transferase pi, Ya, or Yb1 confers resistance to alkylating agents. *Proceedings of the National Academy of Sciences of the USA*, **87**, 2443–2447.

Redmond, A., *et al*. (1990). Establishment of two new multidrug resistant variants of the human tumor line Hep-2. *Cell Biology and Toxicology*, **6**, 293–302.

Scala, S., *et al*. (1997). P-glycoprotein substrates and antagonists cluster into two distinct groups. *Molecular Pharmacology*, **51**, 1024–1033.

Schneider, E., *et al*. (1995). Buthionine sulphoximine-mediated sensitisation of etoposide-resistant human breast cancer MCF7 cells overexpressing the multidrug resistance-associated protein involves increased drug accumulation. *British Journal of Cancer*, **71**, 738–743.

Smit, J. W., *et al*. (1999). Absence or pharmacological blocking of placental P-glycoprotein profoundly increases fetal drug exposure. *Journal of Clinical Investigation*, **104**, 1441–1447.

Thiebaut, F., *et al*. (1987). Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proceedings of the National Academy of Sciences of the USA*, **84**, 7735–7738.

Townsend, A. J., *et al*. (1992). Expression of human mu or alpha class glutathione S-transferases in stably transfected human MCF-7 breast cancer cells: effect on cellular sensitivity to cytotoxic agents. *Molecular Pharmacology*, **41**, 230–236.

Tsuruo, T., *et al*. (1981). Overcoming of vincristine resistance in P388 leukemia *in vivo* and *in vitro* through enhanced cytotoxicity of vincristine and vinblastine by verapamil. *Cancer Research*, **41**, 1967–1972.

Versantvoort, C. H., *et al*. (1995). Regulation by glutathione of drug transport in multidrug-resistant human lung tumour cell lines overexpressing multidrug resistance-associated protein. *British Journal of Cancer*, **72**, 82–89.

Wang, K., *et al*. (1999). Glutathione S-transferases in wild-type and doxorubicin-resistant MCF-7 human breast cancer cell lines. *Xenobiotica*, **29**, 155–170.

Wijnholds, J., *et al*. (1997). Increased sensitivity to anticancer drugs and decreased inflammatory response in mice lacking the multidrug resistance-associated protein. *Nature Medicine*, **3**, 1275–1279.

Wijnholds, J., *et al*. (1998). Multidrug resistance protein 1 protects the oropharyngeal mucosal layer and the testicular tubules against drug-induced damage. *Journal of Experimental Medicine*, **188**, 797–808.

Wijnholds, J., *et al*. (2000). Multidrug resistance protein 1 protects the choroid plexus epithelium and contributes to the blood-cerebrospinal fluid barrier. *Journal of Clinical Investigation*, **105**, 279–285.

Yang, C. P., *et al*. (1989). Progesterone interacts with P-glycoprotein in multidrug-resistant cells and in the endometrium of gravid uterus. *Journal of the Biological Chemistry*, **264**, 782–788.

Young, L. C., *et al*. (1999). Expression of multidrug resistance protein-related genes in lung cancer: correlation with drug response. *Clinical Cancer Research*, **5**, 673–680.

Zamora, J. M., *et al*. (1988). Physical-chemical properties shared by compounds that modulate multidrug resistance in human leukemic cells. *Molecular Pharmacology*, **33**, 454–462.

FURTHER READING

Borst, P., *et al*. (1999). The multidrug resistance protein family. *Biochimica Biophysica Acta*, **1461**, 347–357.

Gottesman, M. M. and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual Review of Biochemistry*, **62**, 385–427.

Harrison, D. J. (1995). Molecular mechanisms of drug resistance in tumours. *Journal of Pathology*, **175**, 7–12.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annual Review of Cell Biology*, **8**, 67–113.

Hipfner, D. R., *et al*. (1999). Structural, mechanistic and clinical aspects of MRP1. *Biochimica Biophysica Acta*, **1461**, 359–376.

Holland, I. B. and Blight, M. A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *Journal of Molecular Biology*, **293**, 381–399.

Chapter 86

Molecular Mechanisms of Radiotherapy

William H. McBride and Graeme J. Dougherty
University of California, Los Angeles, CA, USA

Luka Milas
M. D. Anderson Hospital and Tumor Institute, Houston, TX, USA

CONTENTS

- Introduction
- Ionizing Radiation and Free Radical Generation
- DNA Damage and Repair
- DNA Damage and its Relationship to Cellular Radiosensitivity
- Other Initiators and Sensors of Radiation Damage
- Pre-existing Molecular Pathways and the Response to Irradiation
- Conclusion

INTRODUCTION

Radiation therapy (RT) is an important modality in the treatment of cancer. The aim of RT is to deliver radiation doses and schedules that kill cancer cells while preserving normal tissue function. This therapeutic benefit is achieved in two ways. The first is by limiting the dose to the normal tissue by carefully planning the way in which the radiation is delivered. This involves exact calculation of the physical dose each organ receives. The second way is by exploiting the biological differences between tumour and normal tissues in their response to irradiation. This has led to the standard protocol for external beam RT, in which a dose on the order of 2 Gy is given daily to the tumour site for 5–8 weeks, depending on the clinical situation and the tolerance of the normal tissue in the field. This fractionated dose delivery regimen spares slowly proliferating normal tissues at the expense of tumours. (The unit of radiation dose, the gray, is equivalent to 100 rad and is an energy absorption of 1 J/kg.)

The most common forms of RT, whether delivered by external beam devices or radioactive implants, employ ionizing radiation. The biological effects of ionizing radiation vary with the dose, the dose rate, the physical nature of the radiation and the type of cell or tissue being irradiated. In other words, in addition to physical parameters, the manner in which cells sense and respond to radiation damage is critical for the outcome of RT. The molecular pathways that are triggered by radiation exposure act in an integrated fashion with those imposed by the molecular hardwiring of a cell to determine the multiple possible responses, whether they be DNA repair, cell cycle arrest, cell differentiation, cell death, adaptive responses, genomic instability or carcinogenesis. Most important from the perspective of RT is that these pathways provide a molecular explanation of why cell types differ in their response to irradiation.

The importance of the molecular phenotype is seen vividly in the clinic. Certain types of tumours, such as seminomas and lymphomas, respond well to RT while others, such as melanomas and glioblastomas, are relatively radioresistant. A similar rank order of radiosensitivity is obtained for different histological tumour types irradiated *in vitro* (Malaise *et al*., 1986). Furthermore, because molecular pathways within a cell can be altered by biological agents, such as growth factors, cytokines, monoclonal antibodies to cell surface receptors, or by gene transfer, it follows that the response to irradiation can be modified using the same approaches. Chemotherapeutic and physical agents, such as hyperthermia, hypoxia and radiation itself, can activate pathways that can also affect the outcome. The result is numerous novel approaches with potential to increase the efficacy of this powerful anticancer therapy. The molecular pathways that determine the intrinsic and induced cellular response to radiation are the focus of this chapter. For a broader discussion of radiobiological principles as they relate to RT, the reader is referred to standard texts (Withers and McBride, 1998; Hall, 2000).

IONIZING RADIATION AND FREE RADICAL GENERATION

Radiation is considered to have ionizing potential if it has a wavelength of $<10^{-6}$ cm. Such radiation has sufficient

energy in an individual photon to eject orbital electrons from a molecule or atom. The types of radiation used in RT, X-rays and γ-rays, produce several hundred thousand ionization events per cell per gray, but biological molecules are damaged largely indirectly. The absorbed energy causes ejection of primary electrons that go on to ionize other molecules leading to a complex chain reaction that continues until the energy is dissipated. In the process, free radicals, which are neutral atoms or molecules that have an unpaired electron, are generated. Because of their unpaired electrons, free radicals are very reactive and can reduce or oxidize biological molecules and break their chemical bonds.

Since the most abundant molecule in cells is water, the most common free radicals that are generated in a cell after exposure to ionizing irradiation are reactive oxygen species (ROS). Free radicals produced by radiolysis of water are responsible for about 60% of the biological damage caused by X-rays or γ-rays. The hydroxyl radical appears to be the most damaging primary product that is formed. Along with the hydrogen radical and aqueous electrons, it initiates cascade reactions that produce secondary products such as superoxide and hydrogen peroxide. Radicals are formed in organic molecules. In the presence of oxygen, additional radical peroxides, hydroperoxides and peroxides are formed that interfere with self-repair.

Some features of the molecular response to ionizing radiation can be explained by understanding the processes underlying these complex chemical events. First, the generation of ROS places ionizing radiation in the category of oxidative stress stimuli. The molecular signalling pathways that are responsive to ROS and to changes in redox homeostasis are part of the response to ionizing radiation. These pathways are shared with numerous other biological reactions, such as inflammatory processes, respiration, and damage caused by oxygen deprivation and reperfusion. Obviously, not all oxidative stress responses are the same and they do not always use the same signalling pathways. The spectrum of damage, and therefore the outcome, depend on the agent. Ionizing radiation is particularly effective at producing unrepairable lesions in DNA (Ward, 1988) which accounts for its relative efficacy in RT as a cytotoxic anticancer agent (see later).

A second consequence of the involvement of free radicals in radiation damage is that oxygen plays a major role as a modifier of radiation responses. Oxygen influences the nature of the free radicals and the lesions that are formed. The peroxides and hydroperoxides, in particular, inhibit repair. The result is that cells are typically 2.5–3 times more resistant to irradiation under hypoxic conditions than in the presence of oxygen (**Figure 1**). The concentration of oxygen must be low if radiation responses are to be affected; in general, less than 10 mmHg. Hypoxia is often considered significant in RT only when it reaches levels <5 mmHg, and reoxygenation between fractions may mitigate some of its influence. Some normal tissues may contain a small proportion of hypoxic cells, but only hypoxia within tumours is considered relevant in RT. Tumours vary considerably in their hypoxic status. There is also variation within one tumour. The chaotic organization of the tumour vasculature combines with tumour cell proliferation and cell death to form areas of chronic hypoxia with limited diffusion of oxygen, which may precede necrosis. Areas of acute reversible hypoxia also are found, which that may be the result of temporary closure of blood vessels, transient flow instability or changes in fluid pressure.

A number of studies have documented the importance of hypoxia to the outcome of RT. This may be due simply to depriving cells of the radiosensitizing effect of oxygen, but tumour hypoxia has other dimensions that must be considered. At the molecular level, hypoxia induces expression of a number of genes, in particular genetic programmes that are under the control of hypoxia inducible factor (HIF-1). Some of these (such as erythropoietin,

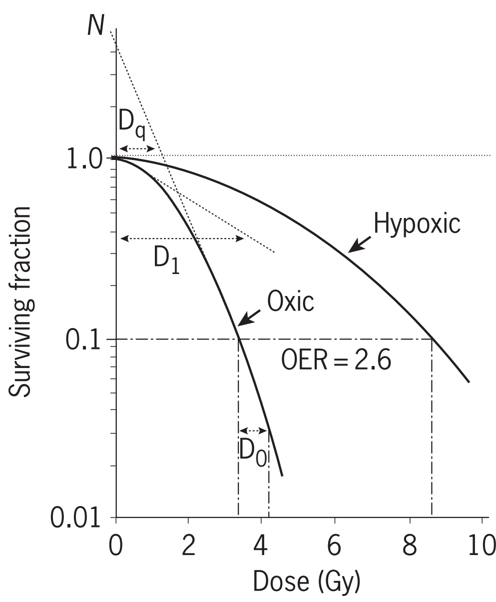


Figure 1 Survival curves for mammalian cells exposed to X- or γ-rays under oxic or hypoxic conditions. Cell survival is plotted on a logarithmic and radiation dose on a linear scale. There are two components to cell killing by radiation. At low doses, the survival curves have an initial linear slope (designated D_1). At higher doses the slope is steeper (designated D_0). D_q is a measure of the size of the shoulder. N is the extrapolation number on the ordinate. The oxygen enhancement ratio (OER) is the ratio of the dose required for a given effect under hypoxic conditions divided by the dose required to produce the same effect in oxic conditions. Since the slopes of the initial and final slopes depend on the presence of oxygen, the OER is 2.5–3.0 for high dose-related effects (2.6 for 10% survival in the figure) but is typically less for lower dose-related effects.

vascular endothelial growth factor and tumour necrosis factor alpha (TNF-α)) are clearly aimed at increasing angiogenesis and increasing oxygen delivery. Hypoxia can therefore play an important role in driving angiogenesis and tumour expansion. Other hypoxia-induced genes (such as *p53*) are part of a stress response that encourages cells to undergo cell death by apoptosis. Acting in this way, hypoxia serves as a selective force to favour expansion of cells mutated in *p53* (Green and Giaccia, 1998). These tumours will have an antiapoptotic and therefore a more malignant phenotype. These, and other, phenotypic and genetic cellular alterations induced by hypoxia, can affect the intrinsic cellular response to radiation through altering the molecular hardwiring of the cell (see later).

A third consequence of free radical involvement in the response to radiation is that agents that scavenge free radicals have potential as radioprotectors whereas those that mimic oxygen in 'fixing' radiation damage can act as radiosensitizers. The role that intrinsic redox active molecules play in the radiation response will depend on their specific intracellular location.

DNA DAMAGE AND REPAIR

Ionizing radiation is nondiscriminatory in that all molecular species in a cell may be damaged. However, DNA is the most important target for RT. The evidence comes from experiments showing that irradiation of the nucleus, but not the cytoplasm, results in cell death. The incidence of radiation-induced chromosome aberrations also correlates with cell death. The implications are that the most relevant radiation-induced free radicals are those generated within 2–4 nm of the DNA and that the concentration of redox-active agents in this location can influence the amount of damage that results.

Each gray of ionizing radiation causes in each cell thousands of single-strand DNA breaks (SSB), cross-links in DNA and with DNA-associated proteins, and 30–40 double-strand DNA breaks (DSB). Most of this damage is repaired. The most significant lesions from an RT standpoint are the unrepairable DSBs that arise from multiple events occurring close together in a short stretch of DNA (Ward, 1998). These lethal lesions are large, being of the order of 10–20 nucleotides in size, and represent less than a few per cent of the total DSB that are formed. Perhaps on average 25% of cells are killed by 1 gray although this percentage varies widely with the cell line. The relationship between survival and dose is generally log-linear (**Figure 1**), but, as the dose is increased, the probability of forming a multiply damaged site by interactions between lesions formed by separate ionization tracks increases and the slope of the dose–survival curve becomes steeper. The random nature of inactivation events means that a much higher radiation dose (by several $\log_{10}$) will be needed to sterilize organisms such as bacteria because they have considerably less DNA than do mammalian cells.

DNA repair mechanisms are, by necessity, highly efficient and come in many different forms. The involvement of any one repair mechanism is determined in large part by the nature of the damage, but other factors are influential. For example, damage in transcriptionally active genes or in S phase is more rapidly repaired. Repair of SSB and base damage is facilitated by the presence of an intact complementary DNA strand that can act as a template. A consequence is that repair of SSB is rapid with a half-time of minutes. However, overall repair of the diverse potentially lethal lesions is a slower process with an average half-time of around 1 h. 'Slow' and 'fast' repair components have been demonstrated in a number of systems (Canney and Millar, 1997). A comprehensive discussion of DNA repair mechanisms is beyond the scope of this chapter, which will concentrate on the relationship between repair and other molecular pathways, and the reader is referred to a more comprehensive review (Haber, 2000).

DNA repair mechanisms have evolved, in large part, to deal with errors produced during DNA replication as well as with damage resulting from free radicals generated through normal metabolism. Base excision repair processes with formation of apurinic/apyrimidinic sites and the sequential action of endonucleases, DNA polymerases and ligases repair most SSB and base damage. One of the earliest events following DNA strand breakage by radiation is the poly(ADP–ribosyl)ation of surrounding proteins by poly(ADP–ribose) polymerase (PARP). PARP binds to SSB and DSB to catalyse the transfer of successive units of ADP–ribose from NAD($^{+}$) covalently to itself and other nuclear acceptor proteins, including p53. PARP cycles on and off DNA, depending on its degree of polyribosylation. PARP-deficient cells are more radiosensitive and exhibit genomic instability and PARP cycling may be important in apoptosis, DNA replication and DNA repair. PARP has recently been identified as one of the intracellular 'death substrates' that can be degraded by proteolytic caspases activated during apoptosis. The finding that several molecules that are involved in the response to DNA damage are caspase substrates suggests a degree of coordination between repair and cell death processes, although PARP itself may play only a secondary role in repair. This may be true of a number of molecules that bind broken DNA ends, including the tumour-suppressor protein p53. Such molecules may influence radiosensitivity more by activating downstream pathways. The relationship between binding of proteins to DNA strand breaks, repair of those breaks and activation of downstream intracellular signalling pathways may be better developed in DSB repair.

The repair template that is used in the process of DSB repair is less easily identified than it is in SSB repair, but physiological models have shed light on how repair is

achieved. DSBs are created during meiosis and during the generation of diversity in the immune system. For example, in developing thymocytes, DSBs are generated by endonucleases RAG-1 and RAG-2 during initiation of V(D)J recombination within the T cell receptor (TCR) locus. Ligation of the DSBs that are formed is essential for the generation of the TCR repertoire. Very similar processes in B lymphocytes are responsible for the generation of immunoglobulin diversity. Failure to generate functional immunoglobulin or T cell receptors results in apoptosis. Ligation involves several proteins that are components of the nonhomologous end-joining pathway (NHEJ) of DNA damage repair (**Figure 2**). Ku autoantigen is a DNA end-binding heterocomplex of Ku70 and Ku80 (Jin and Weaver, 1997) that activates the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to form the active kinase DNA-PK (Chen *et al*., 1996). DNA-PK initiates DSB repair by processes that are poorly understood, although an essential role for binding of DNA ligase IV and XRCC4 has recently been demonstrated. Importantly, DNA-PK also appears to trigger signal transduction pathways that are involved in stress responses. If any of the DNA-PK complex of molecules is defective, both V(D)J rejoining and sensitivity to ionizing radiation are affected (Gu *et al*., 1997). The best known example of this link is the scid (severe combined immune deficiency) syndrome, which in mice and humans is caused by a mutation in DNA-PKcs and is associated with radiation sensitivity (Jhappan *et al*., 1997).

An alternative pathway for DSB repair is homologous recombination (HR). The difference between HR and NHEJ lies in their dependence on DNA homology and the degree of accuracy of the repair. HR uses the sister chromatid or homologous chromosome. NHEJ uses no or limited homology and inevitably makes more errors. Unlike yeast cells, mammalian cells appear to rely more on NHEJ as the principal method of DSB repair. Very little is known about how cells elect to perform HR or NHEJ; the choice may vary with the phase of the cell cycle.

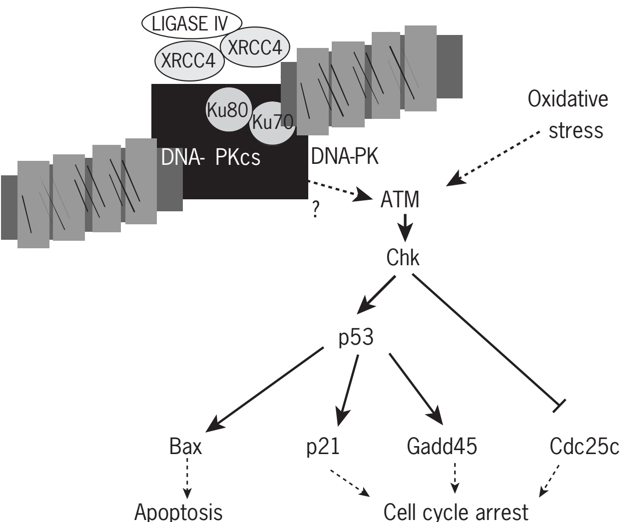


Figure 2 Nonhomologous end joining proteins involved in DSB repair recognize DNA end (Ku) and activate DNA-PKcs to the active kinase DNA-PK. XRCC4 and ligase IV are essential for repair. The ATM gene product is activated by DNA damage and by oxidative stress and initiates downstream p53-dependent pathways leading to apoptosis, cell cycle arrest or cell cycle progression.

Like NHEJ, HR is a complex process with subpathways involving gene conversion, single-strand annealing and break-induced replication. The key proteins were most often identified in yeast before their human homologues were cloned. They include active complexes of Rad51, Rad52 and Rad54, as well as complexes of BRCA1, NBS (Nijmegen breakage syndrome product), MRE11 and Rad50, and complexes of Rad51 with BRCA1 and BRCA2. Identification of BRCA tumour-suppressor gene products as sensors of DNA damage is of interest because these are mutated in hereditary cases of breast and ovarian cancer. Functional loss of BRCA may result in an increase in cellular radiosensitivity (Shen *et al*., 1998) although the extent to which this influences clinical outcome of RT has yet to be fully determined. Mutation in any of the HR gene products increases the risk of carcinogenesis as well as influencing radiosensitivity.

A major unanswered question is the extent to which variation in the level of expression of DNA repair enzymes influences the outcome of RT. Variation in DSB repair enzymes has been observed between different cells and tissues and it is known that considerable control is exerted at the post-translational level. However, so far this variation has not been convincingly shown to correlate with cellular or tissue radiosensitivity (Allalunis *et al*., 1995; Carlomagno *et al*., 2000). A particular problem is that the enzymes that critically determine the rate and extent of DNA repair by the different repair processes have, for the most part, yet to be determined. Such studies may indicate which molecules are most relevant to the outcome of RT.

DNA DAMAGE AND ITS RELATIONSHIP TO CELLULAR RADIOSENSITIVITY

Classically, DNA damage has long been considered the major factor causing cell death. However, recently, an important concept has developed from studies in molecular radiobiology. It is that cells possess sensors of radiation-induced damage that link damage to cellular outcome through stress-responsive signal transduction pathways (reviewed by Schmidt-Ullrich *et al*., 2000). The pathways that are involved vary to an extent with the cell type and the cell's microenvironmental context. However, in spite of these sources of variation and the problems imposed by the complexity of the integrated nature of the circuitry within cells, patterns are emerging that allow critical molecules and pathways to be recognized. Certain patterns

link DNA repair, DNA misrepair or failure to repair to classical outcomes of radiation exposure. Others identify non-genomic initiators of signal transduction pathways as influences.

The extent of initial DNA damage and the persistence of this damage are presumably critical factors in determining the cellular response to radiation, but how this information is usefully coordinated within the cell is still unclear. Assessment is made more difficult by feedback control from downstream pathways, such as those leading to apoptosis. When activated, these can influence DNA damage and repair processes. For example, DNA repair enzymes can serve as substrates of active caspases. Finally, there is a tendency to view these pathways as leading in a linear fashion to transcriptional activation of genetic programmes that mediate responses. This is true only to an extent. Post-transcriptional and post-translational mechanisms are also involved, in particular in the initial response to stress stimuli. Stabilization of mRNA and protein expression is the most rapid way to initiate a stress response and expression of many of the molecules involved in DNA repair, cell cycle and cell death are controlled, at least in part, at this level (Moll *et al*., 1999). There is little known about how radiation affects these levels of control but recently radiation-induced inhibition of proteasome function has been demonstrated (Pajonk and McBride, to be published) and this may be a mechanism of potential importance.

The two cellular responses to irradiation that have long been recognized as being highly relevant to RT are cell cycle arrest and cell death. Key transitions between cell cycle phases are regulated by the activities of various protein kinase complexes composed of cyclin and cyclin-dependent kinase (Cdk) molecules. Radiation activates multiple gene products that interfere with their interactions resulting in arrest at G1/S, S, or G2 phase checkpoints. The radiation dose dependence of arrest at the different checkpoints varies, with G2 arrest being the most dose sensitive and consistent response. It seems logical to suppose that a function of checkpoint arrest is to allow cells time to repair DNA damage before proceeding in their cell cycle. DNA damage is, however, not a prerequisite for cell cycle arrest, which can also be caused by exposure to growth regulatory cytokines. Another interpretation of the function of cell cycle arrest, which may be particularly applicable to G1 arrest, is that signals from all sources are evaluated so that cells with inappropriately activated pathways can be eliminated prior to proceeding through growth factor-independent cell-cycle phases that result in replication. Cells that fail to arrest at G1, such as cells with mutated *p53*, are less likely to undergo apoptosis at the G1/S checkpoint and may replicate lesions in S phase resulting in increased DNA mutations and fragility. For cells entering mitosis, the integrity of the chromosomes and the replication apparatus determines if the cell dies and repair during G2 arrest may restore this. In contrast with the G1/S checkpoint, when radiation-induced G2 arrest is abrogated, cells are more likely to die. The use of mice deficient in specific genes has been particularly useful in elucidating the cascade linking DNA damage to cell cycle arrest and cell death. Three major links in the chain are the ATM (ataxia telangiectasia mutated), the checkpoint kinase Chk2 and the p53 tumour-suppressor gene proteins (**Figure 2**).

Ataxia telangiectasia (A-T) is the archetypal human radiation sensitivity syndrome (reviewed by Lavin and Khanna, 1999) and ATM plays important roles in the maintenance of the cell homeostasis in response to oxidative damage (Takao *et al*., 2000). ATM is a member of a family of large phosphatidylinositol 3-kinase domain-containing proteins that includes DNA-PKcs. DNA-PK can phosphorylate ATM *in vitro*, but its role as the initiator of the cascade *in vivo* still needs substantiation. Some studies have indicated that p53-mediated apoptosis and cell cycle arrest are normal in DNA-PKcs knockout mice (Jongmans *et al*., 1996; Jhappan *et al*., 2000). Others suggest that DNA-PKcs selectively regulates p53-dependent apoptosis but not cell-cycle arrest (Wang *et al*., 2000). In studies of this type, the control and experimental cells have different levels of DNA damage/repair and radiosensitivity and this may complicate interpretation of dose–response relationships.

The role of ATM in the response to radiation-induced DNA damage has been investigated in some depth. ATM-deficient cell lines, unlike normal cells, do not up-regulate p53 expression after exposure to ionizing radiation, which indicates that ATM is upstream of p53. They do not undergo proper G1, S and G2 phase checkpoint arrest and they are hypersensitive to irradiation. Cells from A-T patients express the same phenotype. A number of target proteins have been identified for ATM. They include c-Abl, Chk2, BRCA-1, Ik$B\alpha$, MDM2 and replication protein A (RPA). This spectrum of proteins indicates the key position ATM holds in control of cell cycle arrest, cell replication, cell survival and death. The checkpoint kinase Chk2/hCds1 (Caspari, 2000) is involved in G2 arrest. This is because Chk2 phosphorylates the phosphatase Cdc25C, which is required to activate the Cdk for entry into mitosis (Chaturvedi *et al*., 1999).

Chk2 also phosphorylates p53 on serine 20, which interferes with the binding of its negative regulator Mdm2 (Mirao *et al*., 2000). This could affect its degradation through the ubiquitin/proteasome system and be responsible for stabilization of p53 expression, which is a feature of the normal response of cells to ionizing radiation. p53 is a multifunctional molecule. Amongst its many functions, it acts as a transcription factor to activate many downstream pathways, including ones that lead to cell-cycle arrest and apoptosis. It also induces production of its own negative regulator Mdm2, which could be viewed as an adaptive response to irradiation. The fact that *p53* is mutated in approximately 50% of human tumours indicates the importance of this pathway in preventing tumorigenesis.

The mechanism by which elevated p53 levels leads to apoptosis is not fully known, but transcriptional activation of apoptotic genes, such as *BAX*, is one mechanism. p53 activation can also cause cell cycle arrest at the G1/S transition. Pathways to cell-cycle arrest converge to induce or phosphorylate Cdk inhibitors (CKIs) and antagonize the formation of cyclin–CDK complexes to block entry into S phase (Shackelford *et al*., 1999). CKIs come from three families, Cip (e.g. p27)/Kip (e.g. p21), Ink4 (e.g. p16) and the pRb pocket-protein family. p53-dependent cell cycle arrest in G1 is largely through the p21/WAF1 pathway that inhibits CDK2 kinases. Other targets for p21 have been identified. For example, p21 forms complexes with PCNA to inhibit DNA synthesis. It may also play a role in maintaining G2 arrest after irradiation, although Gadd45, which is another p53-regulated stress protein, may be more important for arrest at the G2/M checkpoint (Jin *et al*., 2000).

Cell cycle arrest and apoptosis can be initiated through pathways other than those going through p53, but p53 knockout mice have distinct features that clearly indicate the central importance of this tumour-suppressor gene in radiation responses. Mouse embryo fibroblasts from *p53−/−* mice, lack radiation-induced G1/S arrest in addition to having other cell cycle checkpoint abnormalities, and are more radioresistant than wild type cells (Whang *et al*., 2000). On the other hand, damage accumulated consequent to *p53* mutation may also affect cellular radiation sensitivity (Denko *et al*., 2000). The *in vivo* relationship between p53 and apoptosis after irradiation has been investigated in several tissues. Apoptosis occurs spontaneously in cells going through specific developmental stages within tissues. It has a homeostatic role in regulating the numbers of cells in stem cell areas or sites of rapid proliferation. Irradiation strikingly increases the extent of apoptosis in these specific populations (Potten *et al*., 1994). This suggests that cells predisposed to apoptosis are most readily and most rapidly killed by irradiation. Radiation-induced apoptosis is frequently associated with rapid up-regulation of p53 expression, e.g. in the stem cell areas of the jejunal crypt, although the p53 dependence is not present in all sites. Introduction of wild-type *p53* into tumour cells lacking *p53* generally results in an increase in radiation sensitivity (McBride and Dougherty, 1995) that is frequently associated with an increased tendency towards apoptosis.

What has not been fully explained by use of *p53−/−* mice is the relationship between apoptosis and the final tissue response to irradiation. *In situ* assays show that survival of stem cell clonogens following irradiation may be unaltered in *p53−/−* mice, even if radiation-induced apoptosis is defective (Hendry *et al*., 1997). Since p53 has more than 100 downstream target genes, knockout of this gene affects many functions and the response to irradiation could be influenced through pathways other than those involved in apoptosis. The same argument holds for lack of correlation that is often found between *p53* mutations in human tumours and response to RT. Multiple pathways are affected by loss of one gene and, in cancer, multiple genes are mutated. While *p53*, like *DNA-PK*, *ATM*, *CHK2*, *BAX*, *p21* and other genes in this pathway, are important in mediating the cellular response to irradiation, signal transduction pathways have many tributaries. Branches of different pathways may neutralize or accentuate each other. The outcome can vary with the cell type depending on the way radiation-induced signals integrate into the hardwiring in the cell, and upon the overall gestalt of the multiple signals the cell is receiving through multiple pathways (Bouvard *et al*., 2000).

In spite of these reservations, there is little doubt that the molecular pathways linking DNA damage to cell cycle arrest and cell death are important in determining the outcome of RT. Indeed, the relationship extends further. Cancer susceptibility genes have been classified into two groups; caretakers or gatekeepers. In general, caretakers are DNA repair genes whose inactivation leads to genetic instability and a marked increase in chromosomal aberrations, including breakage, translocations and aneuploidy. This does not normally lead to a dramatic increase in the risk of cancer unless gatekeeper genes that control cell proliferation and death are also affected. For example, loss of Ku80 synergizes with loss of *p53* to promote early tumorigenesis (Difilippantonio *et al*., 2000). Communication between molecular programmes involved in DNA repair, cell cycle and apoptosis is therefore essential for the integration of DNA integrity and cellular homeostasis.

OTHER INITIATORS AND SENSORS OF RADIATION DAMAGE

Although DNA may be the most important initiator of signalling responses following cellular exposure to ionizing radiation, there is increasing evidence for critical contributions from other sources. Oxidative damage to cytoplasmic, mitochondrial and membrane structures, changes in the homeostatic redox balance, alterations in protease activity and the production of lipid peroxides in membranes following radiation have all been suggested as influences on the cellular response to irradiation. Mitochondrial damage may be particularly important in amplifying some of the ROS effects of ionizing radiation. The molecular pathways that mediate these effects have in some cases been identified, but their importance is still a matter of some controversy and will vary with the cell type and its microenvironmental context.

Ionizing radiation shares with cytokine death signals, such as TNF-α and Fas ligand (FasL), the ability to generate the second messenger ceramide through a sphingomyelin pathway (Haimovitz-Friedman *et al*., 1997). The phospholipid sphingomyelin is hydrolysed in the cell

membrane by activation of neutral or acidic sphingomyelinases or ceramide synthase. The consequences of ceramide production depend on the cell type, but it most commonly serves to initiate an apoptotic response. Recent work has defined two cascades activated preferentially by pro-inflammatory cytokines, such as TNF-α and IL-1, as well as ceramide, ionizing radiation and stress (**Figure 3**). One converges on the mitogen-activated protein kinase (MAPK) subfamily known as the 'stress activated' protein kinases (p38 SAPKs). The second pathway recruits the Jun N-terminal kinases (JNK)s. These signalling pathways are sensitive to intracellular thiol-redox and protease–antiprotease status and are linked to activation of at least two well-defined transcription factors, NF-κB and AP-1 (c-Fos and c-Jun). The apoptotic response resulting from activation of the sphingomyelin pathway, at least in some cell types, involves release of cytochrome *c* from mitochondria, and is particularly influenced by the actions of members of the Bax/Bcl-2 pro- and antiapoptotic family of proteins that functions at this level (Tepper *et al*., 1999).

The close association between radiation exposure and inflammation is also seen in radiation-induced responses. *In vitro* and *in vivo* irradiation of cells and tissues frequently increases expression of TNF-α, FasL, IL-1, IL-6 TGF-β and other cytokines and cytokine receptors, as well as cell adhesion molecules and proteases/antiproteases (Kasibhatla *et al*., 1998; Hong *et al*., 1999). This 'danger' response plays an important role in initiating tissue recovery and remodelling. NF-κB-dependent pathways are major components that are further amplified by infiltrating host inflammatory and immune cells. The extent of the response varies with tissue and dose. The classical dose of 2 Gy does not induce a particularly strong pro-inflammatory response, which may be part of the reason for its use in the clinic. The other more substantiated reason is that 2 Gy spares late-responding, slowly proliferating normal tissue at the expense of more rapidly proliferating tumours. The proinflammatory response to radiation may be particularly relevant in situations where bystander effects of radiation have been observed and in adaptive responses to subsequent exposures. Proinflammatory cytokines, in consort with their up-regulated receptors, provide a positive feedback loop that may affect subsequent responses to radiation.

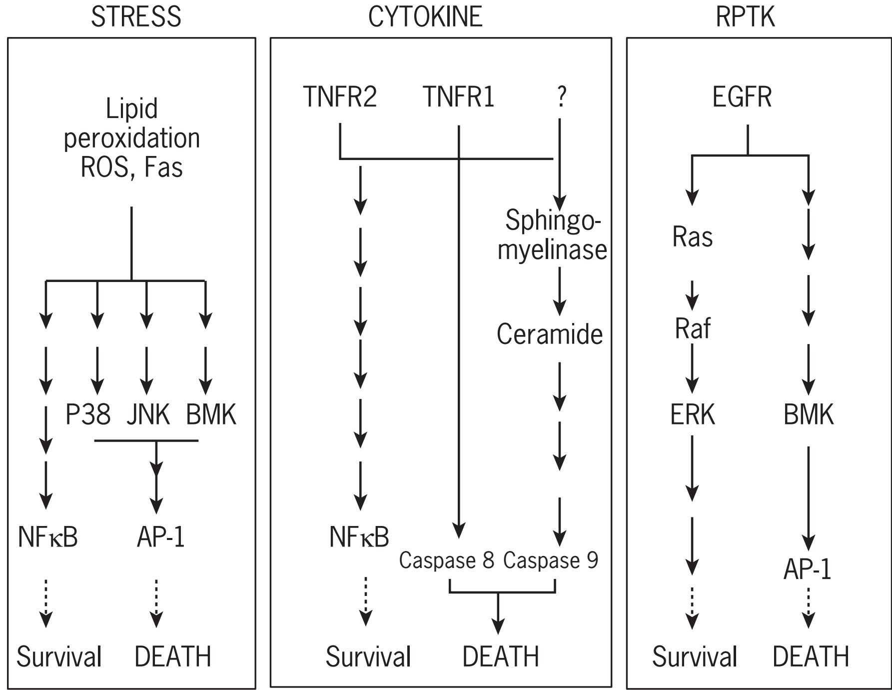


Figure 3 Responses to radiation can be initiated by several mechanisms other than DNA damage. This oversimplified diagram indicates three general pathways that influence the cellular outcome of irradiation. The pathways are distinguished by their preferential utilization by distinct classes of stimuli, but there is considerable cross-talk. For example, the mitogen activated kinase (MAPK) pathways involving c-Jun N-terminal kinase (JNK) and p38/stress-activated protein kinase (SAPK) are activated by many stimuli, including cytokines and lipopolysaccharide. The big MAP kinase pathway (BMK) is highly redox-sensitive, but is also activated through protein tyrosine kinase receptors (RPTK) such as epidermal growth factor receptor (EGFR), as are extracellular signal-regulated kinases (ERK1/2). The sphingomyelinase pathway feeds into a stress response involving ROS. The pathways converge on multiple downstream factors that initiate transcriptional and nontranscriptional mechanisms in a complex combinatorial fashion, but preferentially appear to favour cell death or survival, although the outcome is, to an extent, cell-type dependent.

Two major transcription factors, AP-1 and NF-κB, mediate responses to free radicals, ceramide, proinflammatory cytokines and ionizing radiation. AP-1 has been implicated in certain apoptotic responses. In contrast, NF-κB has generally been found to encourage survival of many cell types, most likely through activation of inhibitors of apoptosis (IAP) family members (You *et al*., 1997). The classical pathway for NF-κB activation is in response to signals like TNF-α. IκBa, which is bound to NF-κB in the cytoplasm, becomes phosphorylated on serine residues and is degraded through the ubiquitin/proteasome system. NF-κB is freed to translocate to the nucleus to initiate transcriptional programs. Ionizing radiation activates NF-κB, at least in some cell types. However, an alternative pathway that may involve tyrosine phosphorylation may operate (Raju *et al*., 1998) as it does in response to certain other oxidative stresses. This emphasizes the danger of generalizing about common pathways activated by diverse agents. The consequences of constitutive or radiation-induced AP-1 and NF-κB on the response to radiation is not completely established. Inhibition of NF-κB expression induces apoptosis in many cell lines (Duffey *et al*., 1999) and may radiosensitize, but this is not always the case (Pajonk *et al*., 1999).

Another signal transduction cascade generated at the plasma membrane by radiation involves activation of cell membrane receptors, including that for epidermal growth factor (EGFR). This tyrosine kinase receptor has been shown to autophosphorylate shortly after irradiation of tumour cells both *in vitro* and *in vivo* (Akimoto *et al*., 1999; Todd *et al*., 1999). Other cytoplasmic and membrane-bound kinases are also probably activated after radiation exposure. The mechanism of receptor activation is not clear. ROS action on cysteine residues may lead to cross-linking and receptor aggregation. Alternatively, inactivation of phosphatases might be involved. EGFR-mediated pathways trigger the prototypical p44 and p42 MAP kinases (ERK1, ERK2) and the newly described big mitogen-activated kinase-1 (BMK1). The former pathway is activated by many growth factors and oncogenes, including *ras* and *raf*, which a number of studies have shown to be associated with cellular resistance to radiation. Both p44/p42 and BMK pathways may be activated by ROS, but the latter seems particularly redox sensitive. Both seem to generate cytoprotective responses, at least in certain cell types (Schmidt-Ulrich *et al*., 2000).

PRE-EXISTING MOLECULAR PATHWAYS AND THE RESPONSE TO IRRADIATION

The fact that cell types vary in their response to irradiation indicates the importance of molecular pathways that are hardwired into the cell. The best example of this comes from gene transfer studies aimed at disrupting these pathways. Expression or knockout of proto-oncogenes, tumour-suppressor genes, cytokines, cytokine receptors, cell adhesion molecules, redox-active genes and many other genes that are important in determining cell behaviour, can influence the outcome of irradiation (McBride and Dougherty, 1995; Dougherty *et al*., 1996; Biaglow *et al*., 1997; Chiang *et al*., 1998). For example, radiation resistance can often be achieved by transfer of genes for growth factors or growth factor receptors that cause cell proliferation, tyrosine kinase genes or products such as Ras and Raf on MAPK1/2 pathways, or mitochondrial proteins such as Bcl-2 or manganese superoxide dismutase. A dominant negative approach (Reardon *et al*., 1999) antisense, or antibody (Milas *et al*., 2000) directed against, for example EGFR, can result in cellular radiosensitization and the level of expression of EGFR by a tumour can determine radiocurability (Akimoto *et al*., 1999). High constitutive NF-κB levels may also affect radiocurability and targeting this pathway using a COX-2 inhibitor (Kishi *et al*., 2000), or other approaches, seems feasible. Radiosensitization often results from transfer of cytokine genes or receptors that slow cell cycle progression or encourage apoptosis.

It follows that carcinogenesis-related molecular pathways in tumours may determine the outcome of RT. They do this not because they are directly involved in transformation *per se*, but because they influence pathways downstream of damage recognition (Chiang *et al*., 1998) As has been mentioned, carcinogenesis is a multistep process and interactions between pathways will occur that may not always give a predictable outcome in terms of radiation response. Assessments of patterns of gene expression are likely to give more reliable information than examination of any one gene product. The outcome of ongoing studies to determine the value of gene array studies for predicting outcome of RT is awaited with interest.

If constitutive gene expression is important in radiation responses, a further logical extrapolation is that genetic differences between individuals may result in different responses to irradiation. A classical example is the differences between C57Bl and C3H mouse strains in response to irradiation (Dileto and Travis, 1996; Franko *et al*., 1997). Several genes are responsible that have been mapped, some within the major histocompatability complex. In addition, mice lacking the TNFR2 molecule, which can trigger NF-κB responses, have recently been shown to be radiosensitive (Daigle *et al*., 2000). Genetic differences in the radiation response of fibroblasts from different individuals have also been observed (Eastham *et al*., 1999). Since the dose that is normally given in RT is often based on the dose that the normal tissue in the field can be expected to tolerate, identification of radiosensitive and radioresistant individuals would allow more individual tailoring of the maximum dose that can be delivered. This could increase the cure rate. Screening individuals to determine if they are carriers of the *ATM* gene, or other

genes known to confer radiosensitivity, might help in this regard, but recognition of patterns of response by multiple genes may be more universally valuable.

CONCLUSION

Recognition by a cell that it has received radiation damage has numerous downstream consequences. DNA lesions that can not be repaired may result in chromosome abnormalities that will interfere with cell function and cause death directly. At the same time and in addition, radiation initiates genomic and non-genomic molecular pathways that network with constitutive pathways to form integrated signals that are channelled into making concerted and meaningful cellular responses. Such responses may, in part, be characteristic of a given cell type and its environment but, all other things being equal, the failure or success of RT is determined by the characteristics of these pathways. Specific groups of genes are activated, progression through cell cycle checkpoints is blocked, replication and transcription are affected, apoptosis may be induced, tissue recovery and remodelling processes are initiated and the tissue is adaptively primed in preparation for future insults.

The functional relevance of any single molecular response that is observed may not always be clear, but the purpose of the response overall is presumably damage control at the level of the individual cell and signalling damage at the tissue and whole animal level to initiate healing. Apoptosis following radiation exposure is an altruistic response that may prevent carcinogenesis from misrepaired or unrepaired DNA damage. In support of this suggestion are findings that mice lacking proapoptotic genes generally have an increased incidence of cancer. On the other hand, only a small proportion of irradiated normal and tumour cells die by rapid apoptosis and this pathway can often be subverted by stimulation by growth factors. It seems that most cells that die after irradiation die a mitotic death, sometimes after several division cycles, depending on the dose. Mitotic death can have an apoptotic component, but more often it is necrotic. This failure to undergo rapid apoptosis in response to irradiation has important implications for tumour sterilization; as cells divide, a proportion of the progeny may die, but the clone may still survive. A more important role for radiation-induced apoptosis in a tissue may be to initiate repopulation, although solid evidence for this hypothesis is still lacking.

Although radiation-induced death is in many cases nonapoptotic, the order of radiosensitivity of tumours of different histological types suggests that those cells with a propoptotic tendency are more radiation sensitive. Based on what we know about the level of integration and cross-talk between molecular pathways, it is possible that the proapoptotic phenotype overlaps at the molecular level with other processes, such as DNA repair or proliferation. If this is the case, apoptosis may not be the best measure of radiosensitivity, but the pathways that are involved would still be useful targets for therapeutic intervention. On the other hand, if the apoptosis and repair pathways are distinct, increasing the fraction of cells that apoptose following irradiation may have only limited overall impact on cure rates.

Molecular radiobiology has clearly reached the stage where preliminary conclusions can be reached as to the impact of specific signalling pathways on the outcome of RT. The studies have exposed numerous potential targets for therapeutic manipulation of radiation responses. The general proven efficacy of ionizing radiation as therapy against many tumours, together with the knowledge that has accumulated over the last century of treatment, gives an excellent base from which to launch these novel approaches to extending the usefulness and increasing the efficacy of RT. Combined with molecular evaluation of the probability of tumour cure and normal tissue complications in individual patients, as well as more precise delivery methods, the future of RT looks promising.

REFERENCES

Allalunis, T. M., *et al*. (1995). Lack of correlation between DNA-dependent protein kinase activity and tumor cell radiosensitivity. *Cancer Research*, **55**, 5200–5202.

Akimoto, T., *et al*. (1999). Inverse relationship between epidermal growth factor receptor expression and radiocurability of murine carcinomas. *Clinical Cancer Research*, **5**, 2884–2890.

Biaglow, J. E., *et al*. (1997). Effect of oncogene transformation of rat embryo cells on cellular oxygen consumption and glycolysis. *Biochemical and Biophysical Research Communications*, **235**, 739–742.

Bouvard, V., *et al*. (2000). Tissue and cell-specific expression of the *p53*-target genes: *bax*, *fas*, *mdm2* and *waf1/p21*, before and following ionising irradiation in mice. *Oncogene*, **19**, 649–660.

Canney, P. A. and Millar, W. T. (1997). Biphasic cellular repair and implications for multiple field radiotherapy treatments. *British Journal of Radiology*, **70**, 817–822.

Carlomagno, F., *et al*. (2000). Comparison of DNA repair protein expression and activities between human fibroblast cell lines with different radiosensitivities. *International Journal of Cancer*, **85**, 845–849.

Caspari, T. (2000). How to activate p53. *Current Biology*, **10**, R315–R317.

Chaturvedi, P., *et al*. (1999). Mammalian Chk2 is a downstream effector of the ATM-dependent DNA damage checkpoint pathway. *Oncogene*, **18**, 4047–4054.

Chiang, C., *et al*. (1998). Oncogene expression and cellular radiation resistance: a modulatory role for c-myc. *Molecular Diagnosis*, **3**, 21–28.

Chen, F., *et al*. (1996). Disruption of DNA-PK in Ku80 mutant *xrs-6* and the implications in DNA double-strand break repair. *Mutation Research*, **362**, 9–19.

Daigle, J. L., *et al*. (2000) Acute molecular and cellular responses of TNF receptor knockout mice to brain irradiation. *Proceedings of the American Association for Cancer Research Annual Meeting* (41), 478.

Denko, N. C., *et al*. (2000). p53 checkpoint-defective cells are sensitive to X-rays, but not hypoxia. *Experimental Cell Research*, **258**, 82–91.

Difilippantonio, M. J., *et al*. (2000). DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature*, **404**, 510–514.

Dileto, C. L. and Travis, E. L. (1996). Fibroblast radiosensitivity *in vitro* and lung fibrosis *in vivo*: comparison between a fibrosis-prone and fibrosis-resistant mouse strain. *Radiation Research*, **146**, 61–67.

Dougherty, G. J., *et al*. (1996). Gene therapy-based approaches to the treatment of cancer: Development of targetable retroviral Vectors. *Transfusion Science*, **17**, 121–128.

Duffey, D. C., *et al*. (1999). Expression of a dominant-negative mutant inhibitor-kappaBalpha of nuclear factor-kappaB in human head and neck squamous cell carcinoma inhibits survival, proinflammatory cytokine expression, and tumor growth *in vivo*. *Cancer Research*, **59**, 3468–3474.

Eastham, A. M., *et al*. (1999). Fibroblast radiosensitivity measured using the comet DNA-damage assay correlates with clonogenic survival parameters. *British Journal of Cancer*, **79**, 1366–1371.

Franko, A. J., *et al*. (1997). Immunohistochemical localization of transforming growth factor beta and tumor necrosis factor alpha in the lungs of fibrosis-prone and 'non-fibrosing' mice during the latent period and early phase after irradiation. *Radiation Research*, **147**, 245–256.

Green, S. L. and Giaccia, A. J. (1998). Tumor hypoxia and the cell cycle: implications for malignant progression and response to therapy. *Cancer Journal from Scientific American*, **4**, 218–223.

Gu, Y., *et al*. (1997). Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proceedings of the National Academy of Sciences of the USA*, **94**, 8076–8081.

Haber, J. E. (2000). Partners and pathways repairing a double-strand break. *Trends in Genetics*, **16**, 259–264.

Haimovitz-Friedman, *et al*. (1997). Ceramide signaling in apoptosis. *British Medical Bulletin*, **53**, 539–553.

Hall, E. J. (2000). *Radiobiology for the Radiologist*, 5th edn. (J. B. Lippincott, Philadelphia).

Hendry, J. H., *et al*. (1997). p53 deficiency sensitizes clonogenic cells to irradiation in the large but not the small intestine. *Radiation Research*, **148**, 254–259.

Hirao, A., *et al*. (2000). DNA damage-induced activation of p53 by the checkpoint kinase Chk2 [see comments]. *Science*, **287**, 1824–1827.

Hong, J. H., *et al*. (1999). Rapid induction of cytokine gene expression in the lung after single and fractionated doses of radiation. *International Journal of Radiation Biology*, **75**, 1421–1427.

Jhappan, C., *et al*. (1997). DNA-PKcs: a T-cell tumour suppressor encoded at the mouse scid locus. *Nature Genetics*, **17**, 483–486.

Jhappan, C., *et al*. (2000). The p53 response to DNA damage *in vivo* is independent of DNA-dependent protein kinase. *Molecular and Cellular Biology*, **20**, 4075–4083.

Jin, S. and Weaver, D. T. (1997). Double-strand break repair by Ku70 requires heterodimerization with Ku80 and DNA binding functions. *EMBO Journal*, **16**, 6874–6885.

Jin, S., *et al*. (2000). The GADD45 inhibition of Cdc2 kinase correlates with GADD45-mediated growth suppression. *Journal of Biological Chemistry*, **275**, 16602–16608.

Jongmans, W., *et al*. (1996). The role of Ataxia telangiectasia and the DNA-dependent protein kinase in the p53-mediated cellular response to ionising radiation. *Oncogene*, **13**, 1133–1138.

Kasibhatla, S., *et al*. (1998). DNA damaging agents induce expression of Fas ligand and subsequent apoptosis in T lymphocytes via the activation of NF-κB and AP-1. *Molecular Cell*, **1**, 543–551.

Kishi, K., *et al*. (2000). Preferential enhancement of tumor radioresponse by a cyclooxygenase-2 inhibitor. *Cancer Research*, **60**, 1326–1331.

Lavin, M. F. and Khanna, K. K. (1999). ATM: the protein encoded by the gene mutated in the radiosensitive syndrome ataxia-telangiectasia. *International Journal of Radiation Biology*, **75**, 1201–1214.

Malaise, E. P., *et al*. (1986). Distribution of radiation sensitivities for human tumor cells of specific histological types: comparison of *in vitro* to *in vivo* data. *International Journal of Radiation Oncology and Biological Physics*, **12**, 617–624.

McBride, W. H. and Dougherty, G. J. (1995). Radiotherapy for genes that cause cancer. *Nature Medicine*, **1**, 1215–1217.

Milas, L., *et al*. (2000). *In vivo* enhancement of tumor radioresponse by C225 antiepidermal growth factor receptor antibody. *Clinical Cancer Research*, **6**, 701–708.

Moll, U., *et al*. (1999). DNA-PK, the DNA-activated protein kinase, is differentially expressed in normal and malignant human tissues. *Oncogene*, **18**, 3114–3126.

Pajonk, F., *et al*. (1999). Inhibition of NF-κB, clonogenicity, and radiosensitivity of human cancer cells [see comments]. *Journal of the National Cancer Institute*, **91**, 1956–1960.

Potten, C. S., *et al*. (1994). Characterization of radiation-induced apoptosis in the small intestine and its biological implications. *International Journal of Radiation Biology*, **65**, 71–78.

Raju, U., *et al*. (1998). IkappaBalpha degradation is not a requirement for the X-ray-induced activation of nuclear factor kappaB in normal rat astrocytes and human brain tumour cells. *International Journal of Radiation Biology*, **74**, 617–624.

Reardon, D. B., *et al*. (1999). Dominant negative EGFR-CD533 and inhibition of MAPK modify JNK1 activation and enhance radiation toxicity of human mammary carcinoma cells. *Oncogene*, **18**, 4756–4766.

Schmidt-Ullrich, R. K., *et al*. (2000). Signal transduction and cellular radiation responses. *Radiation Research*, **153**, 245–257.

Shackelford, R. E., *et al*. (1999). Cell cycle control, checkpoint mechanisms, and genotoxic stress. *Environmental Health Perspectives*, **107** (Suppl. 1(2)), 5–24.

Shen, S. X., *et al*. (1998). A targeted disruption of the murine *Brca1* gene causes gamma-irradiation hypersensitivity and genetic instability. *Oncogene*, **17**, 3115–3124.

Takao, N., *et al*. (2000). Protective roles for ATM in cellular response to oxidative stress. *FEBS Letters*, **472**, 133–136.

Tepper, A. D., *et al*. (1999). Ordering of ceramide formation, caspase activation, and mitochondrial changes during CD95- and DNA damage-induced apoptosis. *Journal of Clinical Investigation*, **103**, 971–978; Erratum, **103**, 1363.

Todd, D. G., *et al*. (1999). Ionizing radiation stimulates existing signal transduction pathways involving the activation of epidermal growth factor receptor and ERBB-3, and changes of intracellular calcium in A431 human squamous carcinoma cells. *Journal of Receptor and Signal Transduction Research*, **19**, 885–908.

Wang, S., *et al*. (2000). The catalytic subunit of DNA-dependent protein kinase selectively regulates p53-dependent apoptosis but not cell-cycle arrest. *Proceedings of the National Academy of Sciences of the USA*, **97**, 1584–1588.

Ward, J. F. (1988). DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation and repairability. *Progressing Nucleic Acids Research and Molecular Biology*, **35**, 96–128.

Whang, Y. E., *et al*. (2000). c-Abl is required for development and optimal cell proliferation in the context of p53 deficiency. *Proceedings of the National Academy of Sciences of the USA*, **97**, 5486–5491.

Withers, H. R. and McBride, W. H. (1998). Biologic basis of radiation therapy. In: *Principles and Practice of Radiation Oncology*, 79–118. Perez, C. A. and Brady, L. W. (eds). (Lippincott-Raven, Philadelphia).

You, M., *et al*. (1997). ch-IAP1, a member of the inhibitor-of-apoptosis protein family, is a mediator of the antiapoptotic activity of the v-Rel oncoprotein. *Molecular and Cellular Biology*, **17**, 7328–7341.

FURTHER READING

Hall, E. J. (2000). *Radiobiology for the Radiologist*, 5th edn. (Lippincott, Williams and Wilkins, Philadelphia).

Chapter 87

Antibodies and Recombinant Cytokines

Mario Sznol
Vion Pharmaceuticals, New Haven, CT, USA

Thomas Davis
National Cancer Institute, National Institutes of Health, Rockville, MD, USA

CONTENTS

- Introduction
- Cytokines
- Antibodies
- Conclusion

INTRODUCTION

An increasingly sophisticated understanding of cancer biology and immunology, advances in molecular biology and the improved capacity for production of recombinant proteins have made it possible to employ cytokines and antibodies for the treatment of cancer. Several products including the cytokines interferon-alpha and interleukin-2 and the monoclonal antibodies rituximab (antibody to the lymphoma surface antigen CD20) and trastuzumab (antibody to the cell surface receptor Her2/neu) have found a place in the standard treatment of certain human malignancies. In addition, a large number of antibodies and cytokines have undergone clinical evaluation or are currently in clinical trials, and several will likely prove to have benefit and become part of the standard practice of oncology.

CYTOKINES

Biological Basis for Application of Cytokines in Cancer Treatment

Cytokines are naturally produced proteins that interact with specific cell surface receptors, which then generate intracellular signals that produce transcription factor binding to DNA. The resulting increases or decreases in expression of certain genes modify cell function, proliferation or survival. The many and diverse biological effects of most cytokines *in vivo* have made it difficult to define and understand fully the mechanisms underlying their anticancer activity. For example, administration of a single cytokine will induce the production of a cascade of other cytokines that can originate from a variety of cells in the body, including the tumour cells. Cytokines are also able to induce their own production or production of soluble receptors that can enhance or limit their biological effects. Depending on the distribution of receptors, an individual cytokine or secondary induced mediators can have effects directly on tumour cells, immune effector cells, the tumour vasculature and/or the tumour stroma. The biological and antitumour effects of a cytokine may be highly dose and schedule dependent, and may vary with the specific histological type of tumour and stage of growth, the site of metastases and the underlying condition of the host. Some biological effects of cytokines have the capacity to promote tumour growth.

Perhaps the most important biological property of cytokines that impacts the growth of a tumour is the recruitment and activation of immune effector cells. In animal models, the antitumour activity of many of the cytokines can be abrogated by depletion of one or more specific lymphocyte cell subsets. Much attention has been focused on nonantigen-specific natural killer (NK) lymphocytes, since cytokine activation of these cells can enhance their cytotoxicity against tumour cells and production of secondary cytokines, and therefore increase their potential for mediating antitumour activity *in vivo*. However, the majority of animal models involving cytokine administration indicate that antitumour effects require the activation, at some point, of CD4+ and/or CD8+ T lymphocytes that specifically recognize a tumour antigen. Once within the tumour, lymphocytes are capable of eliminating tumour cells directly through cell–cell interactions, or by producing additional cytokines, some of which attract other immune effector cells such as monocytes and macrophages to the tumour microenvironment.

In the modulation of antigen-specific immune responses, cytokines may have pleiotropic effects, e.g. by activating or inducing the maturation of professional antigen-presenting cells such as dendritic cells, by promoting the proliferation, survival and cytokine production

of various T cell subsets or by directly increasing tumour cell expression of antigen and antigen presentation molecules. The role of cytokine administration in inducing production of tumour-specific antibody responses has not been fully explored.

Cytokines in the Treatment of Cancer

Cytokines used to treat cancer have been delivered by systemic and local administration, or have been added to *ex vivo* cultures for activation and expansion of immune effector or antigen-presenting cells. The genes encoding the production of a cytokine can be also transferred to a tumour or tumour stromal cells *in vitro* or *in vivo* using various delivery vectors. Cytokine production by gene-modified cells, which is used primarily to increase the immunogenicity of autologous or allogeneic cancer cell vaccines, will be covered in the chapters on gene therapy and cancer vaccines.

Only three recombinant cytokines (interleukin 2 (IL-2), interferon-alpha, tumour necrosis factor (TNF)) have reached the level of general acceptance for systemic or regional treatment of cancer, as determined by regulatory agency approval in one or more countries. Various other cytokines have been tested in clinical trials but were found not to have sufficient activity for clinical development to be continued. In the now substantial clinical experience with various cytokines, several obstacles to rational and efficient clinical evaluation have become apparent and are common to most cytokines, including the following: incomplete understanding of mechanism of antitumour activity in animal models and, consequently, a lack of a surrogate marker for antitumour activity that can guide the choice of an optimal dose and schedule for patients; inability to define tumour characteristics that would predict for responsiveness to cytokine antitumour effects; inability to administer doses equivalent to the effective doses in animal models because of excessive toxicity; and uncertainty as to whether the cytokines must demonstrate activity in advanced disease before evaluation in the adjuvant (micrometastatic) settings, where they are most likely to show a beneficial effect.

The clinical activity of antitumour cytokines is summarized in **Table 1**.

Interleukin 2

With the possible exception of interferon-alpha, perhaps no other antitumour cytokine has received more attention in basic studies, animal models and clinical trials than IL-2.The initial experiments to treat cancer in animals and patients were based on *in vitro* observations that IL-2 could activate and expand lymphocytes with cytotoxic effects against tumour cell lines. The mouse models demonstrated that IL-2 could be effective in treating small, moderately to strongly immunogenic tumours, although optimal antitumour activity required the administration of near-maximum tolerated doses. The effective IL-2 dose in some animal models could be reduced, or overall antitumour effects increased, if IL-2 was administered together with adoptive transfer of lymphokine-activated killer cells (LAK) or tumour-infiltrating lymphocytes (TIL). LAK are generated by short-term (days) *in vitro* culture of natural killer cells with high concentrations of IL-2. TIL are composed of CD8+, CD4+ lymphocytes that expand over periods of weeks when tumours are resected and placed in *in vitro* culture with IL-2. In contrast to the IL-2/LAK combination, prior treatment with cyclophosphamide was required in order to observe the anti-tumour activity of IL-2 combined with TIL in some tumour-bearing animals. The exact role of cyclophosphamide in these models has not been determined with certainty, but has been attributed to depletion of a suppressor T cell population, or creation of ‘space’ *in vivo* for the ‘take’ of adoptively transferred cells, or partial tumour bulk reduction.

Although direct and indirect evidence in most animal models points to activation or expansion of tumour antigen-specific CD8+ lymphocytes as a necessary requirement for IL-2’s antitumour activity, the exact mechanisms by which this occurs are unclear. IL-2 may be activating and expanding dormant or anergic pre-existing tumour antigen-specific T cells, similar to effects observed *in vitro* when tumours are cultured in IL-2. IL-2, either alone or through induction of secondary cytokines, could also be activating antigen-presenting cells that lead to primary T cell immune responses directed towards tumour antigens. In addition, secondarily induced cytokines such as TNF and interferon-gamma may be necessary to increase expression of antigen and self-major histocompatibility complex (MHC) molecules by tumour cells, thus allowing their recognition by T cells. The role of IL-2 *in vivo* may be substantially more complex, since IL-2 knockout mice have the phenotype of lymphocyte accumulation and autoimmune disease, suggesting perhaps that high-dose IL-2 could also be eliminating suppressive immune cells that inhibit the anti-tumour immune response. In different human tumours, different biological effects may be more important for anti-tumour response, e.g. induction of secondary cytokines or activation of natural killer cells, rather than activation of antigen-specific cytotoxic T cells (CTL). Furthermore, some biological effects such as activation of NK cells may be more optimal for combinations, for example, with monoclonal antibodies, since NK cells have receptors for the Fc portion of antibodies and mediate antibody-dependent cellular cytotoxicity.

Incomplete understanding of tumour–host biology, and its modulation by IL-2 relevant to an antitumour immune response, has made rational application of IL-2 to the

Table 1 Clinical activity of antitumour cytokines

Cytokine	Major common toxicities (including toxicity at highest doses)	Clinical antitumour activity[a]	Comments[a]
Interleukin 1	Constitutional symptoms, fever, chills, nausea, vomiting, diarrhoea, acute hypertension, hypotension, dyspnea, confusion, creatinine elevation, abdominal pain, leucocytosis	Low response rate noted in nonvisceral metastases of metastatic melanoma	Not fully evaluated, no longer in clinical trials
Interleukin 2	High dose: constitutional symptoms, fever, chills, nausea, vomiting, diarrhoea, hypotension, cardiac arrythmias, vascular leak, weight gain/oedema, pulmonary insufficiency/hypoxia, creatinine elevations, decreased urine output, liver function test abnormalities (bilirubin, transaminases), confusion/coma/altered mentation, metabolic abnormalities, skin erythema, thyroid dysfunction	High dose produces 15–20% ORR in MM and mRCC, 5–10% durable CR; combination with interferon-alpha and chemotherapy increases ORR in MM; objective responses reported in several other malignancies, but no consistent activity; in trials to increase relapse-free survival following chemotherapy-induced CR in haematological malignancies	Approved for treatment of MM and mRCC
Interleukin 4	High dose: constitutional symptoms, fever, nausea, diarrhoea, nasal congestion, erythema, liver function test elevations, increased creatinine, oliguria, dyspnea, cardiac arrythmias and ischaemia, gastrointestinal bleeding, fluid retention/vascular leak syndrome, disorientation	Low response rates noted in NHL, rare response in MM and mRCC trials	Primary use now *in vitro* to generate dendritic cells for cancer vaccine trials
Interleukin 6	Constitutional symptoms, fever, nausea, vomiting, liver transaminase and bilirubin levels, hypotension, atrial fibrillation, neurological toxicities (motor and mentation), anaemia, neutropenia, thrombocytosis	Low response rate noted in MM and mRCC	No longer in clinical trials
Interleukin 12	High dose: constitutional symptoms, fever, chills, headache, lymphopenia, thrombocytopenia, neutropenia, stomatitis, liver transaminases and biliribin, autoimmune phenomema, renal insufficiency, confusion/tremors, gastrointestinal bleeding	In early studies, high rates of ORR in CTCL and KS; low response rates observed in MM and mRCC	Remains in clinical trials, studies in lymphoma, myeloma, breast cancer, and combinations with other biologics

(Continued)

Table 1 (Continued)

Cytokine	Major common toxicities (including toxicity at highest doses)	Clinical antitumour activity[a]	Comments[a]
Interferon-alpha	Acute: constitutional symptoms, transient leucopenia, anorexia, nausea and vomiting, reversible transaminase elevations Chronic: hypothyroidism or other autoimmune phenomena, fatigue, depression, other neurological complaints including difficulty in concentrating	High haematological and low (~30%) major cytogenetic remission rate in chronic phase CML, prolongs survival; ~50% complete and partial remission rate in HCL; 15% ORR in MM and mRCC; increases ORR in MM in combination with chemotherapy and IL-2; prolongs PFS in Stage 2b/3 melanoma; prolongs PFS in combination with chemotherapy in poor prognosis indolent lymphoma; has activity in multiple myeloma and KS	
Interferon-gamma	Constitutional symptoms, fever, chills, nausea, vomiting, hypotension, cardiac arrythmias, hepatic and renal dysfunction, granulocytopenia, coagulopathy	Minimal activity noted, although low response rates in MM and mRCC, responses reported with intraperitoneal or regional administration (ovarian carcinoma)	Approved for noncancer indications
Tumour necrosis factor	Acute hypertension, hypotension, liver function test elevations, thrombocytopenia, transient leucopenia, fever, chills, nausea, vomiting, diarrhoea and constitutional symptoms	Minimal activity noted with systemic administration; high (>80%) ORR in limb perfusion studies of MM and sarcoma when used with melphalan and hyperthermia, also high ORR in liver metastases in liver perfusion study	
GM-CSF	At clinically used doses: rare and mild constitutional symptoms and fever	Low responses when tested alone in mRCC, report of increased PFS in fully resected MM patients compared with historical control, responses noted in neuroblastoma in combination with anti-GD2 antibody	Phase 3 trials in stage 4 NED melanoma, with anti-GD2 in neuroblastoma; used primarily as adjuvant in cancer vaccine studies, or *in vitro* with IL-4 to generate dendritic cells for cancer vaccines
FLT3-ligand	Minimal systemic effects at clinically used doses	In phase 2 clinical trials	

[a]ORR, objective response rate, complete plus partial tumour regression; CR, complete response; MM, metastatic melanoma; mRCC, metastatic renal cell carcinoma; HCL, hairy cell leukaemia; CML, chronic myeloid leukaemia; CTCL, cutaneous T cell lymphoma; KS, Kaposi sarcoma; NHL, non-Hodgkin lymphoma; PFS, progression-free survival; NED, no evidence of disease.

treatment of cancer more difficult. This has led to a large number of IL-2 regimens tested in patients, with substantial variations of dose and schedule in an attempt to maximize certain biological effects felt to be associated with antitumour activity. In addition, a very large number of studies have examined the antitumour activity of IL-2 combined with various types of adoptively transferred effector cells, other cytokines, monoclonal antibodies and chemotherapeutic agents.

Clinical Experience

The vast clinical experience with IL-2 alone and in combination with standard and investigational agents and with adoptive immunotherapy still leaves many questions unanswered as to its optimal application to the treatment of cancer. A major limitation in systemic administration of IL-2 is toxicity, which is dependent on dose and duration of exposure. For any given dose, administration by continuous infusion over 24 h is more toxic than administration by short intravenous bolus or subcutaneous injection. At the maximum tolerated doses of any given schedule, IL-2 affects nearly every organ system and produces systemic changes reminiscent of sepsis. Despite the development of severe adverse effects, toxicities are generally completely reversible once IL-2 is discontinued. Proper selection of patients with well-preserved organ function and good performance status has minimized the overall morbidity and mortality from high-dose IL-2 regimens (Kammula *et al*., 1998).

Consistent anti-tumour activity of IL-2 in clinical trials has only been demonstrated for metastatic melanoma and renal cell carcinoma. Although few randomized studies have been conducted, the clinical activity of IL-2 appears to be dose related. Very low-dose chronic administration regimens have not been associated with objective responses in metastatic disease, except in rare circumstances. Nevertheless, the low-dose regimens produce serum concentrations of IL-2 that are sufficient to activate high-affinity IL-2 receptors and substantially expand a subset of natural killer cells (Caligiuri *et al*., 1990; Soiffer *et al*., 1996). The latter underscores the complexity of IL-2 antitumour mechanisms *in vivo*, and indicates that the additional biological effects of higher doses are probably necessary for activity in patients with advanced or metastatic cancer. Higher dose regimens of IL-2 that are associated with modest toxicity and usually administered in an outpatient setting have had variable but low response rates in metastatic renal cell carcinoma with few durable complete responses. A broad examination of the data suggests that the regimens with modest toxicity may be relatively inactive in metastatic melanoma.

IL-2 regimens that require inpatient administration because of their more severe toxicity appear consistently to produce low-level response rates in metastatic melanoma and mRCC in the region of 15%, but with a fraction of responding patients achieving long-term durable responses (Fyfe *et al*., 1995; Atkins *et al*., 1999). The standard regimen involves the administration of two courses of 600 000–720 000 IU kg^{-1} by intravenous infusion every 8 hours for up to 14 doses separated by a 9–12-day rest, and repeated every 8–12 weeks. The administration of IL-2 is unique in that individual doses may be skipped for acute and rapidly reversible severe toxicity, and patients continue to receive doses until they have reached their individual tolerance. Long-term follow up for separate cohorts in excess of 250 patients with metastatic RCC and melanoma demonstrate that between 5 and 10% of all patients have durable relapse-free complete responses lasting years. Equitoxic regimens administering IL-2 by continuous infusion are also active, although direct comparisons with the bolus regimen have generally not been made, and the impression from the combined data is that the response rates and number of durable responses may be lower for continuous infusion regimens.

A formal evaluation of the activity of IL-2 in other malignancies has not been completed, primarily because of the toxicity of the high dose and some of the modest dose regimens. Nevertheless, objective responses have been observed in patients with metastatic colon cancer, ovarian cancer, breast cancer, lung cancer, head and neck cancer, lymphoma, acute myeloid leukaemia and several other malignancies treated with IL-2 either systemically or by regional administration. Randomized trials are ongoing to assess the activity of IL-2 following high-dose chemotherapy and stem-cell support for relapsed lymphoma and AML.

None of the many measured immunological parameters in IL-2 trials have been consistently associated with objective antitumour responses, although some small trials have found correlations between certain variables such as maximum lymphocyte counts following treatment and response. The best predictor for IL-2 response in renal cell carcinoma and metastatic melanoma has been excellent performance status and lack of symptoms. For patients with melanoma, response rates are highest in patients with disease limited to cutaneous and/or cutaneous sites.

IL-2 and Adoptive Therapy

Many of the initial clinical trials of IL-2 included the adoptive transfer of lymphocytes. Patients received a 3–5-day course of IL-2, which produced a marked lymphocytosis approximately 24 h later. The lymphocytes were harvested from the patient and placed in culture with high concentrations of IL-2 for 3–4 days, during which the lymphocytes were activated and developed broad cytotoxicity against tumour cell lines. The LAK were transferred back to the patient together with IL-2. After many trials including some randomized studies, LAK did not appear to contribute substantially to the antitumour activity of high-dose IL-2 (Rosenberg *et al*., 1993).

Subsequent studies involved the adoptive transfer of TIL with IL-2. TIL were generated by placing a surgically

excised tumour in *in vitro* culture with IL-2 for several weeks, during which the infiltrating lymphocytes would expand and tumour cells were eliminated. In some animal models, the resulting lymphocyte population is CD8+, oligoclonal and specifically recognizes and kills the tumour from which they are grown. Similar experiments conducted with human tumours have demonstrated consistent expansion of tumour antigen-specific CD8+ TIL only in patients with metastatic melanoma. Clinical studies conducted with IL-2 and TIL suggest that in some patients with metastatic melanoma the addition of TIL to IL-2 may have benefit, as demonstrated by objective responses in patients who had previously progressed on IL-2 alone. Studies conducted by the Surgery Branch, National Cancer Institute, also demonstrated overall higher response rates with IL-2 and TIL, although the process required to make TIL may also have selected for patients more likely to respond to IL-2 alone. The few studies of TIL conducted in patients with other malignancies, particularly renal cell carcinoma, have not demonstrated superior activity over that expected with IL-2 alone. Clinical studies of IL-2 combined with adoptive transfer of cloned and expanded tumour antigen-specific CTL are now ongoing in patients with metastatic melanoma.

IL-2 and Other Cytokines, Antibodies, Vaccines

In animal models, the addition of systemic IL-2 improves the antitumour activity of various other biologics including multiple different cytokines (interferon-alpha, interferon-gamma, TNF, IL-12), antibodies and cancer vaccines. Toxicity and activity can be highly dose and schedule dependent and can vary between animal models. Various regimens of IL-2 in combination with interferon-alpha have been examined extensively, including some randomized trials in metastatic melanoma and renal cell carcinoma, but these studies have not demonstrated a meaningful improvement in activity over the single agents. Combinations of IL-2 with TNF and IL-2 with interferon-gamma did not progress past phase 1 and early phase 2 trials. Previous trials of murine antibodies raised to human cancer antigens and combined with IL-2 were not promising, but these studies have been reinitiated with humanized monoclonals that have longer serum half-lives and do not raise human antimouse antibody responses (HAMA), thus permitting repeated administration of the antibody. Phase 1 clinical trials of IL-2 and IL-12 were recently initiated on the basis of marked antitumour activity observed in animal models. The combination was also fairly toxic in the animal models and required a specific schedule of pulse IL-2 followed by more chronic IL-12 administration to achieve improved levels of antitumour activity safely.

IL-2 and Chemotherapy

Cytotoxic agents have been combined with IL-2 in several clinical settings. The chemotherapy, often cyclophosphamide at low dose, is given prior to IL-2 alone or IL-2 with cellular adoptive transfer, as a means to reduce *in vivo* immunosuppressive elements, or cause tumour bulk reduction, or 'create space' for 'take' of adoptively transferred cells. Despite many trials, evidence for the benefit of adding cyclophosphamide to IL-2 is equivocal.

In several studies, IL-2 and interferon-alpha have been given either sequentially or concurrently with a combination chemotherapy regimen for melanoma. These studies have almost invariably produced higher objective response rates than chemotherapy alone and, indeed, the higher response rates are being confirmed in randomized trials (Legha *et al*., 1998; Rosenberg *et al*., 1999). The higher response rates, however, translate into small or no survival advantage for patients with metastatic disease. Studies in the high-risk surgical adjuvant setting are ongoing. The mechanisms responsible for the higher response rates with the bio-chemotherapy are not yet defined. Although activation of a melanoma-specific immune response is possible, there is some evidence to suggest that IL-2 induced cytokines together with interferon-alpha may have direct effects on tumours cells that sensitize to the cytotoxic effects of the chemotherapy.

Tumour Necrosis Factor

TNF is another complex cytokine with multiple biological activities *in vivo*. Much has been learned regarding the intracellular signalling pathways for the TNF receptor and other members of the TNF receptor family. Signalling through the receptor activates both a cell death pathway that involves cleavage of caspases and a cell survival pathway that involves activation of the Nf-kappaB transcription factor.

Administration of TNF to mice bearing subcutaneous implanted tumours produces marked central haemorrhagic necrosis and granulocyte infiltration within the tumours, through mechanisms that are still not fully understood today. These effects are observed independent of the sensitivity of the tumour cell lines to TNF *in vitro* and independent of the *in vivo* immunogenicity of the tumour. Furthermore, induction of central haemorrhagic necrosis, presumably through effects on tumour vasculature and perhaps involving cell-mediated events, is observed in both nude and fully immune-competent mice. TNF also has direct antiproliferative or cytotoxic effects against some tumour cell lines, and can sensitize some tumour cells lines to the effects of chemotherapy. The role that TNF's many other biological effects play in its antitumour activity, e.g. effects on different subsets of lymphocytes, granulocytes, endothelial cells, monocytes/macrophages and dendritic cells, is unclear.

Many clinical trials were conducted with single- and multiple-dose and continuous infusion schedules of recombinant TNF in the 1980s, with little observed antitumour activity. The maximum tolerated doses in patients were far less than doses tolerated by mice if normalized

to weight or body surface area. Major clinical toxicities included hypotension, liver function test elevations, thrombocytopenia, transient leucopenia, fever, chills and constitutional symptoms. Investigators subsequently explored the administration of TNF into isolated limb vascular perfusion circuits, for patients with sarcoma or melanoma localized to an extremity. The isolated perfusions allowed delivery of high TNF concentrations to the affected area while minimizing systemic exposure. The response rate was low in the few patients receiving TNF alone by isolated limb perfusion, but the combination of melphalan and TNF, under conditions of hyperthermia and together with a dose of systemic interferon-gamma, produced very high complete and overall tumour response rates in perfused limbs (Eggermont *et al*., 1996; Fraker *et al*., 1996). Some subsequent studies have eliminated the interferon-gamma without an apparent effect on response rate. Although randomized studies to determine the role of TNF have not been completed, on the basis of historical data and the high response rates in the single-arm trials, isolated limb perfusion with TNF, melphalan and hyperthermia is now approved in Europe.

Early clinical trials combining systemically administered TNF with IL-2 or interferon-gamma did not yield sufficient activity to pursue further development. Similarly, some clinical trials combining systemically administered TNF with chemotherapy were performed, but none provided compelling evidence for continued clinical trials. TNF is being explored currently for *in vitro* maturation of dendritic cells used in cancer vaccine studies.

Interferon-alpha

As with other cytokines, many different mechanisms probably account for the antitumour activity of interferon-alpha in patients, and perhaps different mechanisms are active depending on the patient, the disease and stage. Interferon-alpha directly inhibits the growth of some tumour cell lines *in vitro*. Interferon-alpha also enhances immune-mediated antitumour responses, by enhancing NK cell activity, by modulating survival of T cells and the function of professional antigen-presenting (dendritic) cells and by increasing expression of MHC molecules on tumour cells, thus enhancing their recognition by T cells. Low-dose chronic interferon-alpha has been shown to inhibit tumour angiogenesis by down-regulating production of bFGF and matrix metalloproteinases by tumour cells.

A large range of doses and various schedules and routes of interferon-alpha administration have been explored in clinical trials. Acute toxicity is mostly manifest as constitutional symptoms, transient leucopenia, anorexia, nausea and vomiting and reversible transaminase elevations, although severe hepatic toxicity has been reported. The acute effects of flu-like symptoms and occasional elevations in transaminase levels often resolve with continued administration. However, chronic administration of interferon-alpha can be associated with hypothyroidism or other autoimmune phenomena, debilitating fatigue, depression and other neurological complaints including difficulty in concentrating.

As a single agent, interferon-alpha has antitumour activity in several malignant diseases including hairy cell leukaemia, chronic myeloid leukaemia, non-Hodgkin lymphoma, multiple myeloma, Kaposi's sarcoma, renal cell carcinoma and metastatic melanoma. In hairy cell leukaemia, administration of interferon-alpha produces a high percentage of partial to complete haematological remissions (approximately 50%), usually occurring gradually over several months (Grever *et al*., 1995). Discontinuation of treatment eventually results in relapse in most patients, although many will be responsive to reinstitution of treatment. In chronic-phase CML, interferon-alpha prolongs overall survival, with the major benefit to the approximately one-third of patients who achieve major cytogenetic responses (Silver *et al*., 1999). Median time to response in CML is 12 months.

Response rates in the other interferon-alpha sensitive diseases are modest and few durable complete remissions are observed. The benefit of interferon-alpha in stage 2b/3 melanoma patients (following resection of deep primaries or involved regional lymph nodes) remains controversial. In randomized trials, a high-dose interferon-alpha regimen appears to produce some improvement in disease-free but not overall survival, based on the combined results of two randomized trials conducted by the Eastern Oncology Cooperative Group. For stage 2a patients, some randomized trials have shown improved DFS for low-dose interferon regimens, but an effect on overall survival has not been demonstrated unequivocally.

Many different combinations of interferon with standard and investigational agents have been tested in the clinic, and new combinations involving interferon with newer cytokines (IL-12) or antiangiogenic agents (thalidomide) are in early clinical trials. In randomized trials, interferon-alpha combined with multiagent chemotherapy for poor-prognosis low-grade lymphoma has demonstrated an improvement in disease-free survival (and, in one study, overall survival), but the results have not been consistent. Other initially promising results with interferon-alpha in combination with chemotherapy have not been unequivocally confirmed in randomized studies.

Other Cytokines

A number of other cytokines have been or will be evaluated in clinical trials. Interleukin 1 (IL-1) was evaluated only in a limited number of trials, and produced toxicities similar to TNF and few antitumour responses. Interferon-gamma underwent extensive testing in humans, alone and in combination with other agents, and although some clinical responses were observed (primarily in melanoma

and renal cell carcinoma), development in cancer has been mostly discontinued. Despite the ability of interferon-gamma to increase antigen and MHC expression in human tumours, it did not appear to improve clinical response rates when used in combination with IL-2, IL-2 and adoptive immunotherapy, or radiolabelled antibodies in single-arm phase 1 and 2 studies.

Much of the initial impetus to develop interleukin 4 (IL-4) in cancer stemmed from the *in vitro* demonstration that addition of IL-4 and IL-2 in cultures of tumour-infiltrating lymphocytes enhanced the generation of antigen-specific CTL, and suppressed the development of non-specific NK/LAK activity. IL-4 was also shown to inhibit directly the growth of some tumour cell lines *in vitro* (although at high concentrations), to enhance LAK activity in certain situations and, more recently, to inhibit angiogenesis. The opposing effects of IL-4 sometimes reported from different experiments could be explained in part by the sequence with which cells were exposed to other cytokines and their activation status prior to IL-4 exposure. Thus, for example, LAK activity was suppressed by culture of lymphocytes concurrently with IL-2 and IL-4, but enhanced when the lymphocytes were exposed to IL-2 followed by IL-4. Various doses/schedules of IL-4 were examined in the clinic, ranging from daily subcutaneous injections to high-dose thrice-daily intravenous injections. Although some clinical responses were observed, particularly in lymphoma, the activity was not sufficient to merit continued clinical development. Several studies of IL-2 in combination with IL-4, again testing various sequences and regimens, did not demonstrate promising activity. Similarly, broader clinical experience using IL-2 in combination with IL-4 *in vitro* did not improve the yield, specificity or antitumour activity of TIL.

While a role for systemically administered IL-4 has not been found, there is substantial interest in the use of IL-4 *in vitro*, in combination with GM-CSF, to generate dendritic cells for cancer vaccine trials (Romani *et al*., 1994). The dendritic cells generated from PBMC cultured in IL-4 and GM-CSF are potent activators of antigen-specific T-cells. Several trials of peptide-pulsed dendritic cells have demonstrated the ability to induce clinical regression of advanced melanoma in a small fraction of patients. Recent work indicates that a vaccine composed of autologous tumour cells fused to allogeneic dendritic cells (generated in IL-4 and GM-CSF) can produce complete and partial regressions in approximately one-third of patients with advanced renal cell carcinoma. Further exploration of dendritic cell cancer vaccines may depend on the continued availability of IL-4, although a related cytokine, interleukin 13 (IL-13), may have similar properties.

Interleukin 6 (IL-6) exhibited single-agent antitumour activity in various murine tumour models and was therefore examined in phase 1 and phase 2 clinical studies. Although a low rate of partial responses was observed in metastatic melanoma and renal cell carcinoma, at the relatively high doses employed IL-6 produced unacceptable cardiac (atrial arrthymias) and neurological toxicities.

Interleukin 12 (IL-12) has several unique features among the cytokines, including its two chains and a prolonged serum half-life of several hours. In murine tumour models, IL-12 showed substantial antitumour activity alone and in combination with other cytokines (Brunda *et al*., 1993). The antitumour activity is abrogated in most models by antibodies against interferon-gamma or depletion of CD8+ T cells, and partially reduced by antibodies against the interferon-gamma induced chemokines IP-10 and Mig (Nastala *et al*., 1994). IP-10 has been shown to inhibit angiogenesis in various assays.

Depending on dose, schedule and animal model, IL-12's antitumour effects *in vivo* may be mediated by different mechanisms. In the stimulation of antigen-specific T cell responses, either as a single agent or as a vaccine adjuvant, single low doses may be sufficient and perhaps optimal. However, in many animal models, chronic administration of maximum tolerated doses appears to have the greatest antitumour activity. The latter regimens appear to inhibit the induction of antigen-specific T cell responses, at least temporarily, and probably cause tumour regression by inducing high levels of interferon-gamma that cause tumour cell death or production of antiangiogenic cytokines within the tumour. Other mechanisms are likely to be involved since direct administration of interferon-gamma does not have similar antitumour activity. The complexity of these cytokine biological effects has made effective translation into the clinic more difficult.

A variety of dose/schedule regimens for IL-12 have been explored in the clinic. An additional level of complexity was added to the clinical development by the finding that an initial dose of IL-12 could decrease the toxicity and biological effects (i.e. induction of interferon-gamma) of subsequent treatment with IL-12. The induction dose of IL-12 affected subsequent IL-12 administration even if given up to 2 weeks earlier. The mechanisms involved are not well understood. In some animal models, the induction dose increases the amount of IL-12 that can be administered subsequently, but does not necessarily improve the therapeutic benefit. A regimen consisting of an i.v. induction dose, followed 2 weeks later by intravenous administration, daily for 5 days, every three weeks, did not have sufficient antitumour activity in metastatic melanoma and renal cell carcinoma, as measured by objective tumour regression, to justify continued clinical evaluation of that schedule.

Attempts to give IL-12 on a continuous basis have been limited by toxicity, but once and twice weekly schedules have been reasonably well tolerated in advanced cancer patients. High rates of objective partial tumour regression have been reported in early trials of twice weekly subcutaneous administration in patients with early-stage CTCL and with Kaposi sarcoma. In a phase 1 trial of twice

weekly i.v. administration, evidence of tumour regression was observed in a small fraction of patients with metastatic renal cell carcinoma, although patients required months of therapy before response was documented (Gollob *et al.*, 2000). Of note, clinical activity was correlated with the ability of lymphocytes to remain responsive to IL-12 with continued administration. The blunted *in vitro* response to IL-12 in some patients could be enhanced by addition of IL-2, forming the basis of a clinical trial combining IL-12 and IL-2.

Based on *in vitro* and *in vivo* data, studies of IL-12 in combination with other agents have also been initiated, including interferon-alpha, IL-2, monoclonal antibodies and vaccines. A variety of mechanisms serve as the basis for the combinations: enhanced intracellular signalling and tumour antiproliferative effects with interferon-alpha, enhanced NK cell production of interferon-gamma in the presence of tumour and a tumour-specific antibody and potentiation of vaccine-induced antitumour immune responses.

Although not commonly thought of in terms of its potential antitumour effects, Granulocyte–monocyte colony-stimulating factor (GM-CSF) has been reported to enhance the production and activation of monocytes/ macrophages and to promote the development of antigen-presenting cells from monocytes (in combination with IL-4). Based on its effects on monocytes, GM-CSF has been combined with several monoclonal antibodies in clinical trials, but response rates in metastatic disease have been low. Nevertheless, clinical responses were observed in combination with an anti-GD2 antibody in some patients with neuroblastoma, and these results have prompted the initiation of a randomized phase 3 trial of the GM-CSF and anti-GD2 in patients that have been cytoreduced with high dose chemotherapy. As a single agent, GM-CSF has been associated with a low response rate in metastatic renal cell carcinoma, and appeared to increase progression-free survival compared with historical controls in melanoma patients with fully excised metastatic disease. The latter observations formed the basis of an ongoing randomized phase 3 trial comparing GM-CSF to observation (placebo) in stage 4 melanoma patients that have had all metastatic disease resected. GM-CSF has also been administered systemically in combination with IL-4 to increase circulating dendritic cells, and its use together with IL-4 for the same purpose *in vitro* was described above. GM-CSF is a commonly used adjuvant in experimental cancer vaccine trials.

Another cytokine, FLT3-ligand (flt3l), was shown to expand circulating and tissue dendritic cells in animal models (Maraskovsky *et al.*, 1996). As a single agent, flt3l had antitumour activity in several murine models. Although the presumed mechanism of antitumour activity is induction of tumour antigen-specific responses by the flt-3l expanded DC, depletion of NK cells in one tumour model partially reduced the antitumour effect of flt3l administration. Phase 1 trials have confirmed the ability of flt3l to expand circulating DC in patients, and phase 2 studies have been initiated. The ability of flt3l to enhance immune responses to tumour antigens in cancer vaccines is currently being explored. Clinical and preclinical studies to date have not reported an increased incidence of autoimmune disease, and the flt3l has in general been very well tolerated by patients. A fusion protein of flt3l attached to GM-CSF is also in clinical trials.

Several other cytokines are likely to enter or have already entered clinical trials, including interleukin-7, interleukin-15 and interleukin-18. Rational clinical development of these cytokines will depend on careful characterization of their biological properties and relevant mechanisms of antitumour activity in animal models.

ANTIBODIES

Biological Basis for Application of Antibodies to Cancer Treatment

Antibodies are large proteins made by specific blood cells called B lymphocytes, or B cells, which can attach to specific targets with great selectivity and strength, and can then recruit innate immune defence mechanisms to eliminate the bound protein, and any cell to which it is attached. Since it is critical that the immune system does not attack 'self', humans do not normally produce antibodies against normal human proteins. It is possible, however, to generate many different antibodies by injecting viral, bacterial or human proteins into other species, such as mice. Following these injections, the animals will produce abundant antibodies in their blood, and this 'immune serum' has been used to treat several human diseases. In the 1970s, Kohler and Milstein first developed technology to isolate the individual B cells from the immunized animals that were injected with selected antigens, and to induce those B cells to produce a single antibody clone ('monoclonal') against selected antigens (Kohler and Milstein, 1975). It was apparent that specific mouse (murine) antibodies could be produced against human antigens, and the potential to target the human immune system against specific antigens and diseases has driven tremendous research efforts in subsequent decades.

Antibodies mediate their anticancer effects through various mechanisms that depend in part on the particular antigen target. Upon recognition and binding to the tumour antigen, an antibody can mediate complement (CDC) or cell dependent (ADCC) lysis of tumour cells. Cell-mediated cytotoxicity is dependent on the presence of granulocytes, natural killer cells or monocyte/macrophages that bear a receptor for the constant region (Fc) portion of the antibody. The ability of an antibody to mediate CDC or ADCC in humans is dependent on its class and isotype.

In recent years, the mechanisms by which antibodies can mediate anti-cancer activity have broadened to include blocking growth factor receptors on tumour cells, binding and sequestering growth factors, inducing an apoptotic intracellular signal by binding to a cell surface molecule, disrupting tumour vasculature or preventing angiogenesis and modulating an anticancer immune response.

Further improvements in antibody-mediated direct antitumour effects can be obtained by recombinant engineering techniques that join the antigen-binding portions of the antibody to various effector molecules. The tumour-targeting capacity of antibodies permits the delivery of molecules that would otherwise be too toxic to administer systemically. Various toxins have been attached to antibodies, and one, an antibody to CD25 carrying a diphtheria molecule, has been approved in some countries for the treatment of cutaneous T cell lymphoma. Some effector molecules do not have a requirement for internalization into the tumour cell, e.g. cytokines, radioactive isotopes or prodrug-converting enzymes. The latter provide a 'broad field' or bystander effect, and bypass the need for each cancer cell to express the target for the antibody, as long as the antigen-negative tumour cells are in relatively close proximity to the target antigen-expressing tumour cells.

Antigen Targets

Identification of an appropriate target antigen is critical to effective anticancer antibody therapy. It is relatively straightforward to identify targets on nonhuman pathogens such as viruses or bacteria. However, in cancer, the targets are generally common to both normal and cancerous cells, so that any antibody against a common antigen has the potential to affect both normal and malignant tissues, leading to toxicity. Therefore, an appropriate antigen should be tumour specific, expressed uniquely on the tumour or at a much higher level in tumour tissues than on normal cells. It should also be readily available on the cell surface to allow for antibody binding, since under normal conditions antibodies are unable to penetrate into cells. A large number of potential antigens that meet these specified criteria have been identified. Furthermore, many more potential target antigens are being identified by techniques that characterize differential gene expression between normal and tumour cells. Some of the tumour or tumour-associated antigens currently targeted in clinical trials are listed in **Table 2**.

Appropriate selection of an antibody or antibody-targeted therapeutic is also dependent on certain characteristics of the antigen and the intended mechanism of antibody cell killing (**Table 3**). Antibodies that utilize the immune system to kill cancer cells require antigens that are maintained on the cell surface. Antibodies attached to certain toxic agents that must enter cells for their effect must recognize antigens that are internalized after antibody binding. An additional consideration includes the selection of targets that mediate some internal reaction within the cell – either initiating a cell death pathway (apoptosis) or altering cellular mechanisms in some way that may make the cell more susceptible to other therapeutic modalities such as chemotherapy or radiation.

Antibody Types

The fact that all antibodies to human targets are made in nonhuman species presents a significant problem. These nonhuman antibodies are recognized as being alien when they are infused into patients, and are rejected by the recipient over a period of about 2 weeks after initial exposure, and almost immediately after subsequent treatments. Further, the effector arms of the immune system do not interact optimally with these alien molecules, so that their ability to recruit the host's immunity to eliminate a bound target may also be limited. During the last two decades, technologies have been developed that allow the conversion of these murine antibodies into chimeric molecules, where half or more of the protein is replaced with equivalent human parts while the antigen recognition sequence, the part of the molecule that binds to a target, remains unchanged. The standard reference to a chimaeric antibody describes a form that contains approximately 60% human protein sequence, while 'humanized' antibodies contain 90–95% human sequence. These formulations appear to have distinct advantages over their murine counterparts. They do not induce immune reactions and are therefore not rejected by the patient's own immune system, and can consequently persist in a recipient's blood for weeks to months. They can also be given repeatedly. Most antibodies currently in the clinic are either chimaeric or humanized.

As technology has developed further, it is now possible to create completely human molecules. The human immunoglobulin genes have been inserted into mice, which then produce human rather than murine antibodies. An alternate technology involves cloning immunoglobulin genes from human B cells and expressing them on the surface of bacteriophage which can be easily used to select for antibodies against specific targets, including human antigens. Antibodies to specific targets can then be selected and manipulated to create a desired molecule with specific antigen-binding properties. These fully human molecules are relatively simple to make, and could significantly decrease the time necessary to produce new therapeutic antibodies. Whether fully human antibodies will offer improved antitumour effects has yet to be determined.

The source species for an antibody is defined in the agent's generic name. Murine antibodies contain the suffix 'momab,' while chimaeric antibodies use 'ximab' and humanized antibodies are called 'zumab.' The suffix 'umab' identifies fully human antibodies.

Table 2 Tumour targets for antibodies currently used in patients

Antigen	Disease	Normal tissue expression	Antibody	Comments
Idiotype (immunoglobulin receptor)	B cell NHL	None	Anti-idiotype antibodies (murine)	Patient-specific agents requiring individualized production
CD20	B cell NHL or CLL	B cells	Rituximab (chimaeric)	Approved
			B1 (murine)	Radiolabelled
CD22	B cell NHL or CLL	B cells	Epratuzumab (humanized)	Bare and RIT
CD19	B cell NHL or CLL	B cells		Fusion protein
CD25 (IL-2 receptor)	B or T cell lymphoma or leukaemia	Activated B and T cells	Daclizumab (humanized)	Approved as immunosuppressive for renal transplantation
CD52	Lymphoma or leukaemia	Leukocytes	CAMPATH-1H (humanized)	Approved for CLL
CD33	Acute leukaemia	Myeloid stem cells	Hum195 (humanized)	Advanced development as bare antibody
			CMA-676 (humanized)	Immunotoxin approved for AML
CD45	Acute leukaemia	Myeloid stem cells	BC8 (murine)	Radiolabelled
HLA-DR	Lymphoma, leukaemia, and potentially solid tumours	Variable	Hu1D10 (humanized)	Early trials for lymphoma
Her2/*neu*	Overexpressed in epithelial malignancies	Skin, GI tract	Trastuzumab (Herceptin®) (humanized)	Approved for breast cancer. In trial for other diseases
EGF receptor	Overexpressed in epithelial malignancies	Skin, GI tract	C225 (chimaeric)	Advanced clinical development
VEGF	Many malignancies	Regenerating/ healing tissues	Bevacizumab (humanized)	Advanced clinical development
EpCAM	Epithelial malignancies	GI tissues	Edracolomab (17-1A) (murine)	Advanced clinical development. Mechanism unclear
Tag72	Epithelial malignancies	GI tissues	CC49 (humanized)	Radiolabelled
$\alpha v \beta 3$ integrin	Tumour vasculature	None	Vitaxin (humanized)	Early clinical development
GD2 ganglioside	Melanoma, neuroblastoma	Peripheral nerve	Chimaeric 14.18	Advanced clinical development

Table 3 Mechanisms of action for antibody antitumour activity

Mechanism	Description
CDC (complement-dependent cytotoxicity)	Recruits innate complement mechanisms to kill bound cells
ADCC (antibody-dependent cellular cytotoxicity)	Recruits cellular killing via monocytes and NK cells
Apoptosis	Induces spontaneous cell death via signalling mechanism induced by binding. Often related to antigen cross-linking
Signalling inhibition	Blocks ligand/receptor interactions by binding either molecule. Can deprive cell of growth stimulus, etc.
Vaccination	Induces immunity against bound antigen
Immune stimulation	Augments immune recognition of bound cell, or nonspecifically stimulates innate antitumour immunity

General Toxicities of Antibody Infusions

The infusion of biological proteins into humans is associated with a unique syndrome characterized by fevers, chills, low blood pressure, shortness of breath and an assortment of other symptoms. This syndrome is often observed with the initial infusion of antibodies and with subsequent infusions of non-humanized antibodies. In general, it can be controlled by common medications such as acetaminophen or antihistamines.

Antibodies that target immunologically active cells can also cause a more dramatic infusional syndrome that presents the same spectrum of side effects, but often in a more severe form. This syndrome is likely due to a 'cytokine storm,' where the cells rapidly release an assortment of biologically active molecules (interleukin-6, interferon-gamma, TNF) in response to antibody binding, and these cytokines then cause a broad spectrum of side effects. Slowing or delaying the infusion, administering standard prophylactic medications or splitting the dose can reduce these complications. When the target cells have been depleted, or their cytokine load is reduced, the infusional syndrome is reduced or eliminated. Thus, subsequent infusions of the antibody induce few to no side effects.

Unlabelled Antibodies

Many different types of monoclonal antibodies have been tested against tumours. Initial trials suggested limited therapeutic effects, prompting the development of antibodies attached to toxic agents that could mediate cell death in target cells. However, with the selection of new antigen targets and the introduction of chimaeric and humanized antibodies, several unmodified monoclonal antibodies have demonstrated the capacity to induce meaningful tumour regressions in patients.

Stevenson's group first suggested that the antibody made and expressed by B cells could actually be an antigen itself (Hough *et al*., 1976). In B cell lymphomas, the antibody made by the malignant B cell (the 'idiotype') is not found on other cells and can serve as a specific target for antibody therapy. Levy's group first established the use of murine anti-idiotype antibodies as treatment for B cell lymphoma in 1981 (Miller *et al*., 1982; Meeker *et al*., 1985). Up to 75% of patients who were treated on initial clinical trials developed significant reductions in their tumours that lasted a median of approximately 9 months. Some patients experienced complete remissions lasting over 10 years. The fact that a different therapeutic monoclonal antibody had to be produced for each patient made this approach impractical for general patient use. Additionally, many patients developed their own immune response against the murine antibodies (human antimurine antibodies (HAMA)), which rapidly neutralized the therapeutic antibodies from the bloodstream, limiting the amount of time that the therapeutic agent could remain in the patient. A small percentage of patients developed tumour recurrences that had lost or mutated the idiotype target in such a way that the therapeutic anti-idiotype antibody no longer recognized the tumour and could not be used again. Effectively, this treatment selected for malignant cells that did not have the antigen and thus escaped treatment. Such antigen loss is still a critical issue in antibody therapy.

Rituximab

CD20 is a cell membrane antigen that is expressed on mature B cells, but not on B cell precursors or other cells. It appears to be a calcium channel that is involved in activation and proliferation of B lymphocytes. It does not internalize with antibody binding, so that bound antibodies are available on the cell surface for immune system interactions. In addition, CD20 is expressed by over 90% of B cell malignancies. These features make it a good target for antibody therapy, and several different monoclonal antibodies have been developed against this antigen.

Rituximab is a chimaeric mouse–human antibody that binds CD20. It can activate both ADCC and CDC, and has been shown to induce apoptosis in CD20 expressing B cell lymphoma cell lines. Infusion into nonhuman primates resulted in profound depletion of circulating B cells with little toxicity. The initial phase 1 clinical trials did not identify any dose-limiting side effects, and the standard clinical dose was defined by practical issues such as infusion duration and the ability to sustain antibody levels in the blood with depletion of normal B cells. Rituximab demonstrated clear activity against indolent lymphomas, which are generally not cured by available therapies and make up about 40% of all lymphomas. In the pivotal studies, 50% of patients experienced significant benefit, demonstrated in part by a delay of a year or more in chemotherapy treatment for patients who had objective responses (McLaughlin *et al*., 1998).

Rituximab-induced tumour shrinkage progresses over several months after completion of the four weekly antibody infusions, possibly owing to the long duration of rituximab levels in the blood that persist for 2–6 months. Rituximab can also be effective in patients who have responded and then progressed after a previous course of rituximab, with about half of patients responding to a second or third course, thus further delaying other treatments and potentially prolonging lives. Based on at least additive anti-tumour activity *in vitro*, clinical trials are evaluating rituximab and chemotherapy combinations in all types of CD20-positive lymphomas, and initial results are very promising, suggesting that adding the antibody to more traditional treatments may significantly increase patient benefit. Definitive randomized trial evaluations are still in progress.

The use of rituximab in other diseases that express CD20 is still under evaluation. Malignant diseases in

which this drug may provide benefit include acute and chronic lymphocytic leukaemia, small lymphocytic lymphoma, Waldenstrom macroglobulinaemia, high-grade non-Hodgkin lymphoma, hairy cell leukaemia and multiple myeloma. Many different autoimmune diseases may also be responsive to rituximab. Augmenting rituximab's effects is a subject of intense research that includes elucidating the details of mechanism of action and optimizing the clinical dose and schedule.

Trastuzumab (Herceptin®)

The antigen HER2/*neu* is a member of a family of epithelial growth factor receptors. Most skin and gastrointestinal (epithelial) tissues express HER2/*neu* at some point in their development, but usually at relatively low levels. When the gene for HER2/*neu* is inserted into cell lines or mice, it can mediate the transformation of cells into a malignant phenotype. HER2/*neu* overexpression is found in a small percentage of many epithelial cancers and presumably is tied to their transformation into malignant cells. Furthermore, HER2/*neu* expression in certain tumours has been correlated with more aggressive disease and worse outcome for the patient. The overexpression of HER2/*neu* is often, though not always, a result of gene amplification that produces multiple copies of the gene in malignant cells. While this antigen is present in many normal cells, the high overexpression in a subset of tumours makes it a reasonably specific tumour target.

Trastuzumab, (Herceptin®) is a humanized monoclonal antibody that specifically targets the HER2/*neu* receptor. Like rituximab, it can mediate CDC and ADCC and can halt the proliferation of HER2/*neu* expressing cell lines *in vitro*, although the latter may involve direct effects on signalling through the receptor. Clinical development of trastuzumab has been focused on HER2-positive metastatic breast cancer, which accounts for approximately 30% of all metastatic breast cancer patients. When used alone, trastuzumab causes tumour shrinkage in 15–20% of patients, but it appears to improve significantly the antitumour effect of concurrently administered chemotherapy and can prolong the survival of patients with advanced or widespread disease (Pegram *et al*., 1998; Baselga *et al*., 1999; Pegram and Slamon, 1999).

Trastuzumab infusions are generally associated with mild infusion reactions, although some reactions are severe and can be fatal. There are relatively few other side effects. However, in clinical studies of trastuzumab combined with chemotherapeutic drugs that are known to decrease heart function (anthracyclines such as doxorubicin), a surprisingly high incidence of cardiac damage was observed. The cause for augmentation of anthracycline-induced cardiac toxicity is not yet known, since normal heart tissue does not express appreciable amounts of HER2/*neu.* It has been proposed that heart muscle which is regenerating after some toxic insult may have increased expression of the HER2/*neu* receptor and could therefore be more susceptible to damage from trastuzumab.

Trastuzumab's role in the treatment of early-stage breast cancer, and other cancers that overexpress HER2/*neu*, has yet to be defined. Complicating the continuing development of the agent is the unsettled controversy regarding the optimal method to identify tumours that overexpress this antigen. (See the chapter *Signalling by Tyrosine Kinases*.)

CAMPATH-1H (Anti-CD52)

The CD52 antigen is expressed on all lymphocytes and some other haematological cells, but not on the majority of normal tissues. An anti-CD52 antibody made by Cambridge Pathology Associates (alemtuzumab or CAMPATH) has been developed for the treatment of lymphoma and leukaemia. This antibody very effectively depletes all lymphocytes, and has shown promise in the treatment of acute and chronic leukaemias and lymphomas. The primary efficacy appears to be in patients with chronic lymphocytic leukaemia, where CAMPATH-1H (the humanized version of the antibody) can induce significant cancer regression in up to 30% of patients. CAMPATH has also been used to suppress immunity, and thus enhance engraftment, in recipients of allogeneic bone marrow transplants. Because of the severe lymphocyte depletion, CAMPATH administration has been associated with significant immune suppression and opportunistic infections, potentially limiting its clinical use. However, anti-infectious prophylaxis may be able to prevent infectious complications. While immunosuppression may still preclude its use in combination with chemotherapy, there is significant potential for use of CAMPATH in combination with other antibodies or cytokines. This antibody has been approved by the US FDA for the treatment of relapsed CLL.

C225 (Anti-epidermal Growth Factor Receptor)

C225 is a humanized antibody that targets the epidermal growth factor receptor. This antigen is present on many solid tumours, but is also expressed on normal skin and gut. C225 appears to have limited antitumour effects when used at clinically tolerable doses, but preliminary experience suggests that the antibody may be able to augment the effects of either radiation or chemotherapy. The most promising results have been reported in pilot studies of C225 used in combination with radiation therapy for patients with cancer of the head and neck, but it may also sensitize tumours to chemotherapy. Indeed, C225 in combination with cisplatin has produced responses in patients who had previously failed treatment with cisplatin-based therapy. More definitive evidence that it provides significant benefit for patients is pending the outcome of randomized clinical trials.

Anti-angiogenic Antibodies

The concept of disrupting angiogenesis as a means to inhibit tumour growth has been validated in numerous animal models. Tumour neovasculature presents several antigens with restricted distribution in normal tissues, and has receptors for growth factors whose effects are also relatively specific to blood vessels. On this basis, several antibodies targeting proteins essential to the angiogenic process have been developed and are being evaluated in clinical trials.

Several antibodies against vascular epithelial growth factor (VEGF), which supports the growth of new blood vessels, or against the VEGF receptor, are currently in development. VEGF is expressed in any tissue that is developing a new blood supply, but is highly expressed by most cancers, and appears to be essential in their survival. VEGF also has complicated interactions with other biological systems. It appears to suppress immune function, including the inhibition of dendritic cell maturation, and could serve dual roles in promoting the development of a blood supply for the tumour and blocking its rejection by the immune system. Antibodies against VEGF or the VEGF receptor have shown substantial antitumour activity in animal models, which is further enhanced when combined with chemotherapy. The initial results from clinical trials of an antibody to VEGF suggest that circulating VEGF can be cleared from the blood, but it is difficult to determine whether all VEGF in the tumour has been inactivated. Some early results suggest that a small percentage of lung, colon and breast cancers may respond to anti-VEGF therapy alone, and small randomized trials indicate that there may be improved effects when this antibody is used with chemotherapy.

An alternative antiangiogenic target is the family of molecules called integrins. Integrins are expressed in blood vessels, and different members of the family can be selectively expressed in different blood vessels. The antibody identified as Vitaxin targets $\alpha v\beta 3$, an integrin that appears to be selectively expressed in new vasculature and tumour vessels. In addition to the usual killing mechanisms of antibodies, blockade of $\alpha v\beta 3$ could directly induce apoptosis in developing neovasculature. Phase 1 trials have been initiated, and further development is planned, but the results of preliminary studies are not yet available. (See the chapter *Angiogenesis*.)

Immune Stimulants and Anti-idiotype Vaccines

Antibodies can serve as nonspecific immune stimulants that result in an effective antitumour immune response. Antibodies can directly stimulate the proliferation of T cells, deliver positive costimulatory signals to T cells in the context of tumour antigen presentation, block negative costimulatory signals, or provide maturation signals for dendritic cells. Various approaches to enhance T cell antitumour immune responses with immune-stimulating antibodies have been attempted both in preclinical models and in the clinic.

The murine T cell stimulating antibody OKT3 (anti-CD3) causes proliferation of T cells *in vitro* and enhances the effect of IL-2. However, in phase 1 clinical trials, OKT3 alone did not produce tumour regressions, and did not appear to enhance the immune effects or antitumour activity of IL-2. Humanized versions of the antibody are now available, and rather than attempting immune stimulation, consideration is being given to its application in CD3-expressing T cell malignancies. Antibodies to CD3 are still in use for *in vitro* expansion of T cells in adoptive immunotherapy trials. Stimulating antibodies to CD28, which transmits a positive T cell costimulatory signal, are also under investigation for *in vitro* expansion of T cells.

Blocking antibodies to CTLA-4 have generated substantial interest, since CTLA-4 transmits a negative costimulatory signal to T cells, and its blockade *in vivo* enhances native and vaccine-induced antitumour immune responses. Preclinical studies of anti-CTLA-4 alone and in combination with a GM-CSF transfected tumour cell vaccine have shown substantial antitumour activity in mice bearing at least weakly immunogenic tumours (Leach *et al*., 1996). The antibody has been humanized and is being evaluated in phase 1 clinical trials in the near future. An alternative approach involves stimulation of the CD40 molecule, which is predicted to enhance function and maturation of dendritic cells. Preclinical models have shown that an anti-CD40 antibody can cure mice with lymphoma, although the contribution of immune stimulation versus direct effects on malignant B cells is unclear (French *et al*., 1999).

Antibodies can form *in vivo* which recognize the antigen-binding site of other antibodies. The antigen-binding portion of the anti-antibody (AB2) can bear a physical resemblance to the antigen target (e.g. a cancer antigen) of the original antibody (AB1). When the AB2 is used as an immunogen in a vaccine, it is capable of inducing antibody responses (AB3) with similar specificity to AB1. The AB2 is referred to as an anti-idiotype, and anti-idiotypes are sometimes able to generate stronger immune responses to the original antigen than would be possible if the original antigen were used as the immunogen in the vaccine. Various anti-idiotype vaccines are in clinical development, including vaccines that target CEA and the ganglioside GD2 and GD3 antigens, each of which is currently in or will enter phase 3 clinical trials. More discussion on anti-idiotype vaccines can be found in the chapter on *Genetic and Cellular Vaccines*.

Antibodies as Therapeutic Targeting Agents

Early evidence that monoclonal antibodies alone had only limited ability to initiate the rejection of tumours led to

the development of antibodies as targeting molecules to deliver toxic payloads. The toxic payload can either be a radioisotope that provides intense local radiation to damage cellular DNA, or a toxin that can kill the cell if it is internalized. Where possible, it is preferable to deliver therapeutic agents that can affect antigen-negative bystander cells. The choice of antigen is also critical, since toxins may require internalization into the cell, whereas the effect of bound radioisotopes may be diminished by cellular processing.

Radiolabelled Antibodies

The delivery of radioisotope-labeled antibodies for both tumour imaging and targeted radiotherapy has been pursued for several decades. The theoretical benefits of this approach include the ability of the radioactivity to kill both the cells bound by the antibody and also bystander tumour cells that might not express the targeted antigen. Important considerations for effective treatment include tumour radiosensitivity, an appropriate surface antigen and the selection of radioisotope. For the last, availability and cost, radioactive half-life and type and energy of the emitted particle (which affects pathlength) are all relevant issues. A comparison of some available radioisotopes is included in **Table 4**.

Various radiolabelled antibodies have been investigated in solid tumours, either as single agents or in combination with cytokines such as interferon-gamma that upregulate expression of the surface antigen target. In general, few objective responses have been observed, perhaps because of the radio-resistance of tumours, and the relatively small doses of radiotherapy that have been delivered to tumours with acceptable toxicity. Furthermore, most trials were conducted with murine antibodies, and re-treatment was not possible owing to the formation of HAMA. To overcome these problems, investigators are developing new antibody constructs that are humanized, have lower molecular masses to enhance penetration into tumours, and maintain relatively short circulating half-lives in order to achieve high tumour to normal tissue ratios.

Radiolabelled antibodies, and in particular antibodies targeting CD20, have shown substantial antitumour activity in non-Hodgkin's lymphoma. Extensive experience has been developed with anti-CD20 antibodies carrying the isotope ^{131}I, which is readily available and comparatively inexpensive. ^{131}I produces primarily gamma irradiation (photons) that can penetrate deeply into tissue, and has a half-life of approximately 8 days. It does localize to the thyroid gland, so that some preventative measures must be taken to protect the thyroid prior to radioiodine infusions. The therapeutic agent [^{131}I]tositumomab (murine anti-CD20) has been well tolerated with little symptomatic toxicity, although uptake in bone marrow produces marrow suppression. When given in doses that do not eliminate or ablate the bone marrow, up to 80% of patients with indolent lymphomas can experience profound tumour shrinkage that lasts for 1 year or more (Kaminski *et al*., 1993; Vose *et al*., 2000). Blood counts are suppressed for up to 12 weeks, but the complications can generally be managed. HAMA develops in a significant number of patients, which prevents repeated therapy.

When [^{131}I]tositumomab is used at doses that ablate the marrow, a stem cell transplantation procedure is necessary, and toxicity to other organs, such as liver and lung, become limiting. The response rate with this intensive dosing may be higher than with the lesser doses (86% reported response rates in small trials) (Liu *et al*., 1998), and response durations of up to 10 years have been reported, but the complicated therapeutic procedure has limited extensive development of high-dose radioimmunotherapy. The possibility of prolonged toxicity, including thyroid damage and marrow dysfunction (myelodysplastic syndrome), is also a concern with this approach.

An alternative therapy incorporates a different anti-CD20 antibody, the murine parent of rituximab (2B8), bound to ^{90}Y. This isotope produces primarily β particles (electrons) that travel with high energy but penetrate only a

Table 4 Radioisotopes used for radioimmunotherapy

Radioisotope	Radiation[a]	Decay half-life	Maximum pathlength (mm)	Comments
^{131}I	β, γ	8 days	2	Inexpensive, images, significant whole-body radiation with localization to thyroid
^{90}Y	β	2.5 days	12	Minimal normal tissue irradiation, poor imaging, localizes to bone
^{67}Cu	β, γ	2.5 days	2	Scarce, images well, localizes to liver but less whole-body irradiation than iodine
^{125}I	Electron capture	60 days	0.02	Does not image, cell kill in immediate proximity only
^{212}Bi	α, β	1 h	0.08	Alpha particles provide potent cell kill in immediate proximity

[a]Radiation particles: α, two protons (equivalent to a helium ion); β, electron; γ, photon.

small distance (1 cm) into the tissues. The latter produces more intensive local irradiation while decreasing the radiation to normal tissues and to the environment. The half-life is 2.5 days, so the radioactivity clears from the system more quickly. ^{90}Y–2B8 (Y2B8) has also been tested extensively in B cell lymphomas with evidence of similar toxicity and equivalent efficacy to [^{131}I]tositumomab. Both agents will be evaluated by the US Food and Drug Administration (FDA) and may be available for general usage in the near future.

Immunotoxins

Antibodies or ligands for cell surface receptors can be coupled chemically to carry certain toxins or engineered as recombinant DNA products that have a biological toxin attached to the targeting aspect of the molecule. In many, but not all, cases the toxin's mechanism of entry into cells is removed and replaced by the antibody or ligand-targeting molecule, thus markedly reducing its potential nonspecific toxicity for normal tissues. Two immunotoxins have been approved for use in the USA.

Two different immunotoxin therapies that target CD33, an antigen expressed on most cases of acute myeloid leukaemia (AML) and on precursors of some white blood cells, are currently in development. One of immunotoxins carries ozogomycin (calicheamicin), an extremely potent analogue of doxorubicin that can permanently damage cellular machinery. Gemtuzumab ozogomycin efficiently kills bound myeloid and leukaemic cells *in vitro*, and when administered systemically causes complete disappearance of detectable leukaemia in approximately 30% of treated AML cases (Bernstein, 2000). Although the responses are not durable, patients in remission could then be eligible to receive high-dose chemotherapy with stem cell support, a potentially curative treatment for this disease. Gemtuzumab ozogomycin has been approved in the US for clinical use in elderly patients with relapsed AML.

Several other immunotoxins have been extensively evaluated. Their use has been limited by nonspecific damage to the vascular system, leading to increased vessel permeability and fluid leakage, which can seriously impair normal organ function. Another limitation to their usage has been the patient's neutralizing immune response against both the murine antibody and the toxins, which are often derived from bacteria or marine organisms and are therefore immunogenic. This immune response effectively limits the dosing of these agents to a single cycle of therapy, whereas most drugs effective against cancer are given repeatedly in order to achieve benefit. As the toxins are improved to eliminate nonspecific toxicity and supportive agents become available to suppress the immune response, immunotoxins may be administered in dose/schedule regimens with potential for increased antitumour activity.

Despite current limitations, several immunotoxins are showing significant efficacy. The DAB389–IL-2 fusion toxin targets the CD25 IL-2 receptor present on several B and T cell lymphomas and leukaemias, and has been approved for treatment of patients with relapsed mycoses fungoides (cutaneous T cell lymphoma). The molecules LMB-2 and BL-22, both developed at the US National Cancer Institute, are antibody-based immunotoxins that target CD25 and CD22 (a B lymphocyte marker), respectively. CD22 is also present on lymphomas and leukaemias. LMB2 and BL-22 have shown a reasonable toxicity profile with distinct efficacy against advanced chronic lymphocytic leukaemia and hairy cell leukaemia (Kreitman *et al*., 2000). Fortunately, the disease-related immune suppression present in diseases such as CLL and HCL limits immunity against the therapeutic molecule so that repeated doses can be given.

Pre-targeting Antibody Therapeutic Approaches

While antibody conjugates circulate in the blood, they cause damage to the normal tissues and organs through which they circulate. Several different approaches are being developed to enhance the rate and specificity of localization of the toxic molecules into the targeted tumours and limit exposure for normal tissues. In the 'pretargeting' technology, an antibody is labelled with one side of a pair of linking molecules, such as biotin and streptavidin, which will function as a specific glue, and infused into the bloodstream. Over approximately 24–48 h the antibody is allowed to localize to the tumour. A separate compound can be infused to clear the antibody remaining in the blood. Subsequently a toxic compound, either a radioisotope or toxin that is attached to the molecule that will bind to its partner on the antibody, is administered systemically. When this compound is infused, it is able to enter the tumour rapidly owing to its small size, and will attach to the antibody that is already bound to the tumour cells. This approach allows for very rapid localization of the toxic agents into the tumour, and will hopefully reduce the toxicity that results from circulating toxins.

An alternate approach relies on toxins that need to be activated in order to kill cells. The antibody can be attached to a nontoxic enzyme or other activating compound and infused. The unbound antibody can be cleared from the circulation. A prodrug, or inactive toxin, is then infused and allowed to circulate throughout the system. Only the prodrug that penetrates into the tumour will be activated by the enzyme on the antibody, which can lead to a high local concentration of toxic drug within the tumour, but negligible levels throughout the normal tissues. However, backleak from the tumour of activated drug and low-level conversion in normal tissues may lead to toxicity. The general approach has been termed antibody-directed prodrug enzyme therapy (ADEPT). Enzymes being investigated in preclinical systems include carboxypeptidase G2 (CPG2), nitroreductase, cytosine deaminase and DT diaphorase. A CPG2-based ADEPT approach to

activate an alkylating agent prodrug has entered phase 1 clinical trials.

Bispecific Antibodies

A bispecific antibody contains two connected antibodies to make a single molecule that targets two different antigens. Thus a single molecule functions as a bridge or glue, bringing together other molecules, structures or whole cells.

Some bispecific antibodies target a tumour antigen and a toxic molecule or its carrier. The bispecific antibody can be infused with little innate toxicity. The secondary toxic molecule can then be infused at low concentrations, perhaps repeatedly, in such a way that it too has minimal toxicity. The homing mechanism will then concentrate the toxin at the tumour to kill tumour cells selectively.

A second approach uses the second binding site of the bispecific antibody to attract cytolytic killing cells from the immune system. If an antibody against a tumour antigen is coupled to an antibody against a molecule on NK cells, or T cells, the immune effector cell can be attracted to the tumour cell and activated to kill the malignant target. Bispecifics have been produced which include anti-CD16 and anti-CD64 (Fc receptors on NK cells) and anti-CD3 (non-specific activation of T-cells). Anti-Her2/neu × anti-CD16 or × anti-CD64 have been evaluated in the clinic. As expected, the binding to and activation of NK cells and other cells bearing CD16 or CD64 produce cytokine release and its attendant toxicity, but in general, the bispecific antibodies can be administered safely. The effectiveness of the approach has been clearly demonstrated in laboratory experiments, and clinical trials are ongoing.

CONCLUSION

Cytokines and antibodies have become an established part of the cancer treatment armamentarium. Several agents are approved by regulatory agencies, and particularly for antibodies, many more are likely to demonstrate clinical efficacy. There are substantial hurdles that remain to develop effectively optimal therapeutic regimens with cytokines, and progress will come only with an increased understanding of tumour–host biology and the mechanisms of antitumour activity in animal models and in patients. Even for antibodies, the mechanisms of antitumour activity in animal models and humans are not well defined, thus complicating the choice of antigen and antibody, and limiting otherwise very rapid progress. Nevertheless, the rapid elucidation of tumour antigen targets and increased sophistication in creating and manipulating antibodies and delivering toxic payloads suggest that these agents will become increasingly important in the treatment of cancer.

REFERENCES

Atkins, M. B., *et al*. (1999). High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: analysis of 270 patients treated between 1985 and 1993. *Journal of Clinical Oncology*, **17**, 2105.

Baselga, J., *et al*. (1999). Phase II study of weekly intravenous trastuzumab (Herceptin) in patients with HER2/neu-overexpressing metastatic breast cancer. *Seminars in Oncology*, **26**, 78–83.

Bernstein, I. D. (2000). Monoclonal antibodies to the myeloid stem cells: therapeutic implications of CMA-676, a humanized anti-CD33 antibody calicheamicin conjugate. *Leukemia*, **14**, 474–475.

Brunda, M. J., *et al*. (1993). Antitumour and antimetastatic activity of interleukin 12 against murine tumours. *Journal of Experimental Medicine*, **178**, 1223–1230.

Caligiuri, M. A., *et al*. (1990). Functional consequences of interleukin 2 receptor expression on resting human lymphocytes. Identification of a novel natural killer cell subset with high affinity receptors. *Journal of Experimental Medicine*, **171**, 1509–1526.

Eggermont, A. M., *et al*. (1996). Isolated limb perfusion with high-dose tumour necrosis factor-alpha in combination with interferon-gamma and melphalan for nonresectable extremity soft tissue sarcomas: a multicenter trial. *Journal of Clinical Oncology*, **14**, 2653–2665.

Fraker, D. L., *et al*. (1996). Treatment of patients with melanoma of the extremity using hyperthermic isolated limb perfusion with melphalan, tumour necrosis factor, and interferon gamma: results of a tumour necrosis factor dose-escalation study. *Journal of Clinical Oncology*, **14**, 479–489.

French, R. R., *et al*. (1999). CD40 antibody evokes a cytotoxic T-cell response that eradicates lymphoma and bypasses T-cell help. *Nature Medicine*, **5**, 548–553.

Fyfe, G., *et al*. (1995). Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. *Journal of Clinical Oncology*, **13**, 688–696.

Gollob, J. A., *et al*. (2000). Phase I trial of twice-weekly intravenous interleukin 12 in patients with metastatic renal cell cancer or malignant melanoma: ability to maintain IFN-gamma induction is associated with clinical response. *Clinical Cancer Research*, **6**, 1678–1692.

Grever, M., *et al*. (1995). Randomized comparison of pentostatin versus interferon alfa-2a in previously untreated patients with hairy cell leukemia: an intergroup study. *Journal of Clinical Oncology*, **13**, 974–982.

Hough, D. W., *et al*. (1976). Anti-idiotype sera raised against surface immunoglobulin of human neoplastic lymphocytes. *Journal of Experimental Medicine*, **144**, 960–969.

Kaminski, M. S., *et al*. (1993). Radioimmunotherapy of B-cell lymphoma with [^{131}I]anti-B1 (anti-CD20) antibody. *New England Journal of Medicine*, **329**, 459–465.

Kammula, U. S., *et al*. (1998). Trends in the safety of high dose bolus interleukin-2 administration in patients with metastatic cancer. *Cancer*, **83**, 797–805.

Kohler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, **256**, 495–497.

Kreitman, R. J., *et al*. (2000). Phase I trial of recombinant immunotoxin anti-Tac(Fv)-PE38 (LMB-2) in patients with hematologic malignancies. *Journal of Clinical Oncology*, **18**, 1622–1636.

Leach, D. R., *et al*. (1996). Enhancement of antitumour immunity by CTLA-4 blockade . *Science*, **271**, 1734–1736.

Legha, S. S., *et al*. (1998). Development of a biochemotherapy regimen with concurrent administration of cisplatin, vinblastine, dacarbazine, interferon alfa, and interleukin-2 for patients with metastatic melanoma. *Journal of Clinical Oncology*, **16**, 1752–1759.

Liu, S. Y., *et al*. (1998). Follow-up of relapsed B-cell lymphoma patients treated with iodine-131-labeled anti-CD20 antibody and autologous stem-cell rescue. *Journal of Clinical Oncology*, **16**, 3270–3278.

Maraskovsky, E., *et al*. (1996). Dramatic increase in the numbers of functionally mature dendritic cells in Flt3 ligand-treated mice: multiple dendritic cell subpopulations identified. *Journal of Experimental Medicine*, **184**, 1953–1962.

McLaughlin, P., *et al*. (1998). Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. *Journal of Clinical Oncology*, **16**, 2825–2833.

Meeker, T. C., *et al*. (1985). A clinical trial of anti-idiotype therapy for B cell malignancy. *Blood*, **65**, 1349–1363.

Miller, R. A., *et al*. (1982). Treatment of B-cell lymphoma with monoclonal anti-idiotype antibody. *New England Journal of Medicine*, **306**, 517–522.

Nastala, C. L., *et al*. (1994). Recombinant IL-12 administration induces tumor regression in association with IFN-gamma production. *Journal of Immunology*, **153**, 1697–1706.

Pegram, M. D. and Slamon, D. J. (1999). Combination therapy with trastuzumab (Herceptin) and cisplatin for chemoresistant metastatic breast cancer: evidence for receptor-enhanced chemosensitivity. *Seminars in Oncology*, **26**, 89–95.

Romani, N., *et al*. (1994). Proliferating dendritic cell progenitors in human blood. *Journal of Experimental Medicine*, **180**, 83–93.

Rosenberg, S. A., *et al*. (1993). Prospective randomized trial of high-dose interleukin-2 alone or in conjunction with lymphokine-activated killer cells for the treatment of patients with advanced cancer. *Journal of the National Cancer Institute*, **85**, 622–632. Erratum. *Journal of the National Cancer Institute*, 1993, **85**, 109.

Rosenberg, S. A., *et al*. (1999). Prospective randomized trial of the treatment of patients with metastatic melanoma using chemotherapy with cisplatin, dacarbazine, and tamoxifen alone or in combination with interleukin-2 and interferon alfa-2b. *Journal of Clinical Oncology*, **17**, 968–975.

Silver, R. T., *et al*. (1999). An evidence-based analysis of the effect of busulfan, hydroxyurea, interferon, and allogeneic bone marrow transplantation in treating the chronic phase of chronic myeloid leukaemia: developed for the American Society of Hematology. *Blood*, **94**, 1517–1536.

Slamon, D. J., *et al*. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *New England Journal of Medicine*, **344**, 783–792.

Soiffer, R. J., *et al*. (1996). Expansion and manipulation of natural killer cells in patients with metastatic cancer by low-dose continuous infusion and intermittent bolus administration of interleukin 2. *Clinical Cancer Research*, **2**, 493–499.

Vose, J. M., *et al*. (2000). Multicenter phase II study of iodine-131 tositumomab for chemotherapy-relapsed/refractory low-grade and transformed low-grade B-cell non-Hodgkin's lymphomas. *Journal of Clinical Oncology*, **18**, 1316–1323.

FURTHER READING

Atkins, M. B. and Mier, J. W. (eds) (1993). *Therapeutic Applications of Interleukin-2*. (Marcel Dekker, New York).

Gollob, J. A. and Atkins, M. B. (1999). Clinical trials of interleukin-12 in oncology. *Current Opinion in Oncologic, Endocrine, and Metabolic Investigational Drugs*, **1**, 260–271.

Green, M. C., *et al*. (2000). Monoclonal antibody therapy for solid tumours. *Cancer Treatment Reviews*, **26**, 269–286.

Levy, R. (2000). A perspective on monoclonal antibody therapy: where we have been and where we are going. *Seminars in Hematology*, **37** (4, Suppl. 7), 43–46.

Maloney, D. G. (2000). Monoclonal antibodies in lymphoid neoplasia: principles for optimal combined therapy. *Seminars in Hematology*, **37** (4, Suppl. 7), 17–26.

Rosenberg, S. A. (ed.) (2000). *Principles and Practice of the Biologic Therapy of Cancer*. (Lippincott Williams & Wilkins, Philadelphia).

Rosenberg, S. A. (2000). Interleukin-2 and the development of immunotherapy for the treatment of patients with cancer. *Cancer Journal, Scientific American*, **6** (Suppl 1), S2–S7.

WEB Sites

Cytokines Online Pathfinder Encyclopaedia. http://www.copewithcytokines.de/cope.cgi/.

Cytokines WEB. http://www.psynix.co.uk/cytweb/.

The International Cytokine Society. http://bioinformatics.weizmann.ac.il/cytokine/.

Chapter 88

Genetic and Cellular Vaccines

Li Yang and David Carbone
Vanderbilt-Ingram Cancer Center, Nashville, TN, USA

CONTENTS

- Introduction
- Cancer Vaccines
- Cellular Vaccines
- Genetic Vaccines

INTRODUCTION

Vaccines have been powerful tools for the prevention of human infectious disease for decades. These vaccines are protein based, containing attenuated or killed forms of microorganisms, or, more recently, defined natural or recombinant protein components of whole organisms. Even successful immunoprevention strategies for infectious diseases have had very limited applications in the therapy of existing disease, however. The development of immunotherapies for the treatment of existing cancer is an even more difficult problem since there are likely to be fewer significant antigenic differences between tumour and normal cells than between microorganisms and human cells. Cancer therapeutic vaccines may well require approaches very different from infectious disease vaccines if they are to be successful. With the limited effectiveness of conventional treatments for most cancers (particularly the solid tumours) there is a great need for novel approaches. It is hoped that increased understanding of the genetic basis of cancer, in addition to continuing advance in our understanding of the immune response on a molecular level, will allow the design of successful immunotherapeutic strategies.

In this chapter, we will survey the use of cell-based and genetic vaccines for the therapy of cancer. In brief, cell-based vaccines utilize intact mammalian cells as the vaccine, and genetic vaccines rely on expression of the target antigen from transferred nucleic acids. These approaches (and their combination) will be discussed.

Cancers are derived from a clone of normal cells that have undergone malignant transformation. During the process of malignant transformation, a normal cell becomes a neoplastic cell and a series of genetic lesions occur that result in the loss of normal growth regulation and the acquisition of the ability to avoid recognition and destruction by the immune system. Malignant transformation is usually accompanied by the expression of protein, lipid or carbohydrate molecules on the tumour cells that are not expressed at all or are present in much lower quantities than on normal cells. Such molecules, called tumour antigens, potentially allow the immune system to distinguish cancer cells from the normal ones (Wang and Rosenberg, 1999). There are likely to be fewer such distinguishing antigens in a cancer cell than in bacteria or viruses, and this may be one reason why clinically evident tumours manage to avoid induction of an effective immune response, and why anticancer vaccine development has been so difficult. In fact, very little evidence of immune responses against tumours can be detected in most patients. Thus, in addition to defining tumour antigens, understanding the reasons for weak or absent host immune responses induced by clinical tumours is an important step in the development of effective cancer immunotherapy. Several of the new approaches described below are aimed at augmenting weak host immune responses to tumour antigens using potent antigen presenting cells or a variety of genetic manipulations.

CANCER VACCINES

Many approaches have been tried to induce or boost immunity against cancer. The major approaches are shown in **Table 1**. Most cancer vaccines attempt to treat tumours that have already demonstrated this ability to avoid immune recognition. Since a primary requirement for cancer vaccines is to overcome the issues that caused the immune response to fail in the first place, all of these approaches use some strategy to overcome this problem in addition to simply providing tumour antigen. These strategies take the form of adjuvants or cytokine modulation in active immunotherapy or provision of preformed immune effectors in passive strategies.

It has become clear that induction and regulation of immune responses is a complex process involving positive

Table 1 Classification of cancer immunotherapy[a]

Classification	Examples
Active immunotherapy	
Nonspecific	Immune adjuvants such as BCG - activate macrophages. Anti-CD3 antibodies - polyclonal Activation of T cells. Cytokines: IL2, TNF, IFNα, IFNγ, GM-CSF, etc. - to enhance one or more components of cellular immune function
Specific	Immunization with tumour vaccine Cellular vaccine: tumour-specific antigen Genetic vaccine: introduction of genes (cytokines and costimulatory molecules) into tumour cells
Passive immunotherapy	
Antibodies	Monoclonal or polyclonal antibodies alone or coupled with toxins or radiolabels or drugs as 'magic bullets'
Cells	Tumour-infiltrating lymphocytes (TIL), lymphokine-activated killer (LAK)
Indirect	Removal or blocking factors, inhibition of growth factors or angiogenic factors

[a]Adapted from DeVita, V. T., Jr, *et al.*, 1997, *Cancer, Principles and Practice of Oncology* (5th edn) (Lippincott-Raven: Philadelphia).

and negative regulatory factors. Improved understanding of these processes could help design improved immunotherapeutics. These response are both innate, or antigen non-specific, and antigen specific. Antigen-specific responses can be against protein, carbohydrate or lipid antigens expressed uniquely or at least primarily on tumour cells. To be recognized by the cellular immune system, protein antigens from the cell surface or within the cell must be broken down into peptides and presented on the cell surface bound to class I and II major histocompatibility complex (MHC) molecules. While antigen presented this way is necessary for specific responses, it is not sufficient, and a large array of 'costimulatory' molecules modulate the stimulation of specific antibodies and activation of the cytotoxic lymphocytes whose job it is to kill cells expressing these foreign antigens. To induce and regulate these effectors efficiently, there are specialized cells in the body called professional antigen presenting cells (including Langerhans cells and dendritic cells) whose job it is to take up antigen, effectively process it and present it to T cells. Once the antigen is effectively presented on these cells, the T cells must then recognize this antigen, activate and proliferate to result in an effective response. Finally, the target cells must themselves present the antigen in a form that is recognizable by the antigen-specific activated T cells. There are therefore many potential mechanisms, both structural and functional, by which this response can fail.

Tumours may escape immune recognition by either avoiding expression of tumour antigens or by interfering with the induction of an effective response to them. In recent years, investigators have found many mechanisms that may contribute to the failure of immune control of tumour growth. First, tumour cells often have decreased expression of MHC molecules on their cell surface (Seliger *et al.*, 1998). MHC molecules are primarily responsible for antigen presentation and cytotoxic T lymphocyte (CTL) induction. Essential components of the antigen-presenting pathway in tumour cells, including the MHC molecules themselves, may be mutated or deleted, resulting in a complete lack of recognition by CTL. Second, naturally arising tumours often express very 'weak' antigens that are not readily targeted by the immune system (Restifo *et al.*, 1993). Third, most tumour cells lack critical costimulatory molecules, such as CD40, CD80 and CD86, or adhesion molecules, such as ICAM-1 and LFA-3 (Costello *et al.*, 1999), which can contribute to activating T cells. Fourth, recent data show that tumour cells produce and secrete many factors such as TGF-β10, PGE_2 and vascular endothelial growth factor (VEGF), which function to inhibit the function of antigen presenting cells and immune effector cells (Costello *et al.*, 1999, p. 266).

If a tumour has completely lost the ability to process and present antigens to CTL, antigen-specific immunotherapy is likely to be unsuccessful. It is only when immune induction fails for other reasons that cancer vaccine approaches may succeed, and there are data showing that the majority of solid tumours fall into this category. Thus, in addition to identification of the most potent antigens, most immunotherapeutic approaches have focused on optimum methods of immune induction to a particular antigen.

Virtually all modern vaccine strategies include both antigenic and immunostimulatory components, aimed at

inducing or augmenting host immune responses to these weak naturally occurring tumour antigens. In this chapter we will discuss genetic and cellular vaccines in tumour immunotherapy. Cellular vaccines rely on intact tumour cells or antigen-presenting cells activated or 'loaded' with tumour-derived antigens. With genetic vaccines, instead of using protein, genetic material encoding for antigens or cytokines is introduced into cells in culture or directly into a patient.

CELLULAR VACCINES

Tumour cell vaccines can consist of disaggregated autologous or allogeneic tumour cells or tumour cell lines. These can be administered with a chemical or biological adjuvant and there is some evidence of efficacy in several retrospectively controlled clinical trials in colon cancer and melanoma (Rosenberg *et al*., 1998; Restifo and Rosenberg, 1999). Large-scale randomized clinical efficacy trials are currently under way using an autologous tumour vaccine for colon cancer and a mixture of allogeneic cell lines for melanoma. The use of allogenic tumour cell lines in melanoma is based on the finding that many of the observed T cell responses in patients target common antigens shared by many tumours. The complexities of preparing autologous tumour cell vaccines can be avoided by using well-characterized cell lines known to express the common antigens.

One strategy to increase the efficacy of tumour cell vaccines is to engineer them genetically to produce immunostimulatory cytokines (Adris *et al*., 2000). Early-phase clinical trials have been completed using intact autologous tumour cells modified to secrete large quantities of various cytokines, especially GM-CSF, to enhance their immunogenicity, and some immune induction has been observed in several tumour systems (Simons and Mikhak, 1998). Coimmunization of autologous tumour cells with autologous fibroblasts or allogeneic cells engineered to produce cytokine has also met with some technical success at induction of immune responses (Elder *et al*., 1996).

In addition to intact tumour cells or gene-modified tumour cells, another approach is to improve the potency of immune induction by using autologous professional antigen-presenting cells (APCs). The most potent of the antigen-presenting cells is the dendritic cell (DC). These are the cells whose role is to induce and modulate natural immune responses, and thus they are a logical choice for attempts at cancer immunotherapy. Antigen-presenting cells in culture can be isolated, purified, activated or loaded with antigen and administered in such a way as to optimize their potency for activating specific T cells. The potential advantage of this approach is the reduced reliance on antigen uptake, processing and presentation by endogenous antigen-presenting cells, thought to be required for approaches that use non-cell-based or tumour cell-based vaccines. Endogenous APCs may be ineffective owing to the effects tumour-derived suppressive factors. Current therapeutic approaches include the isolation and *in vitro* culture of dendritic cell progenitors with activating cytokines followed by their loading with nucleic acid, peptide or protein antigens either for single protein antigens or complex mixtures derived from cell extracts. There has even been some success when APCs have been fused with intact tumour cells to form antigen-presenting chimaeras (Trefzer *et al*., 2000). We will briefly discuss issues of endogenous dendritic cell dysfunction and then their loading with protein antigens here and then discuss the use of nucleic acid antigens under genetic vaccines.

Dendritic Cells in Cancer

APCs play a central role in the induction of antitumour immune responses. Dendritic cells, as the most potent APC, have high levels of expression of MHC class I and II molecules and express a variety of costimulatory and adhesion molecules required for effective immune induction. Autologous dendritic cells can be cultured from cancer patients and used as vehicles to deliver tumour-specific antigens for the immunotherapy of cancer. It has been observed that these *in vitro*-produced DCs can effectively stimulate specific T cell responses (Gilboa *et al*., 1998; McArthur and Mulligan, 1998).

The principle underlying the *ex vivo* growth or manipulation of dendritic cells for reinjection is based on the assumption that tumours fail to induce specific responses due to a failure of effective antigen presentation. *In vitro* expanded and activated dendritic cells may be effective by virtue of their number and the concentration of antigen, but it has also recently been observed that tumours inhibit the host immune system, at least in part, by interfering with the effectiveness of antigen presentation (Seliger *et al*., 1996). The introduction of DC cultured, activated and loaded with antigen in the absence of inhibitory tumour-associated factors may overcome this problem. In recent years, several groups have shown decreased numbers and functional activity of DC in tumour-bearing mice and human cancer patients. Evidence shows that tumours have a major systemic impact on DC differentiation, maturation and function. Tumour-derived soluble factors such as IL-10, TGF-β, PGE2, VEGF and others directly have the ability to inhibit DC function. DC isolated from peripheral blood, lymphoid tissue and tumours from cancer patients and tumour-bearing mice contain large numbers of DC-like cells which express abnormally low levels of MHC class II molecules and low or undetectable levels of costimulatory molecules . They are unable or impaired in their ability to process and present antigen and therefore do not induce effective antigen-specific or antitumour CTL responses. Defective DC function may therefore be a major factor in the defective immune response observed in patients with

cancer. However, functionally competent DCs can be generated from progenitors isolated from the peripheral blood of cancer patients and cultured *in vitro* in the presence of appropriate cytokines. When loaded with the appropriate antigen, these antigen-loaded DCs form a vaccine that may be more potent than simple introduction of the antigen into a tumour-bearing host.

In Vitro-produced Autologous Dendritic Cells or Monocyte-derived APCs (MD-APCs) as a Tumour Antigen Delivery Candidate

DCs pulsed with tumour-specific peptide can be used as a cellular vaccine if a tumour-specific peptide has been identified (Jager *et al*., 2001). Specific examples of peptides currently under investigation include several from melanoma derived from melanoma-associated antigens such as gp100 and tyrosinase, and also MUC1, Her2/neu, Bcr-Abl fusion peptides, mutant p53 and mutant Ras. These will be discussed in more detail below. Induction of peptide-specific responses is regularly observed in these studies, associated with occasional clinical responses, especially in melanoma (Thurner *et al*., 1999), but the absolute clinical impact of these responses is still uncertain. Multiple doses of pulsed DC vaccines or coadministration with systemic cytokines may be required for efficacy, and the optimum conditions for preparing and delivering antigen-loaded DCs are unclear and are currently being addressed.

Peripheral blood mononuclear cells processed outside the patient's body (*ex vivo*) can differentiate into effective antigen-presenting cells. These monocyte-derived APCs (MD-APCs) can then be grown in the presence of crude tumour protein extracts, tumour cell lysates or tumour apoptotic bodies from allogeneic cell lines or autologous tumour. These autologous DCs, or MD-APCs derived from the patients' own monocytes and then pulsed with tumour-specific peptide or fragments of autologous tumours, or of a relevant tumour cell line, comprise an 'autologous cellular vaccine'. Autologous cellular vaccines may have the capability to more effectively present specific tumour antigens than endogenous APCs provided with exogenous tumour antigen (Nouri-Shirazi *et al*., 2000).

A major problem in studies of whole tumour cell vaccines (or any antitumour vaccine) is the difficulty of monitoring relevant specific immunity induction. This is particularly difficult when the specific antigen is unknown, as is the case with whole tumour cell or extract vaccines. The efficiency of immune induction is often monitored by testing specific T-cell immune cytolytic or cytokine release responses against autologous tumour targets or by more crude *in vivo* assays such as DTH. In addition to being very cumbersome to use as 'intermediate markers' of immune response, the presence of contaminating antigens in the antigen preparation (such as bovine serum albumin (BSA) from foetal calf serum or collagenase used to disaggregate tumour cells) limits the specificity of these assays to measure biologically relevant responses. Importantly, none of these assays has been strictly validated as associated with improved clinical outcomes.

One should note that MD-APCs used in autologous cellular vaccine are only one type of DCs useful for clinical trial. Mature DCs obtained from circulating precursor cells or from $CD34^+$ cells can only be pulsed with peptide antigens. They may not phagocytose or process macroscopic cellular debris like MD-APCs. In addition to macrophage-like properties, MD-APCs have also been shown to present exogenous antigens on both MHC-I and MHC-II complexes.

Tumour Antigens — Shared Tumour-associated Antigens and Mutated Proteins as Vaccine Targets

A major challenge in the development of an antigen-specific tumour vaccine is to identify clearly antigens that will elicit an efficient and effective immune response, and be relatively tumour specific to avoid host autoimmune toxicities. A good tumour antigen can be recognized by its selective expression on tumour cells, and ideally also plays a direct role in malignant transformation (so its expression will not be readily lost). This will allow the immune system to be able to distinguish the cancer cells from normal cells, and prevent the tumour from escaping the immune response. Many candidate antigens are being tested preclinically and clinically.

Protein tumour antigens can be targeted with synthetic peptide vaccines. A variety of strategies have been used to enhance the immunological efficacy of peptide-based vaccines, including chemical, bacterial and mammalian cell-based adjuvants. Even so, the use of synthetic peptides as immunogens is complicated by their weak inherent immunogenicity and variable chemical and physical properties. Some peptides are nearly impossible to synthesize for chemical reasons. The use of autologous DC to present peptides as one form of cellular vaccines appears to be the most effective to date (Nestle *et al*., 1998).

Many melanoma antigens have been identified and include MAGE 1 and 3, gp100 tyrosinase and MART-1/Melan-A. Some of these are also widely expressed on other tumour types. MART-1 is widely represented in human melanomas and is recognized by the majority of HLA-A2 restricted cytotoxic T lymphocytes. One of the immunodominant peptides is MT27-35 from the MART-1 118 amino acid protein (13 kDa) and expressed by HLA-A2.1 melanoma cells (Storkus and Zarour, 2000).

Other examples include the HER-2/neu or c-erbB-2 oncogene product, PSA, CEA and MUC1. HER2/neu is overexpressed in various cancers such as breast, ovarian, lung, gastric, etc., and its overexpression was found to be

associated with high metastasis, chemotherapy resistance and poor clinical outcome. CEA and MUC1 are expressed on many adenocarcinomas.

Many somatically altered oncogenes or tumour-suppressor genes have been identified in human cancer. The mutation of these oncogenes or tumour-suppressor genes is unique to the tumour and also results in a function that contributes to the malignant phenotype, making them reasonable candidates as vaccine antigens. Mutant Ras and p53 are two examples of mutant oncogene and tumour-suppressor gene products with the potential to be good tumour antigens (Ciernik *et al*., 1995). Other oncoproteins that may be good tumour antigens include fusion proteins created by chromosomal translocation such as Bcr-Abl, which is well characterized, relatively common and generated by t(9; 22)(q34; q11) chromosomal translocation in certain leukaemias. Other promising lymphoma antigen targets are the somatically rearranged antibody idiotype sequences in B cell lymphomas. Other oncoprotein candidates are not somatically mutated, but are overexpressed in tumours, such as HER-2/neu.

p53 mutation has been known to be associated with many common human cancers, including lung cancer; 50% of all cancers and 60% of non-small cell lung cancers have p53 mutations. The ideal tumour antigen would be the mutant site in p53, since it is absolutely tumour specific. Other somatically mutated antigens such as the B cell idiotypes also may require customized (patient-specific) vaccine strategies. This approach is very cumbersome, but technological advances in molecular biology make such customized approaches possible, and these are also being tested clinically. For mutant p53-based vaccines, once the specific p53 mutation is determined from the patient tumour sample, a custom mutant peptide can be synthesized. The peptide can then be loaded on autologous dendritic cells produced *in vitro*. The p53 mutant peptide-loaded dendritic cells can be then used to immunize patients to induce cytotoxic T cells that specifically recognize and kill tumour cells expressing this mutant protein (DeLeo, 1998).

Mutant Ras can also be used as a target of cellular vaccine (Khlief *et al*., 1999). Ras mutations are present in 15% of cancers and in at least 40% of colon cancers. Unlike p53, a limited number of specific DNA base substitutions have been described. They are typically present at codons 12, 13, 59 and 61. Further, since Ras mutations are essential for the malignant phenotype, they should not, theoretically, be down-regulated by the tumour to prevent immune detection. Since the number of observed mutations in Ras is limited, all of the common mutant sequences can be manufactured ahead of time rather than being custom made for each patient.

The difficulties in defining tumour-specific peptides appropriate for each patient's class I MHC pattern, and also the observation by several investigators that single-epitope peptide-specific responses can occur in the face of rapid tumour progression, have prompted the continued investigation of polyepitope vaccines. These can consist of mixtures of peptides, but there are several approaches that use complex antigen mixtures (protein or RNA) derived directly from autologous tumour cells.

Cellular vaccines can also be combined with conventional protein-adjuvant immunization to treat cancer patients. In one pilot study by researchers at Stanford University, four patients with follicular B cell lymphoma received a series of three or four infusions of antigen-pulsed dendritic cells followed, in each instance, by subcutaneous injections of soluble antigen 2 weeks later. The tumour-specific idiotype protein produced by each tumour was obtained by cell fusion techniques. All patients developed measurable antitumour cellular immune responses. In addition, clinical responses have been measured with one patient experiencing complete tumour regression, a second patient having partial tumour regression and a third patient resolving all evidence of disease as detected by a sensitive tumour-specific molecular analysis (Hsu *et al*., 1996) (**Table 2**).

Clinical trials involving cancer patients with malignant melanoma (Panelli *et al*., 2000) and prostate cancer

Table 2 Summary of a few clinical trials of cellular vaccine against prostate, melanoma and lymphoma

Clinical trial and reference	Vaccine[a]	No. of patients treated and evaluated	Clinical response status[a]			
			CR	PR	NC	P
Prostate cancer phase II (Murphy *et al*., 1999)	DC pulsed with PSMA	37	1	10	8	18
Melanoma phase I (Trefzer *et al*., 2000)	Autologous melanoma fused with allogenic MHC II-bearing cells	16	1	1	5	9
B-Lymphoma phase I (Hsu *et al*., 1996)	DC pulsed with tumour-specific idiotype protein	4	2	1	0	1

[a]Abbreviations: DC, dendritic cells; PSMA, prostate-specific membrane antigen; CR, complete response; PR, partial response; NC, no change; P, progression.

(Murphy *et al*., 1999) are ongoing (**Table 2**) and are being used to evaluate the feasibility and safety of cellular vaccines, in addition to developing methods for quantitating immune responses as intermediate markers for treatment efficiency. There is cautious optimism that cellular vaccines may provide physicians with powerful new treatment options despite the practical drawbacks to this method. Many of the physical problems of protein- or peptide-based vaccines can be avoided by using a different vaccine approach, the genetic vaccine.

GENETIC VACCINES

A genetic vaccine, also known as DNA or polynucleotide immunization, is defined as using genetic material to encode an antigen that induces immunity. The genetic material can consist of whole cell cDNA or RNA, or cloned cDNAs in expression vectors. These are then commonly introduced into patients directly or into antigen-presenting cells in culture, either stably or transiently. Another type of genetic vaccine uses recombinant viruses such as vaccinia or adenovirus engineered to express the antigen gene of interest.

In its most basic form, a polynucleotide vaccine consists of an expression plasmid vector that drives expression of an antigen encoded by an inserted gene. Gene expression is usually directed by a strong viral promoter. It has several theoretical advantages over synthetic or purified protein antigen vaccines. First, transcription from the genetic vaccine is endogenous, resulting in protein trafficking that mimics that of 'natural' antigens. Second, expression can be continuous over long periods of time rather than relying on short-lived proteins. Third, DNA production, purification and characterization are far easier than for the corresponding proteins. Contamination with bioactive material is much less of a problem. In addition, certain protein structures are unstable or insoluble such as highly hydrophobic regions or those with sulfhydryl groups. This is not a problem with genetic vaccines. Lastly, totally synthetic sequences or fusions of multiple unrelated proteins of significant length can be readily manufactured.

Polynucleotide vaccines were first described in 1992 when Tang *et al*. (1992) showed that antibodies could be efficiently induced by genetic immunization. In 1993, Ulmer *et al*. (1993) showed that plasmid DNA encoding influenza nucleoprotein (NP) induced a cytotoxic T lymphocyte (CTL) response and cross-strain protection in mice. In 1996, Ciernick *et al*. (1996a) used genetic vaccination to induce cytotoxic immunity specific for mutant p53 and showed antitumour efficacy. Since then, the direct injection of plasmid DNA encoding antigenic proteins has been used to elicit humoural and cell-mediated immune responses in a variety of preclinical animal models for cancer, viral, bacterial and parasitic disease.

Another approach for genetic vaccination involves the use of dendritic cells transfected with tumour mRNA (Gilboa *et al*., 1998). It has been reported that mRNA-transfected dendritic cells are potent stimulators of T cell immunity *in vitro* and *in vivo* (Nair *et al*., 2000). A potential advantage of using RNA rather than DNA is that no viral promoter element has to be administered *in vivo* and no integration into the host genome is likely. mRNA should also provide an renewable source of tumour antigen, since modern techniques allow amplification from very small amounts of tumour tissue RNA. Because RNA is very prone to hydrolysis by ubiquitous ribonucleases, different strategies have attempted to prevent RNA hydrolysis in order to improve the efficiency of RNA-based vaccines. These include techniques such as condensation with the polycationic peptide protamine. There are also reports describing specific plasmid vectors engineered for *in vitro* transcription of RNA containing a cap structure and a poly(A) tail, which is then flanked at the 5′ end and 3′ end by untranslated regions (Conry *et al*., 1995). Sindbis virus self-replicating RNA vector engineered to target the antigen to the endosomal/lysosomal compartment via fusion to LAMP-1 (lysosome-associated membrane protein 1) has generated the highest E7-specific T cell-mediated immune responses and antitumour effects relative to RNA vaccines containing either wild-type E7 or Sig/E7 (Cheng *et al*., 2001).

Basic Structure of Vectors for Genetic Vaccine

DNA immunization in its simplest form involves the direct introduction of a plasmid DNA encoding an antigenic protein into the cells of the organism. The gene transferred by the plasmid requires the host cellular machinery to be expressed. DNA-based immunization resembles viral infection except these plasmids are non-replicating and usually lack direct pathological consequences to the expressing cell. Most plasmid DNA constructs are made of three major components: (1) a plasmid backbone for an origin of replication and convenient multiple cloning sites; (2) an antibiotic-resistance gene to ensure that only plasmid-containing bacteria will grow in antibiotic culture medium; and (3) a powerful (and sometimes tissue-specific) enhancer/promoter sequence with an mRNA transcript termination/polyadenylation sequence for directing expression in mammalian cells.

Common regulatory elements for DNA vaccines are those known to mediate high levels of gene expression under mammalian cell culture conditions. Examples include the human cytomegalovirus immediate/early promoter (pCMVIE), the rous sarcoma virus (RSV) LTR and the SV40 early promoter used in conjunction with the SV40 or bovine growth hormone 3′-untranslated region (BGH3′UTR) transcript termination/polyadenylation sequences (Lee *et al*., 1997). Most plasmid vectors also

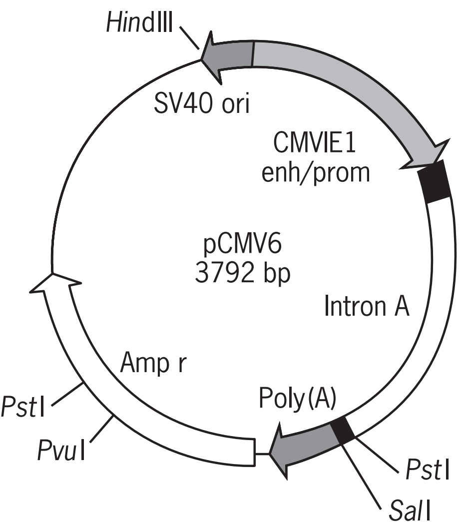


Figure 1 Map of plasmid pCMV6, a typical expression vector containing the CMVIE1 enhancer/promoter with intron A.

contain an intron, because expression of many mammalian genes may be dependent on or may be increased by the inclusion of an intron. To date, one of the most successful plasmids used for eliciting immune responses contains the CMVIE promoter and intron A with the BGH3′UTR (**Figure 1**). It is probable that no single construct will be optimal for all possible genes. One possible concern is the fear of uncontrolled or permanent expression. Owing to the lack of an origin of replication that is functional in eukaryotic cells, these plasmids neither replicate in the mammalian host nor efficiently integrate within the chromosomal DNA of an animal.

Retrovirus and Adenovirus Vectors

Retrovirus-encoded genes depend on the ability of the virus to insert DNA copies of its RNA viral genome into the chromosomes of a host cell. Retroviral vectors take advantage of this ability to insert the gene of interest stably into host genome. The foreign gene replicates along with the host genome. While a definite advantage for gene-replacement strategies, such as the repair of genetic defects, this is only an advantage for genetic immunization when the cells of interest must grow or differentiate after gene transfer, such as when immature DCs are transduced then matured.

Adenovirus vectors are constructed as replication-deficient, E1-deleted versions of these double-stranded DNA respiratory pathogens. The gene of interest (in cDNA form) is usually driven by the CMV enhancer/promoter embedded in the viral backbone. Adenovirus vectors often demonstrate much higher transduction efficiencies than retroviral vectors, but transgene expression is temporary because the adenoviral vectors are replication deficient, and it cannot incorporate into the human genome. It is also able to drive gene expression in non-replicating cells, unlike retroviruses, and is more readily grown *in vitro*.

Genes for Inducing Immunity

Genes that Encode Whole Proteins

Genetic immunization with DNA sequences encoding whole proteins results in antigen production that more closely mimics naturally presented class I epitopes. Briefly, the protein product of genetic vaccine is endogenously synthesized and potentially targeted to the endoplasmic reticulum (ER) like other endogenous proteins. Efficient presentation of peptides derived from endogenously synthesized proteins is thought generally to occur via proteosomal lysis followed by transport into ER. These peptides bind to newly synthesized class I MHC molecules in the ER, which then transport them to the cell surface where they are presented to $CD8^+$ T cells (**Figure 2**). An example of the use of intact protein for immunization is the transduction of dendritic cells by an adenovirus expressing the intact wild-type p53. Normal cells express only very low levels of p53, whereas cells with mutant p53 massively overexpress the mutant protein, thus immunization with the wild-type p53 can generate immunity that can target the many wild-type epitopes in a variety of mutant proteins. Ishida *et al*. showed that this could induce effective antitumour immunity to several different overexpressed mutant proteins (Ishida *et al*., 1999).

There are several studies that have shown that plasmid DNA expression vectors encoding the entire cloned open reading frame of proteins may generate substantial humoural and cellular immunity when directly introduced into living animals (Donnelly *et al*., 1997). In the case of recombinant vaccinia vaccine, the major advantage over purified proteins seems to be extended expression of the protein product from viral genetic material. It is possible, however, that the most potent antitumour effects may occur if responses are directed to some epitopes on the antigen rather than others. Natural immunity or immunity derived from immunization with whole proteins may not induce CTL responses to the desired epitopes.

Minigene Coding for Single Epitopes

When immunity is desired that specifically recognizes one or a few epitopes on a larger molecule, just those portions can be cloned into the expression vector to accomplish this. These ‘minigenes’ or oligonucleotides may only encode a single epitope to obtain the desired defined immune response (An and Whitton, 1997).

A minigene consists of only part of the full open reading frame of a gene. It may encode a single specific T cell epitope that contains a mutant site in an oncoprotein, or simply a single or multiple portions of single or multiple molecules that are known to elicit T cell immunity. An

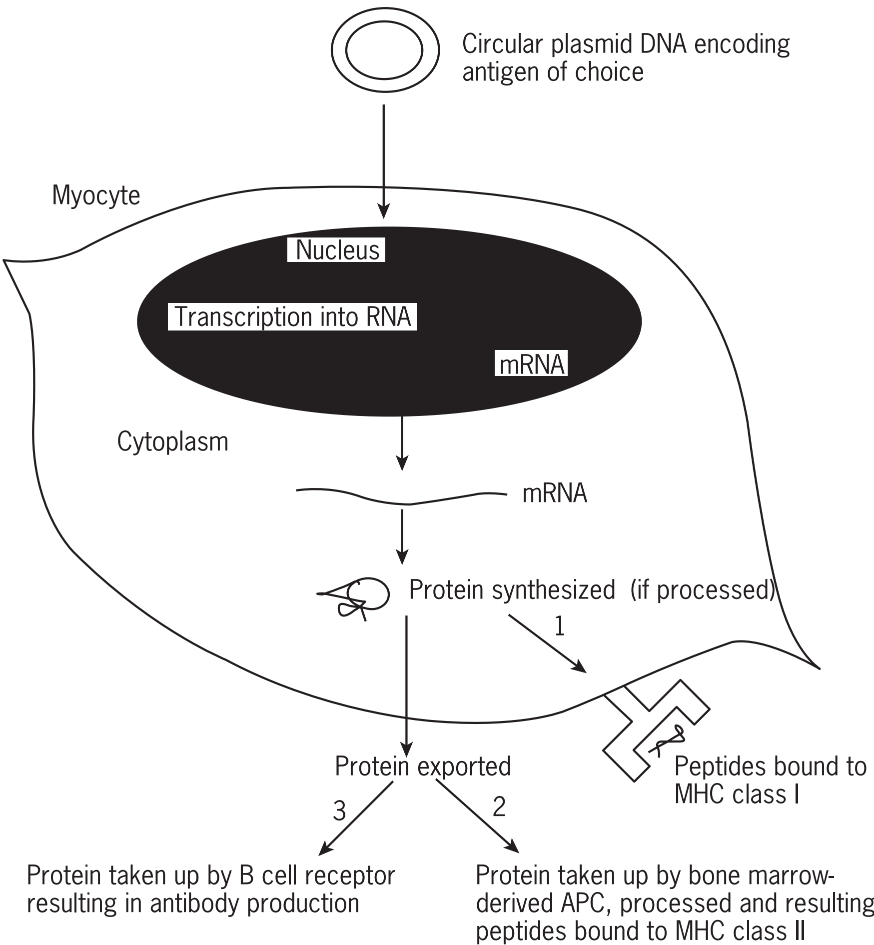


Figure 2 Schematic representation of steps required for immune response to DNA-encoded antigens. Circular plasmid DNA enters the myocyte and is transported into the nucleus. The DNA is transcribed into RNA and the resulting RNA is translocated to the cytoplasm and translated, producing the protein antigen. The antigen is then presented to the immune system by a number of mechanisms. (1) The antigen is processed intracellularly and the resulting peptides are presented on the cell surface bound to major histocompatibility (MHC) class I molecules, where they serve in the selection of cytotoxic T lymphocytes. (2) Protein antigen is transferred to bone marrow-derived antigen-presenting cell (APC), degraded and its peptides are presented on the cell surface bound to MHC II molecules, where they stimulate T-helper cells. (3) The exported protein is taken up by B cells, where it can serve in selection of B cells. (From Tuteja, 1999, *Critical Reviews in Biochemistry and Molecular Biology*, **34**, 1–24.)

example is a mutant *p53* minigene, as reported by Ciernik *et al*. (1996b). The expression vector used in that study only codes for the subunit of p53 from amino acids 128–145. The targeted point mutation is a cysteine to tyrosine at position 135. This sequence was fused in frame to an ER targeting sequence to route the product properly in the APC. Effective and specific T cell immunity was generated after direct injection of this vector. Thus, minigene DNA vaccination is an alternative to peptide-based techniques for eliciting effective immune responses targeted to defined epitopes.

Mutations in p53 are one of the most common somatically acquired genetic abnormalities in human cancer, and most of these mutant genes result in the production of a protein that is altered by a single amino acid substitution. In both p53 and Ras tumour model systems effective induction of mutation-specific CTL may successfully delay or inhibit tumour growth. Because missense mutant protein is present only in tumour cells and not in normal cells, this may make it a good target for tumour-specific cellular immunotherapy.

The use of cloned oligonucleotide epitopes has many theoretical and practical advantages over peptide- or protein-based vaccine. Pure peptides are generally poor immunogens, requiring adjuvants or attachment of lipophilic residues to the peptide for successful induction of CTL responses. One of the major mechanisms of action of adjuvants may indeed be increased antigen persistence. This persistence often causes increased local toxicity as well as efficacy. The production of the epitope from a

DNA template within the cell may result in increased intracellular persistence of the antigen.

Furthermore, it is not clear whether responses to more than a single epitope are necessary for optimal, and therefore biologically significant, immunity in humans. Each situation may require an optimal number and location of epitopes for inducing an efficient immune response.

DNA Vaccine Delivery

One form of cellular vaccine is when genetic material encoding an antigen is introduced into antigen-presenting cells before injection. Genes may be delivered to these cells *in vitro* by any one of the methods (plasmid or viral) mentioned above. These cells may be transduced *in vivo*, however, by a variety of different techniques, including the following.

Intramuscular (i.m.) administration: direct intramuscular and intradermal injection of a DNA vaccine in simple saline solution has been relatively widely used (Hanke *et al*., 1998). There are about 15% myocytes transfected in the vicinity of the injection site in the case of i.m. injection, but these may not be the real APC involved in immune induction. Techniques that are used to transfect cells *in vitro*, such as calcium phosphate precipitates or liposome preparations, do not seem to improve the efficiency of *in vivo* transfer.

Gene gun: devices have been manufactured to fire microscopic gold particles into living tissue. These can penetrate cell membranes and lodge within cells without killing them. DNA can be coated on to the gold particles by mixing the two in the presence of a polycation such as spermidine before firing them into the target tissue, typically the skin. Discharge of the gene gun results in penetration of 0.15 mm in the shaved area of skin. It has shown that this form of DNA injection often transduces cells in both dermis and epidermis (Mahvi *et al*., 1997; Porgador *et al*., 1998).

Other approaches include intravenous, intranasal, intradermal and subcutaneous immunization. A standard multi-tined tatooing device has also been shown to be effective. Since a low efficiency of gene delivery is a major obstacle in DNA vaccination, more studies need to be done to test the best routes for various DNA vaccinations.

Processing and Presentation of Antigens after DNA Vaccination

Generally, protein molecules made or taken up by DC are degraded by proteasomes in the cytosol and short peptide fragments of these proteins are actively transported from the cytosol to the endoplasmic reticulum by a heterodimer protein complex called the transporter of antigenic peptides (TAP). In the ER, these peptides bind to newly synthesized class I MHC molecules, which then carry them out to the cell surface and display them, within their peptide-binding groove, for recognition by $CD8^+$ T cells. The classical chain of events during induction of immune responses includes contact of these DC with antigen, uptake of this antigen and migration of DC to the draining lymph nodes. DCs naturally become activated either by the antigen itself, or by T cells or macrophages through direct contact or release of cytokines. Activated DC (also termed fully mature DC) stimulates antigen-specific T cells. This is followed by T cell migration to the site of antigen entry.

When DNA is directly injected intramuscularly, myocytes at the injection site appear to be the predominant cell type transfected. With gene gunning into the skin, the epidermal Langerhans cell (the cutaneous equivalent of the dendritic cell) may be the important cell type transduced. It is not yet understood how antigen processing and presentation actually occur when cells such as myocytes are expressing the desired antigen. Transfected myocytes may directly mediate antigen presentation, or professional APCs become transfected and serve as APCs, or antigen may be transferred from one to the other, perhaps after death or apoptosis of the myocytes. The mechanisms involved in viral antigen presentation via MHC class I molecules to naive $CD8^+$ T lymphocytes may be much more complex.

Immune Responses after DNA Vaccine Delivery

Immune responses induced by DNA immunization, are very much like those induced by viral infection. These immune responses include antibody production, cytotoxic T cell induction and cytokine secretion.

Humoural Immune Responses

There are a number of studies demonstrating that antibody responses can be induced by DNA vaccination. DNA vaccines encoding influenza NP and HA antigens, HIV envelope protein, bovine herpes virus glycoprotein, carcinoembryonic antigen (CEA), etc., have resulted in the generation of specific antibodies. The duration of antibody responses varies, and can be as long as months to years, or more short-lived. The effectiveness of a DNA vaccine to induce a humoural response, compared with conventional vaccine, appears to be favourable (Conry *et al*., 1996).

Cytotoxic T Cell Responses

Induction of CTL responses can be demonstrated with lymph node or spleen cells from mice that have been injected with plasmid DNA encoding viral antigens. These CTL responses can be detected when these T cells are co-cultured with APC pulsed with purified viral antigens or

infected with the virus. A good example of this is that CD8+ cytotoxic T lymphocytes recognizing an influenza NP peptide can be detected after intramuscular injection with DNA encoding the NP from influenza A virus (Zhou *et al*., 1995). CTL induction has also been shown after gene-gun or intravenous delivery of DNA plasmids encoding specific antigens.

Helper T Cell Responses

Helper T cells are T lymphocytes that secrete cytokines and boost the response of other immune cells such as macrophages, B cells and cytotoxic T lymphocytes. Helper T cells also facilitate the differentiation and development of effector and memory cells. There are two types of T helper cells. Type 1 (Th1) produce cytokines such as IL-2 and interferon-γ. They support the development of cellular immune responses including CTL. Type 2 (Th2) produce cytokines such as IL-4, IL-5, IL-6 and IL-10, which promote B cell activation and antibody production. There are studies showing that DNA intramuscular immunization generates memory T lymphocyte responses manifested by proliferation of antigen-specific T cells and secretion of cytokines during *in vitro* culture of splenocytes from vaccines (Shiver *et al*., 1996; Feltquate *et al*., 1997). High levels of IL-2 and interferon-γ with little or no IL-4 or IL-5 are found in the supernatant of proliferating lymphocytes cultured from immune animals. Therefore, DNA vaccines in these experiments elicited Th1-like cytokine responses (Leclerc *et al*., 1997).

In general, attempts to enhance immune responses against DNA encoded antigens have included varying the vaccination regimen with respect to dose and number of boosts. Coinjection of plasmids encoding cytokines such as GM-CSF, or costimulatory molecules such as B7-1 and B7-2, has also been attempted (Tuteja *et al*., 2000). However, immune resistance can also result from chronic expression of antigen for long periods after DNA injection.

Costimulatory Molecules for Immune Response

Muscle cells express or can be induced to express adhesion molecules, cytokines and MHC class I and II molecules, but they do not seem to express the costimulatory molecules required for efficient antigen presentation. Costimulatory molecules CD80 and CD86 on APC interact with the CD28/CTLA4 molecules on T cells, providing an important second signal in addition to ligation of the T cell receptor to the MHC complex. In a number of experiments, coinjection of CD80 or CD86 expression cassettes along with DNA vaccine resulted in a dramatic increase in CTL response or antibody response (Iwasaki *et al*., 1997).

Summary: the Advantages of DNA Vaccine over Protein- or Peptide-based vaccines

1. DNA can be more easily manipulated and custom designed to meet different needs.
2. In immunocompromised patients, DNA vaccines may be safer than certain live virus vaccines.
3. Antigens from several different proteins can be included on the same plasmid, thus decreasing the number of total vaccinations.
4. DNA is very stable, which is important for the storage, transportation and distribution of DNA vaccine.
5. The equivalent processes for the synthesis, purification and characterization of peptides are at least an order of magnitude more time consuming and expensive.
6. DNA has uniform handling and purification characteristics whereas synthetic peptides differ in a variety of ways, including solubility, optimum diluent and potential for disulfide bonds.
7. The major advantage of genetic vaccine over purified proteins and peptides seems to be extended, endogenous expression of the protein or the peptide product from which results in increased intracellular persistence of the antigen.

Safety Concerns about Genetic Vaccines for Human Use

1. The potential for integration of the plasmid DNA into the human genome and stable, even germ-line, transmission. The integrated antigenic open reading frame may remain silent, but even then could also be mutagenic or disrupt a cellular gene. One particular concern is that the plasmid DNA may cause inactivation of a regulatory gene for cell division or activation of an oncogene. To help combat this potential pitfall, researchers use plasmids without an origin of replication that is functional in eukaryotic cells. Furthermore, the transfected cells, mostly myocytes, sometimes including macrophages and dendritic cells, are for the most part non-dividing. Therefore, the chance of integration and transmission to cellular progeny is limited.
2. Potential for induction of immunological tolerance or unresponsiveness and autoimmunity. There have been various studies on immunological tolerance issues (Mor *et al*., 1997; Czerkinsky *et al*., 1999). Generally, age, species of the subject and the method of administration are key factors related to the ultimate response to DNA immunization. The other important concern is that the destruction of cells expressing foreign genes may lead to the release of cellular constituents theoretically capable of inducing autoimmune responses. However, this event occurs also in the course of viral and bacterial

infection as well as the normal process of tissue remodelling. It appears unlikely that DNA immunization would pose any greater risk in this regard than traditional viral or bacterial vaccine. Such an important issue needs more careful evaluation, however (Overwijk *et al*., 1999)
3. Potential for induction of anti-DNA antibodies, which are known to be associated with various human pathological conditions. It is not yet known whether DNA vaccines induce anti-DNA antibodies in humans, although preclinical animal studies suggest that this is unlikely (Donnelly *et al*., 1997).

The problems and potential of DNA vaccines are outlined in **Table 3**.

Current Human Clinical Trials of DNA Vaccine

There are a few reports detailing the use of DNA vaccines in phase I human clinical trials. One study involved three groups of three subjects (HIV-1 infected but asymptomatic), each of whom were given three doses of 100 μg of plasmid encoding the cDNA for either *nef*, *rev* or *tat* HIV-1 genes intramuscularly in distilled water. Detectable memory cells in all subjects and specific CTLs in eight out of nine subjects that were above preimmunization levels were reported. However, three of these responses were transient. No apparent side effects were detected (Calarota *et al*., 1998). Another group tested the immunogenicity of an HIV env/rev DNA vaccine in HIV-1-seronegative persons. In at least one of multiple assays, the six subjects who received the 300-μg dose had DNA vaccine-induced antigen-specific lymphocyte proliferative responses and antigen-specific production of both interferon-γ and β-chemokine (Ugen *et al*., 1998).

Table 3 The problems and potential of DNA vaccines

Advantages	Concerns and problems
Noninfectious/nonreplicating	Potential for insertion into the genome
Proteins produced in native conformation	Autoimmunity
Physicochemical homogeneity	Immunological tolerance
Stability for storage, transportation and distribution	Induction of anti-DNA antibodies
Easy manipulation/simplified production	
A single vaccine for multiple pathogens	
Extended expression of antigen	

Melanoma is the most immunogenic solid tumour, which makes it an ideal target for immunotherapy. DNA vaccine clinical trials against melanoma are ongoing and a number of genetically modified tumour vaccines (GMTVs), with various immunostimulatory factors, are being tested in phase I/II clinical trials, which include cytokines, tumour antigens (TAs), costimulatory molecules or HLA antigens. Some of these studies have shown enhanced survival of patients developing immune responses to melanoma vaccines (Ollila *et al*., 1998; Nawrocki *et al*., 2000). In addition, a phase III clinical trial of vaccinia melanoma oncolysate (VMO) vaccine showed minimal toxicity and clinical efficacy in patients with melanoma (Wallack *et al*., 1998). A phase I study using direct combination DNA injections for the immunotherapy of metastatic melanoma is ongoing at the University of Colorado Cancer Center. Results of two clinical trials are summarized in **Table 4**.

The release of further human clinical trial data is keenly awaited. Strategies may need to be modified to achieve higher efficiency of DNA vaccination such as enhancing the uptake of DNA, targeting oncoproteins or using a prime-boost strategy (DNA vaccine given in the initial dose followed by recombinant protein or virus).

Table 4 Two clinical trials of DNA vaccine against melanoma

Clinical trial and reference	Vaccine[a]	No. of patients treated and evaluated	Clinical response status[a]			
			CR	PR	NC	P
Phase I/II (Nawrocki *et al*., 2000)	Autologous tumour cells mixed with allogeneic melanoma engineered to produce IL-6 and sIL-6R	41	5	4	13	19
Melanoma Phase I (Soiffer *et al*., 1998)	Autologous melanoma cells engineered to produce GM-CSF	16[b]	3	4	3	5

[a]Abbreviations: sIL-6R, soluble Il-6 receptor; GM-CSF, granulocyte macrophage colony-stimulating factor.
[b]One patient had a scapular metastasis during vaccine preparation.

REFERENCES

Adris, S., *et al*. (2000). Mice vaccination with interleukin 12-transduced colon cancer cells potentiates rejection of syngeneic non-organ-related tumor cells. *Cancer Research*, **60**, 6696–6703.

An, L. L. and Whitton, J. L. (1997). A multivalent minigene vaccine, containing B-cell, cytotoxic T-lymphocyte, and Th epitopes from several microbes, induces appropriate responses *in vivo* and confers protection against more than one pathogen. *Journal of Virology*, **71**, 2292–2302.

Calarota, S., *et al*. (1998). Cellular cytotoxic response induced by DNA vaccination in HIV-1-infected patients. *Lancet*, **351**, 1320–1325.

Cheng, W. F., *et al*. (2001). Enhancement of sindbis virus self-replicating RNA vaccine potency by targeting antigen to endosomal/lysosomal compartments. *Human Gene Therapy*, **12**, 235–252.

Ciernik, I. F., *et al*. (1995). Mutant oncopeptide immunization induces CTL specifically lysing tumor cells endogenously expressing the corresponding intact mutant p53. *Hybridoma*, **14**, 139–142.

Ciernik, I. F., *et al*. (1996a). Human lung cancer cells endogenously expressing mutant p53 process and present the mutant epitope and are lysed by mutant-specific cytotoxic T lymphocytes. *Clinical Cancer Research*, **2**, 877–882.

Ciernik, I. F., *et al*. (1996b). Induction of cytotoxic T lymphocytes and antitumor immunity with DNA vaccines expressing single T cell epitopes. *Journal of Immunology*, **156**, 2369–2375.

Conry, R. M., *et al*. (1995). Characterization of a messenger RNA polynucleotide vaccine vector. *Cancer Research*, **55**, 1397–1400.

Conry, R. M., *et al*. (1996). Polynucleotide-mediated immunization therapy of cancer. *Seminars in Oncology*, **23**, 135–147.

Costello, R. T., *et al*. (1999). Tumor escape from immune surveillance. *Archives of Immunology and Experimental Therapy*, **47**, 83–88.

Czerkinsky, C., *et al*. (1999). Mucosal immunity and tolerance: relevance to vaccine development. *Immunological Reviews*, **170**, 197–222.

DeLeo, A. B. (1998). p53-based immunotherapy of cancer. *Critical Reviews in Immunology*, **18**, 29–35.

Donnelly, J. J., *et al*., (1997). DNA vaccines. *Annuual Review of Immunology*, **15**, 617–648.

Elder, E. M., *et al*. (1996). Successful culture and selection of cytokine gene-modified human dermal fibroblasts for the biologic therapy of patients with cancer. *Human Gene Therapy*, **7**, 479–487.

Feltquate, D. M., *et al*. (1997). Different T helper cell types and antibody isotypes generated by saline and gene gun DNA immunization. *Journal of Immunology*, **158**, 2278–2284.

Gilboa, E., *et al*. (1998). Immunotherapy of cancer with dendritic-cell-based vaccines. *Cancer Immunology and Immunotherapy*, **46**, 82–87.

Hanke, T., *et al*. (1998). Immunogenicities of intravenous and intramuscular administrations of modified vaccinia virus Ankara-based multi-CTL epitope vaccine for human immunodeficiency virus type 1 in mice. *Journal of General Virology*, **79**, 83–90.

Hsu, F. J., *et al*. (1996). Vaccination of patients with B-cell lymphoma using autologous antigen-pulsed dendritic cells. *Nature Medicine*, **2**, 52–58.

Ishida, T., *et al*. (1999). Dendritic cells transduced with wild-type p53 gene elicit potent anti-tumour immune responses. *Clinical and Experimental Immunology*, **117**, 244–251.

Iwasaki, A., *et al*. (1997). Enhanced CTL responses mediated by plasmid DNA immunogens encoding costimulatory molecules and cytokines. *Journal of Immunology*, **158**, 4591–4601.

Jager, D., *et al*. (2001). Vaccination for malignant melanoma: recent developments. *Oncology*, **60**, 1–7.

Khleif, S. N., *et al*. (1999). A phase I vaccine trial with peptides reflecting ras oncogene mutations of solid tumors. *Journal of Immunotherapy*, **22**, 155–165.

Leclerc, C., *et al*. (1997). The preferential induction of a Th1 immune response by DNA-based immunization is mediated by the immunostimulatory effect of plasmid DNA. *Cellular Immunology*, **179**, 97–106.

Lee, A. H., *et al*. (1997). Comparison of various expression plasmids for the induction of immune response by DNA immunization. *Molecular Cells*, **7**, 495–501.

Mahvi, D. M., *et al*. (1997). Phase I/IB study of immunization with autologous tumor cells transfected with the GM-CSF gene by particle-mediated transfer in patients with melanoma or sarcoma. *Human Gene Therapy*, **8**, 875–891.

McArthur, J. G. and Mulligan, R. C. (1998). Induction of protective anti-tumor immunity by gene-modified dendritic cells. *Journal of Immunotherapy*, **21**, 41–47.

Mor, G., *et al*. (1997). Do DNA vaccines induce autoimmune disease? *Human Gene Therapy*, **8**, 293–300.

Murphy, G. P., *et al*. (1999). Phase II prostate cancer vaccine trial: report of a study involving 37 patients with disease recurrence following primary treatment. *Prostate*, **39**, 54–59.

Nair, S. K., *et al*. (2000). Induction of cytotoxic T cell responses and tumor immunity against unrelated tumors using telomerase reverse transcriptase RNA transfected dendritic cells. *Nature Medicine*, **6**, 1011–1017.

Nawrocki, S., *et al*. (2000). Genetically modified tumour vaccines (GMTV) in melanoma clinical trials. *Immunology Letters*, **74**, 81–86.

Nestle, F. O., *et al*. (1998). Vaccination of melanoma patients with peptide- or tumor lysate-pulsed dendritic cells. *Nature Medicine*, **4**, 328–332.

Nouri-Shirazi, M., *et al*. (2000). Dendritic cells capture killed tumor cells and present their antigens to elicit tumor-specific immune responses. *Journal of Immunology*, **165**, 3797–3803.

Ollila, D. W., *et al*. (1998). Overview of melanoma vaccines: active specific immunotherapy for melanoma patients. *Seminars in Surgical Oncology*, **14**, 328–336.

Overwijk, W. W., *et al*. (1999). Vaccination with a recombinant vaccinia virus encoding a 'self' antigen induces autoimmune vitiligo and tumor cell destruction in mice: requirement for CD4(+) T lymphocytes. *Proceedings of the National Academy of Sciences of the USA*, **96**, 2982–2987.

Panelli, M. C., *et al*. (2000). Phase 1 study in patients with metastatic melanoma of immunization with dendritic cells presenting epitopes derived from the melanoma-associated antigens MART-1 and gp100. *Journal of Immunotherapy*, **23**, 487–498.

Porgador, A., *et al*. (1998). Predominant role for directly transfected dendritic cells in antigen presentation to CD8 + T cells after gene gun immunization. *Journal of Experimental Medicine*, **188**, 1075–1082.

Restifo, N. P., *et al*. (1993). Identification of human cancers deficient in antigen processing. *Journal of Experimental Medicine*, **177**, 265–272.

Restifo, N. P. and Rosenberg, S.A. (1999). Developing recombinant and synthetic vaccines for the treatment of melanoma. *Current Opinions in Oncology*, **11**, 50–57.

Rosenberg, S. A., *et al*. (1998). Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nature Medicine*, **4**, 321–327.

Seliger, B., *et al*. (1996). Reduced membrane major histocompatibility complex class I density and stability in a subset of human renal cell carcinomas with low TAP and LMP expression. *Clinical Cancer Research*, **2**, 1427–1433.

Seliger, B., *et al*. (1998). Down-regulation of the MHC class I antigen-processing machinery after oncogenic transformation of murine fibroblasts. *European Journal of Immunology*, **28**, 122–133.

Shiver, J. W., *et al*. (1996). Humoral and cellular immunities elicited by HIV-1 vaccination. *Journal of Pharmaceutical Science*, **85**, 1317–1324.

Simons, J. W. and Mikhak, B. (1998). *Ex-vivo* gene therapy using cytokine-transduced tumor vaccines: molecular and clinical pharmacology. *Seminars in Oncology*, **25**, 661–676.

Soiffer, R., *et al*. (1998). Vaccination with irradiated autologous melanoma cells engineered to secrete human granulocyte-macrophage colony-stimulating factor generates potent antitumor immunity in patients with metastatic melanoma. *Proceedings of the National Academy of Sciences of the USA*, **95**, 13141–13146.

Storkus, W. J. and Zarour, H. M. (2000). Melanoma antigens recognised by CD8 + and CD4 + T cells. *Forum (Genova)*, **10**, 256–270.

Tang, D. C., *et al*. (1992). Genetic immunization is a simple method for eliciting an immune response. *Nature*, **356**, 152–154.

Thurner, B., *et al*. (1999). Vaccination with mage-3A1 peptide-pulsed mature, monocyte-derived dendritic cells expands specific cytotoxic T cells and induces regression of some metastases in advanced stage IV melanoma. *Journal of Experimental Medicine*, **190**, 1669–1678.

Trefzer, U., *et al*. (2000). Hybrid cell vaccination for cancer immune therapy: first clinical trial with metastatic melanoma. *International Journal of Cancer*, **85**, 618–626.

Tuteja, R. (1999). DNA vaccines: a ray of hope. *Critical Reviews in Biochemistry and Molecular Biology*, **34**, 1–24.

Tuteja, R., *et al*. (2000). Augmentation of immune responses to hepatitis E virus ORF2 DNA vaccination by codelivery of cytokine genes. *Viral Immunology*, **13**, 169–178.

Tuting, T., *et al*. (1999). Induction of tumor antigen-specific immunity using plasmid DNA immunization in mice. *Cancer Gene Therapy*, **6**, 73–80.

Ugen, K. E., *et al*. (1998). DNA vaccination with HIV-1 expressing constructs elicits immune responses in humans. *Vaccine*, **16**, 1818–1821.

Ulmer, J. B., *et al*. (1993). Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science*, **259**, 1745–1749.

Wallack, M. K., *et al*. (1998). Surgical adjuvant active specific immunotherapy for patients with stage III melanoma: the final analysis of data from a phase III, randomized, double-blind, multicenter vaccinia melanoma oncolysate trial. *Journal of the American College of Surgeons*, **187**, 69–77.

Wang, R. F. and Rosenberg, S. A. (1999). Human tumor antigens for cancer vaccine development. *Immunological Reviews*, **170**, 85–100.

Zhou, X., *et al*. (1995). Generation of cytotoxic and humoral immune responses by non-replicative recombinant Semliki Forest virus. *Proceedings of the National Academy of Sciences of the USA*, **92**, 3009–3013.

FURTHER READING

Abbas, A. K., *et al*. (2000). *Cellular and Molecular Immunology* (4th edn) (W. B. Saunders: Philadelphia).

DeVita, V. T. Jr, *et al*. (1997). *Cancer, Principles and Practice of Oncology* (5th edn). (Lippincott-Raven: Philadelphia).

Chapter 89

Differentiation Therapy

Reuben Lotan
The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA

CONTENTS

- Introduction and Rationale for Differentiation Therapy
- Preclinical Aspects of Differentiation Therapy
- Clinical Aspects
- Conclusions

INTRODUCTION AND RATIONALE FOR DIFFERENTIATION THERAPY

Cancer morbidity and mortality continue to be major problems worldwide despite improvements in diagnostic, preventive and therapeutic strategies. The improved understanding of genetic, molecular and cellular aspects of carcinogenesis over the last few years has revealed various new targets for intervention. Among these novel approaches is differentiation therapy, the subject of this chapter.

In a developing multicellular organism, cells achieve their specialized set of structural and functional characteristics through a process called differentiation. This process involves concerted changes in gene expression aimed at restricting the expression of a diverse repertory of genes in pluripotential cells to those required to attain the specialized phenotype in unipotent cells in a specific tissue. Cellular differentiation often leads to a mature nondividing cell (terminally differentiated cells). This is the consequence of the modulation of genes that are involved in regulation of cell proliferation, senescence, and programmed cell death (apoptosis).

The development of cancer usually involves aberrations in cellular differentiation; dividing cells that have the potential to differentiate, senesce or undergo apoptosis become immortalized and often fail to complete the differentiation programme owing to a block at a stage preceding terminal differentiation. That this block can be reversed is indicated by reports on tumour cells that undergo 'spontaneous' or induced differentiation *in vitro* and *in vivo* including in patients. Thus, neuroblastoma cells have been shown to differentiate into ganglioneuroma cells, and germ cell tumours (teratocarcinoma) into benign teratomas. In addition, keratin pearls characteristic of terminally differentiated keratinizing squamous epithelial cells are found within squamous cell carcinomas (Pierce *et al.*, 1978).

A series of elegant studies have demonstrated that the malignant phenotype is reversible and can be controlled by physiological signals mediated by soluble factors or cell–cell contacts provided by an appropriate developmental field. In one such experiment, embryonal carcinoma cells injected into mouse blastocysts and implanted into pseudopregnant female mice developed into normal cells in a variety of tissues of the chimaeric offspring. In contrast, the same cells developed into tumours when injected subcutaneously (Illmensee and Mintz, 1976). Similarly, myeloid leukaemia cells injected into embryos *in utero* at 10 days of gestation developed into normal granulocytes in the offspring, neuroblastoma cells have lost their tumorigenic properties when injected into fragments of neurula-stage embryos then transplanted into mice and melanoma cells injected into embryonic skin failed to form tumours. Certain human teratocarcinoma cells were found to differentiate into a variety of somatic cell types after injection into immunodeficient mice.

Additional support for the contention that the malignant phenotype can be reversed has come from *in vitro* studies, which have demonstrated that the aberrant differentiation of tumour cells can be restored at least partially to normal differentiation by interactions with adjacent normal cells (e.g. by epithelial–mesenchymal interactions) or with extracellular matrix. Several cell types have been found to undergo differentiation *in vitro* after treatment with cytokines that regulate normal differentiation or with other compounds, both physiological and pharmacological, that activate various differentiation pathways (**Table 1**).

Because normal terminal differentiation often results in non-proliferating cells that often undergo apoptosis as they complete their normal life span, it was plausible to develop strategies to activate normal pathways of differentiation in premalignant and malignant cells using physiological or pharmacological agents that can bypass the epigenetic and genetic abnormalities that abrogate differentiation. This approach, called differentiation therapy, can be used to

Table 1 Differentiation-inducing agents and their suggested mechanisms of action

Differentiation-inducing agent	Mechanism of action
Short-chain fatty acids (butyrate and derivatives)	Inhibit histone deacetylase, thereby enhancing gene expression, alter the methylation and phosphorylation of nuclear proteins and alter gene transcription
Aromatic fatty acids (phenylbutyrate, phenylacetate)	Inhibition of protein prenylation critical for signal transduction; activation of peroxisome proliferator-activated receptor, a transcription factor related to the steroid nuclear receptor superfamily; hypomethylation of DNA; depletion of circulating glutamine which is required for tumour growth
Trichostain A, SAHA	Inhibit histone deacetylase, thereby modulating gene expression
PPAR ligands	Activate nuclear receptors (PPARs) and modulate gene transcription
Haxamethylenebisacetamide	Modulate cell cycle and differentiation related genes
Vitamin A analogues (retinoids)	Activate nuclear receptors (RARs and RXRs) and modulate gene transcription
Vitamin D_3 analogues	Activate nuclear receptors (VDR) and modulate gene transcription
Cytokines (TNF-α interleukins)	Act via cell surface receptors to induce signal transduction pathways resulting in alterations in gene expression
Interferons	Activate transcription factors (Stats) and thereby alter gene transcription
5-Aza-2′-deoxycytidine	Decreases DNA hypermethylation and thereby restores expression of silenced genes
Diflouromethylornithine	Ornithine decarboxylase inhibition
Staurosporin	Protein kinase C inhibition
8-Choloro-cyclic adenosine monophosphate (8-Cl-cAMP)	Downregulates RI and upregulates the RIIb subunit of protein kinase A
Etoposide nondamaging (ICRF-193)	Topoisomerase II inhibition
Actinomycin D	DNA intercalation
Ultravoilet light; X-irradiation	DNA damage, AP-1 activation
Cytosine arabinoside; cyclophosphamides; aclacinomycin (at subtoxic concentrations)	Inhibition of DNA synthesis

prevent, suppress or reverse the malignant phenotype by inducing differentiation with the associated growth arrest, senescence and apoptosis (**Figures 1–3**). Even if only partially successful, differentiation therapy can convert malignant tumours into benign tumours. In this chapter, we discuss preclinical and clinical data on selected differentiation-inducing agents and their mechanisms of action.

PRECLINICAL ASPECTS OF DIFFERENTIATION THERAPY

A variety of compounds, both natural and synthetic, have been found to induce differentiation *in vitro* in various cell lines, including murine erythroleukaemia cells, human myeloid leukaemia cells and murine and human embryonal carcinoma cells. Some of these agents are presented in **Table 1** and selected ones are described below.

Histone Deacetylase Inhibitors

Histone acetylation status appears to be important in the regulation of differentiation, cell cycle control, hormone responsiveness, senescence and cancer because of its impact on chromatin structure and the control of gene transcription. The histone acetylation status is regulated by the equilibrium of histone acetyltransferase activity and histone deacetylase activity. Histone deacetylases (HDACs) are enzymes that catalyse the removal of acetyl groups from the N-terminal lysine residues of core nucleosomal histones. This action results in restoration of a positive charge to the lysine residues of the histones, which then bind tightly to the phosphate backbone of DNA. The subsequent condensation of the structure of the nucleosome inhibits transcription because transcription factors and their associated cofactors and RNA polymerase do not have access to the gene promoters (Marks *et al*., 2000). Several compounds have been found to inhibit HDACs and this is presumed to relax the structure of chromatin associated with programmed genes allowing them to be expressed. These compounds include butyroids (e.g. butyric acid), hydroxamic acids (e.g. trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA)), cyclic peptides (e.g. trapoxin and apicidin) and benzamides. Most of these agents have been found to induce differentiation and/or apoptosis of transformed cells *in vitro* and some also suppressed tumour growth *in vivo* (Marks *et al*., 2000).

Butyroids

Butyrate, a naturally occurring four-carbon fatty acid found in abundance in milk fat, is also produced by microbial fermentation of dietary fibre, undigested starch

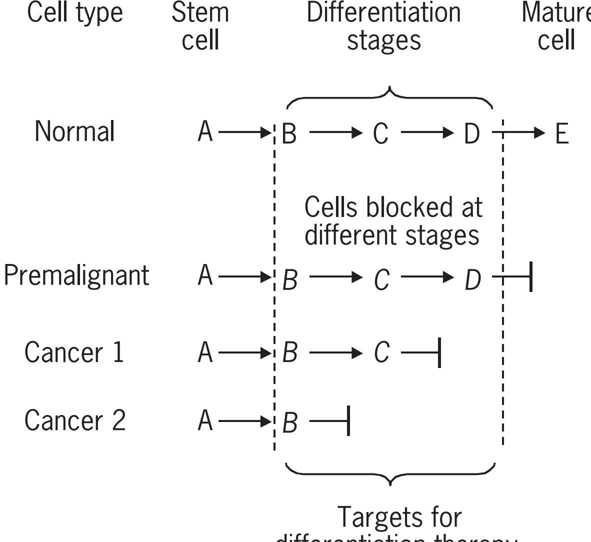


Figure 1 Targets for differentiation therapy. In a normal differentiation pathway, a hypothetical stem cell designated A is required to undergo a series of progressive changes in gene expression and resultant phenotypic changes represented by the letters B, C and D before reaching the terminally differentiated mature cell stage designated E. Cancer development is associated with aberrant differentiation. This can be represented by a block in differentiation that can occur already in premalignant cells. The scheme shows three possible blocks at late stages of the differentiation pathway (e.g. at stage D) and at early stages of the pathway (e.g. at stage B).

and proteins within the colonic lumen of mammals. Butyrate has been shown to inhibit proliferation and induce differentiation and apoptosis in a variety of tumour cell types including erythroleukaemia, embryonal carcinoma, colon carcinoma, pancreatic carcinoma, neuroblastoma, glioblastoma and haematopoietic malignancies (Newmark *et al*., 1994). Butyrate induced the expression of specific differentiation-associated genes characteristic of defined pathways of differentiation when used at concentrations between 0.5 and 10 mmol L^{-1}.

Specific effects of butyrate on gene regulation have been reported, both by transcriptional regulation and post-transcriptional modifications. It has been proposed that butyrate exerts its effects by a noncompetitive inhibition of HDACs and the consequent modulation of gene transcription. Some of the genes modulated in butyrate-treated cells are involved in cell differentiation, growth and apoptosis. Butyrate has been shown to induce cell cycle arrest in G_1. The mechanism of this arrest may be related to the ability of butyrate to increase the transcription of the cyclin-dependent kinase inhibitor $p21^{WAF1}$. In normal and malignant human breast cells, butyrate induced differentiation, decreased expression of cyclin D1, increased expression of $p21^{WAF1}$ and increased hypophosphorylated retinoblastoma protein pRB. Butyrate has also been shown to induce apoptosis in lymphoid cell lines and colon carcinoma cells.

Butyrate, delivered in liposomes to human colon carcinoma tumours growing in immunodeficient mice, was able to suppress tumour growth and enhance differentiation (Otaka *et al*., 1989). This study provided a proof of

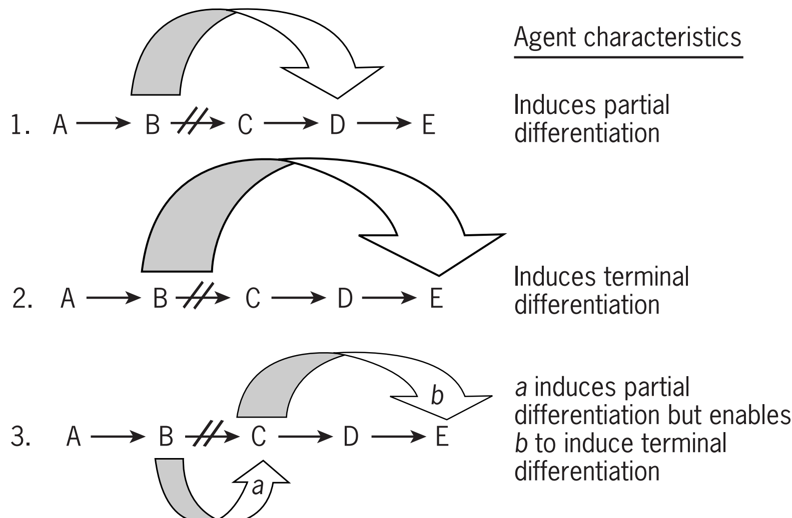
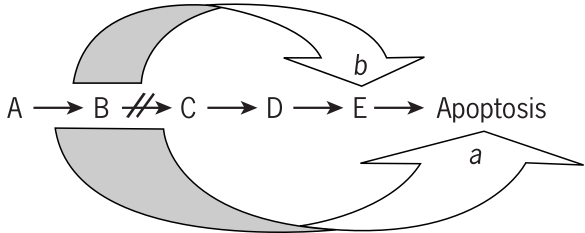


Figure 2 Characteristics of agents that are used to induce differentiation in malignant cells. In scenario 1 an agent can induce partial differentiation of cells blocked at stage B by inducing them to undergo changes to stage D. In scenario 2 the agent can induce cells blocked at stage B to undergo full differentiation to the mature cell E. In scenario 3, agent *a* can induce a partial differentiation (from B to C) but at the same time it renders the resultant cell C capable of responding to agent *b* that can complete the differentiation of cell C to the mature cell E. The latter example can serve as one of the rationales for combination of agents for differentiation therapy. The designation of the letters A–E is the same as in **Figure 1**.



a: Induces apoptosis without inducing differentiation
b: Induces terminal differentiation and the mature cells eventually undergo apoptosis

Figure 3 Inter-relations between differentiation and apoptosis. Because many cell types undergo apoptosis as a normal consequence of terminal differentiation, treatment of cancer cell at stage B with agent *b* induces cell at stage B to undergo differentiation to the mature cell E and eventually to apoptosis. This scenario is contrasted with an agent such as *a*, which can induce apoptosis without inducing differentiation.

the principle that butyrate has a potential for use in differentiation therapy. However, the clinical development of butyrate was limited by the requirement for millimolar concentrations for efficacy combined with its very short metabolic half-life.

To overcome this problem, butyrate derivatives (butyroids) with a longer half-life than butyrate have been synthesized and characterized and some were found to be more potent than butyrate *in vitro* and *in vivo*. Among these derivatives, tributyrin was about fourfold more potent than either butyrate or monobutyrin in inducing monocytic differentiation of HL-60 myeloid leukaemia cells and erythroid differentiation of murine erythroleukaemia cells. Tributyrin has been shown to induce apoptosis in human prostate cancer cell lines irrespective of their androgen receptor expresssion status. Butyramide and monobutyrin also induced differentiation and inhibited DNA synthesis in hepatoma cells and mouse erythroleukaemia cells *in vitro*.

Pivaloyloxymethyl butyrate (AN-9), a prodrug of butyric acid, is a differentiation-inducing agent in a variety of cells. In myelomonocytic cell line (WEHI), AN-9 modulated the expression of the early regulatory genes, c-*myc* and c-*jun*, leading to growth arrest and differentiation. Even a short (4-h) pretreatment *in vitro* resulted in a marked reduction in the tumorigenicity of AN-9-treated cells. Similarly, a short (1-h) AN-9 treatment of Lewis lung carcinoma cells resulted in an almost complete loss of tumorigenicity. In both cell types butyric acid did not affect the tumorigenicity of the cells even when it was used at concentrations 10 times higher than those of AN-9. Thus, some butyroids with longer half-life and greater potency than butyrate might be useful for cytodifferentiation therapy of human malignancies.

Aromatic Fatty Acids

The aromatic fatty acid phenylbutyrate (PB) and its metabolite phenylacetate (PA) are natural derivatives of phenylalanine. When used at millimolar concentrations, which are pharmacologically attainable and nontoxic, these agents have been shown to induce differentiation of a variety of cell types including leukaemia, erythroleukaemia, prostate carcinoma, neuroblastoma, rhabdomyosarcoma, melanoma and neuroectodermal tumours (Samid *et al*., 1997). Although PB and PA are related structurally to butyrate, their mechanisms of action are distinct from butyrate and in some respects also from each other.

PA exhibited a marked activity against a highly tumorigenic human breast cancer cell line both *in vitro* and *in vivo*. *In vitro*, cell growth was inhibited and the anti-apoptotic protein Bcl-2 was down-regulated in conjunction with the appearance of apoptotic cells. *In vivo*, the treated tumours showed growth arrest, Bcl-2 down-regulation and differentiation. PA-induced growth arrest in human breast carcinoma MCF-7 cells was associated with enhanced expression of the cyclin-dependent kinase inhibitor $p21^{Waf1/Cip1}$ and dephosphorylation of pRB. The induction of $p21^{WAF1/CIP1}$ mRNA by PA was independent of the cellular p53 status and appeared to be required for growth inhibition because mouse embryonal fibroblasts from a p21−/− knockout mouse failed to undergo growth arrest by PA.

PA treatment of prostate cancer cell line LNCaP resulted in an increase in prostatic-specific antigen (PSA) mRNA and protein despite inhibition of tumour cell proliferation. A similar increase was noted in rats bearing subcutaneous LNCaP tumour implants that were treated systemically with PA. PA can alter tumour cell lipid metabolism and *ras*-encoded p21 via inhibition of protein isoprenylation. For example, PA inhibited protein isoprenylation and $p21^{ras}$ farnesylation and induced apoptosis in LNCaP prostate cancer cells transfected with the mutated c-Ha-*ras* gene. PA reduced the cellular levels of endogenous farnesyl-PP and inhibited activation of the $p21^{ras}$ downstream target, p42(MAPK)/ERK2. Exogenous farnesyl- and geranylgeranyl-PP reduced the effects of the drug on proliferation and apoptosis in LNCaP(T24-ras) cells.

Although both PB and PA promote differentiation in human prostate cancer cell lines, PB is more active at inhibiting growth and inducing programmed cell death in several human prostate cancer lines than PA at clinically achievable doses. PB caused prostate cancer cells to undergo apoptosis 2–10 times more than PA. Prostate cancer cell lines overexpressing P-glycoprotein or possessing p53 mutations were also sensitive to the effects of PB. *In vivo* experiments demonstrated that PB administered to rats orally delayed growth of an androgen refractory rat prostate cancer subline by 30–45% in a dose-dependent manner. These results demonstrate that PB

induces cytotoxicity via apoptosis in human prostate cancer, in addition to its differentiating properties. The pathway signalling apoptosis in prostate cancer cells exposed to PA included reduction in mitochondrial transmembrane potential, release of cytochrome *c* from the mitochondria to the cytosol, activation by proteolytic processing of caspase-3 and -7 and cleavage of the death substrates poly(ADP–ribose) polymerase and DNA fragmentation factor.

The chimaeric fusion product AML1-ETO, which is formed in acute myeloid leukaemia (AML) by the (8;21) translocation, recruits a transcription repression complex that includes the histone deacetylase 1 enzyme. PB could partially reverse ETO-mediated transcriptional repression, possibly through inhibition of HDAC. PB was also able to induce partial differentiation of the AML1–ETO cell line Kasumi-1. The relief of differentiation repression using PB may prove to be therapeutically beneficial.

Trichostatin A (TSA)

The fungistatic antibiotic TSA, isolated from the culture broth of *Streptomyces platensis*, was found to be a specific inhibitor of the cell cycle of normal rat fibroblasts in the G_1 and G_2 phases. In addition, TSA was found to be a potent inducer of erythroleukaemia cell differentiation at submicromolar concentrations. TSA inhibited the growth or induced differentiation of erythroleukaemia cells, colon carcinoma cells, neuroblastoma cells and pancreatic cancer cells. The level of histone hyperacetylation and $p21^{Waf1}$ expression increased, as reported also for butyrate. The mechanism of these effects was found to be reversible inhibition of HDACs presumably mediated by binding of the hydroxamic acid moiety of TSA to a zinc in a tubular pocket in the catalytic site of HDAC (Marks *et al*., 2000). Although the effect of TSA may appear to be nonspecific and could affect many genes, it is of interest that the activity of only about 2% of expressed genes is affected (Marks *et al*., 2000).

The (8;21) translocation, found in 12% of AML, creates the chimaeric fusion protein AML1–ETO. This protein binds to DNA and recruits a transcription repression complex that includes the histone deacetylase 1 enzyme. TSA could partially reverse ETO-mediated transcriptional repression and this could be the basis for its ability to induce differentiation (Ferrara *et al*., 2001).

In addition to its effects on established tumour cell lines *in vitro*, TSA has been shown to induce or enhance the differentiation of AML cells isolated freshly from clinical samples from patients with different AMLs ranging from M0 to M7.

Suberoylanilide Hydroxamic Acid (SAHA)

SAHA is the prototype of a family of hybrid polar compounds that inhibit histone deacetylase and are potent inducers of cell differentiation and/or apoptosis in certain transformed cells in culture including erythroleukaemia, myeloma, leukaemia, bladder carcinoma, breast carcinoma and prostate carcinoma (Marks *et al*., 2000). They inhibit HDACs and induce cell differentiation at very low concentrations ($0.01–1\ \mu mol\ L^{-1}$). The hydroxamic acid moiety of SAHA can bind to the putative catalytic site of HDACs in the same fashion as TSA (Marks *et al*., 2000).

Like other HDAC inhibitors, SAHA increased the level of $p21^{Waf1}$ by enhancing its transcription (Marks *et al*., 2000). The growth of human prostate tumour cells in immunodeficient mice was suppressed by SAHA at concentrations that exerted no toxicity to the animals although histone hyperacetylation was observed in both normal and tumour cells *in vivo*. The selective effect on tumour cells makes SAHA an excellent candidate for therapy. Further, the ability of SAHA to suppress mammary and lung carcinogenesis makes it also a potential chemopreventive agent (Marks *et al*., 2000).

Peroxisome Proliferator–Activator Receptor Ligands

Peroxisome proliferator–activator receptors (PPARs) are nuclear receptors that function as ligand-dependent transcriptional regulators. The PPARs include three receptor subtypes encoded by separate genes: *PPAR-*α, *PPAR-*δ, and *PPAR-*γ. PPARs form heterodimers with the retinoid X receptors (RXRs), which bind to the consensus sequence of nucleotides in the regulatory regions of PPAR-regulated genes. Activation of PPAR-γ by the naturally occurring ligand 15-deoxy-$\Delta^{12,14}$-prostaglandin J2 (15d-PGJ2), has been linked to adipocyte differentiation, regulation of glucose homeostasis, inhibition of macrophage and monocyte activation and inhibition of tumour cell proliferation. PPARs and their ligands have also been implicated in the control of cell growth and differentiation (Roberts-Thomson, 2000). 15d-PGJ2 and other PPAR-γ activators (e.g. thiazolidinediones such as troglitazone and pioglitazone) were found to induce differentiation in a variety of cancer cell lines including liposarcoma, colon carcinoma, breast carcinoma, and prostate carcinoma. Treatment with pioglitazone induced terminal differentiation in primary human liposarcoma cells. Further, RXR-specific ligands were also potent adipogenic agents in cells expressing the PPAR-γ/RXR-α heterodimer, and simultaneous treatment of liposarcoma cells with both PPAR-γ- and RXR-specific ligands resulted in an additive stimulation of differentiation. These results suggested that PPAR-γ ligands and RXR-specific retinoids may be useful agents for the treatment of liposarcoma. Ligand activation of PPAR-γ in cultured breast cancer cells caused lipid accumulation, changes in breast epithelial gene expression associated with a more differentiated, less malignant state and a reduction in growth rate and clonogenic capacity

of the cells. PPAR-γ activation in colon cancer cells suppressed growth, increased expression of carcinoembryonic antigen and suppressed the expression of many genes associated with colon cancer. In addition, animal studies have shown that transplantable tumours derived from human colon cancer cells show a reduction in growth when mice are treated with troglitazone. These results indicate that the growth and differentiation of the above cancer cells can be modulated through PPAR-γ.

Phenylacetate, which is structurally similar to clofibrate, a PPAR ligand, was found to activate this receptor. In human prostate carcinoma, melanoma and glioblastoma cell lines there was a close correlation between drug-induced cytostasis, increased expression of the endogenous PPAR and PPAR activation. More recently, it was found using *in vitro* models that a close correlation exists between tumour growth arrest by phenylacetate and activation of PPAR-γ. The ability to bind and activate PPAR-γ was common to biologically active analogues of phenylacetate and corresponded to their potency as antitumour agents, whereas an inactive derivative had no effect on PPAR-γ. These results suggest that PPAR-γ may mediate tumour growth inhibition by aromatic fatty acids.

Planar–Polar Solvents

N,N′-Hexamethylenebisacetamide (HMBA) is a hybrid polar compound that was found to induce both the differentiation of murine erythroleukaemia cells and the differentiation of various other cell lines *in vitro* and in patients (Marks *et al.*, 1995). Detailed investigations using murine erythroleukaemia cells have delineated the early and late effects of HMBA (reviewed in Marks *et al.*, 1995). The induction of differentiation by HMBA required treatment at concentations of 3–10 mmol L^{-1} for 3–15 days. Several investigations have been carried out with human cancer cells. HMBA induced the differentiation of human teratocarcinoma, bladder carcinoma and glioma cells. In addition, HMBA suppressed cell growth and decreased the tumorigenicity of the cells. The effects of HMBA were associated with suppression of the expression of growth factors such as transforming growth factor alpha, keratinocyte growth factor and an autocrine factor teratocarcinoma-derived growth factor-1 that is related to epidermal growth factor. HMBA-induced cessation of proliferation was mediated, in part, by increased expression of the cyclin-dependent kinase inhibitor p27, enhanced association of p27 with cyclin E/cyclin-dependent kinase 2 (CDK2 complex) and suppression of kinase activity associated with cyclin E/CDK2. The formation of E2F complexes with pRB and the protein p130 has also been implicated in growth inhibition. Other effects of HMBA include enhancement of gap-junctional intercellular communication and induction of cell death, both of which can contribute to suppression of growth.

Vitamin D_3

The metabolically active form of vitamin D, 1α,25-dihydroxyvitamin D_3 (1α,25$(OH)_2D_3$) plays a key role in calcium homeostasis in the body. In addition, 1α,25$(OH)_2D_3$ can regulate cell growth and differentiation of normal and malignant cells including colon carcinoma cells, prostate cancer cells and neuroblastoma cells. Because the clinical use of 1α,25$(OH)_2D_3$ in differentiation therapy was hampered by its hypercalcaemic activity, various analogues have been synthesized with the hope of identifying some that are as potent as or more potent than vitamin D_3 in inducing tumour cell differentiation but less hypercalcaemic. Indeed, analogues such as 1,25-dihydroxy-23-ynecholecalciferol (23-D), 1,25-dihydroxy-Δ^{16}-23-yne-cholecalciferol (16-23-D) and 1,25-dihydroxy-22,24-diene-24,26,27-tribromovitamin D_3 (Seocalcitol EB 1089) have reduced action on calcium metabolism but high potency in cancer cell growth suppression and differentiation induction (Hansen *et al.*, 2000).

The analogue 19-nor-hexafluoride (1,25-dihydroxy-16-ene-23-yne-26,27-difluoro-19-nor-vitamin D_3) induced cell cycle arrest of prostate cancer cell lines and upregulated the cyclin-dependent kinase inhibitor (CDKI) p21^{Waf1}. Partial differentiation was evidenced by increased expression of PSA in LNCaP cells and E-cadherin in LNCaP and PC-3 cells. In mouse erythroleukaemia cells 1α,25$(OH)_2D_3$ induced haemoglobin and acetylcholinesterase activity. 1α,25$(OH)_2D_3$ and its analogues, 20-epi-22-oxa-24a,26a,27a-tribromo-1,25-dihydroxy vitamin D_3 (KH 1060); 20-epi-1,25$(OH)_2D_3$ 1,25$(OH)_2$-16-ene-D_3, and 1,25$(OH)_2$-16-ene-23-yne-D_3, especially KH 1060, were potent inducers of differentiation of breast cancer cell lines. Apoptosis was induced in four of six cells lines. Vitamin D_3 stimulated the differentiation of LA-N-5 neuroblastoma cells, which was associated with a decrease in N-myc expression and inhibition of cell growth. Two vitamin D_3 analogues, KH 1060 and EB 1089, can cause the same *in vitro* effects on LA-N-5 human neuroblastoma cells at 1% of the concentration required with vitamin D_3, thus making them more attractive candidates for clinical use.

Vitamin D_3 and many of its analogues exert their actions on cell growth and differentiation through the activation of nuclear vitamin D receptor (VDR), a ligand-activated transcription factor, which is a member of the superfamily of steroid hormone receptors. The ligand-activated VDR is found in complex with the retinoid X receptor (RXR) and enhances gene transcription from VD response elements (VDREs). The transcriptional activity can be enhanced by costimulation with vitamin D_3 and 9-*cis*-retinoic acid, the RXR ligand.

The vitamin D receptor has been detected in many types of cancer cell, including colon carcinoma, prostate, lung cancers, osteosarcomas and other types of cancer. Mutations and rearrangements do not appear to be associated with the development of these cancers.

Evidence that the antiproliferarive effects of 1α,25 $(OH)_2D_3$ were mediated by the nuclear receptor VDR was obtained from studies of transfection of VDR into receptorless prostate cancer cells thus restoring their response to the growth inhibitory effects of D_3. The downstream genes that may regulate the effects of D_3 on cell growth may include p21^{WAF1} and c-*fos*. In addition, vitamin D_3 and its analogues are potent inducers of transforming growth factor beta (TGF-β) in breast cancer cells and prostate cancer cells and it was suggested that TGF-β may contribute to the mechanism by which 1α,25$(OH)_2D_3$ promotes cellular differentiation.

U937 and THP-1 cell differentiation with 1α,25$(OH)_2$ D_3 correlated with gene transcription and functional expression of inducible nitric oxide synthase (iNOS). Inhibition of iNOS activity by N^{γ}-monomethyl-L-arginine significantly decreased *in vitro* cell differentiation with 1α,25$(OH)_2D_3$. Thus, it appears that NO plays a role in the growth arrest and terminal differentiation of promonocytic leukaemia cells.

Experiments utilizing 1α,25-$(OH)_2D_3$ analogues and kinase/phosphatase inhibitors suggested that tyrosine kinase and serine/threonine phosphorylation cascades, rather than vitamin D_3 receptor-mediated signals, were involved in some aspects of 1α,25-$(OH)_2D_3$'s actions. Both 1α,25-$(OH)_2D_3$ and the nongenomic analogue 1α,25-$(OH)_2D_3$ (HF) increased expression of PKC-α and PKC-δ and their translocation to the nucleus of the cell. The effects of HF were attenuated by the nongenomic antagonist 1β,25-$(OH)_2D_3$, suggesting that changes in PKC expression are mediated by a nongenomic signalling pathway. Enhanced phosphorylation of a variety of cellular proteins at serine and threonine residues was observed immediately after 1α,25-$(OH)_2D_3$ addition. It was proposed that 1α,25-$(OH)_2D_3$ primes acute promyelocytic leukaemia cell line NB4 cells for TPA-induced monocytic differentiation by increasing the expression of specific PKC isoforms and inducing the specific phosphorylation of key protein signalling intermediates. Indeed, 1α,25-$(OH)_2D_3$ and some of its conformationally locked (6-s-*cis* vs 6-s-*trans*) analogues, which bind poorly to VDR, increased p42MAPK phosphorylation in NB4 cells in a time- and dose-dependent manner, with the earliest response detectable at 30 s, which excludes a role for nuclear VDR in this effect.

Retinoids

Retinoids are naturally occurring and synthetic analogues of vitamin A. The physiological functions of naturally occurring retinoids include regulation of embryonal development and maintenance of the proper differentiation of many tissues in the adult. Retinoids also act pharmacologically to restore regulation of differentiation in certain malignant cells *in vitro*. The effects of retinoids on cell differentiation have been studied extensively in several cultured cell lines derived from embryonal carcinoma (EC), normal and malignant keratinocytes, premonocytic and myeloid leukaemias, neuroblastoma, and melanoma. In most of these cell types, retinoids enhance differentiation. However, in cultured keratinocytes and squamous cell carcinomas (SCCs), retinoids inhibit squamous differentiation. Because in many cases the squamous differentiation of normally nonkeratinizing epithelial cells is aberrant, the effect of retinoids can be viewed as restoration of the normal nonkeratinizing phenotype.

Retinoids appear to enhance predetermined programmes in cells that have the potential to undergo differentiation along one or more specific pathways. For example, in F9 EC cells, retinoids induce endodermal differentiation, whereas in human embryonal carcinoma cells, retinoids induce a neuronal differentiation. In P19 EC cells retinoids can induce both myogenic differentiation and neuronal differentiation depending on the concentration of retinoid used. In HL-60 myeloid leukaemia cells that have the potential to undergo either myeloid or monocytoid differentiation, retinoids can only induce the myeloid pathway. In contrast, all-*trans*-retinoic acid (ATRA) was able to induce three different pathways, ectodermal, mesodermal and endodermal, in a developmentally pluripotent germ cell tumour.

There are only limited reports on induction of differentiation of solid tumour cells *in vivo* and they are restricted to murine embryonal carcinoma cells injected into syngeneic mice. A significant degree of differentiation was observed in teratocarcinomas when RA was injected directly into the tumour (Speers and Altmann, 1984) or when several retinoids were administered in the diet (McCue *et al*., 1988).

The mechanism(s) underpinning the ability of retinoids to inhibit proliferation and clonogenicity of malignant cells and modulate their differentiation *in vitro* are thought to be related to their ability to modulate gene expression. This ability is mediated by retinoic acid receptors (RARs) and retinoid X receptors (RXRs), nuclear receptors which belong to the superfamily of steroid hormone receptors (reviewed in Chambon, 1996). RARs bind both natural retinoids ATRA and 9-*cis*-retinoic acid (9-*cis*-RA), whereas the RXRs bind only 9-*cis*-RA (Chambon, 1996). Retinoid receptors modulate the expression of their target genes by interacting as either homodimers or heterodimers with specific DNA response elements (Chambon, 1996). RXRs dimerize with and enhance the transcriptional activity not only of RARs but also of thyroid hormone receptor, vitamin D_3 receptor, PPARs and several orphan receptors (Chambon, 1996). The RA nuclear receptors may also interact antagonistically with components of other signal transduction pathway such as the AP-1 (jun–fos) transcription factor complex in stromelysin, collagenase promoters possibly by direct protein–protein interactions.

It appears that RARs may be necessary but not sufficient for mediating the effects of most retinoids on cell growth and differentiation. For example, RAR-αs

expressed by most human leukaemias (fresh cells and cell lines) regardless of their responsiveness to RA. Thus, gene regulation by retinoids may depend on the relative concentrations of particular RARs, RXRs and cofactors, the nature of the RAREs in that gene and the presence of a number of other transcription factors.

Many genes are regulated by retinoids. Some may be regulated directly by binding of receptors to their promoter region (e.g. RAR-β, cytochrome P-450RAI, homeobox genes) whereas others may be regulated indirectly by antagonism of the transcription factor activator protein 1 (e.g. collagenase). Different receptors may regulate the expression of distinct subsets of genes related to either growth or differentiation. Recently, a systematic study examined the expression patterns of genes in the cell line NB4 before and after ATRA treatment using complementary DNA array, suppression–subtractive hybridization and differential display polymerase chain reaction. A total of 100 genes were up-regulated 12–48 h after ATRA treatment, while the expression of 59 of 69 down-regulated genes was suppressed within 8 h. The transcriptional regulation of eight induced and 24 repressed genes was not blocked by cycloheximide, which suggests that these genes may be direct targets of the ATRA signalling pathway. Several cytosolic signalling pathways, including JAKs/STAT and MAPK, have been implicated in the effects of ATRA.

Tumour Necrosis Factor Alpha

Tumour necrosis factor alpha (TNF-α) is a 17-kDa cytokine which can induce differentiation, growth arrest or apoptosis. The effects of TNF on cellular differentiation have been demonstrated in malignant haematopoietic cells and cells derived from solid tumours. TNF treatment of a human rectal adenocarcinoma cell line induced features of early-stage differentiation, with goblet-like cell characteristics. The cells were growth-inhibited by TNF and exhibited enhanced expression of high molecular weight mucin glycoproteins. The induction of these differentiation characteristics correlated with reductions in epidermal growth factor receptor (EGF-R) levels and activity. TNF-α reduced proliferation and DNA synthesis in a human pancreatic adenocarcinoma cell line, without loss of viability. Parallel to these changes was an increase in carbonic anhydrase II mRNA, a marker for pancreatic cells of ductal origin. Likewise, TNF-α induced neuroblastoma cell lines to produce neurofilament proteins and adhesion molecules associated with differentiation. This induction was mediated by a calcium-independent inducible form of nitric oxide synthase and was accompanied by growth arrest. In cultured mammary carcinoma cells, TNF-α was shown to stimulate MUC1, isoactin and epithelial cell adhesion molecule expression. The various biological activities of TNF are signalled through two distinct receptors, p55 and p75. The differentiation-inducing activity of TNF-α was signalled through the TNF receptor, p55, via protein kinase C.

Interferons

Interferons (IFNs) are a unique class of cytokines in that they stimulate antiviral, antitumour and antigen presentation. They have been classified as interferon alpha (IFN-α), beta (IFN-β) and gamma (IFN-γ) according to their cellular origin (leukocytes, fibroblasts and lymphocytes, respectively). Because of their ability to inhibit malignant cell growth and phenotypic transformation, promote cellular differentiation and enhance apoptosis, they may be useful for cancer therapy.

Interferons inhibit *in vitro* the proliferation and colony formation in agar of malignant cells including myeloma, Burkitt lymphoma, erythroleukaemia, breast carcinoma, neuroblastoma and lung cancer. IFNs also modulated cell differentiation; in some cell types, IFNs stimulated differentiation and in others inhibited differentiation. IFN-β stimulated the differentiation of human melanoma cell lines as evidenced by increased pigmentation. In human HL-60 myeloid leukaemia cells, all three IFNs inhibited cell proliferation, but only IFN-γ induced differentiation. IFN-β suppressed the growth and enhanced squamous differentiation and apoptosis in cultured NSCLC cells. IFN-β enhanced squamous differentiation in cell lines with squamous features, whereas it induced apoptosis in cell lines with glandular features. IFN-β and IFN-γ modulated the growth and differentiation of head and neck squamous carcinoma in monolayer culture.

The interferons exert their multiple effects on cell proliferation, differentiation and possibly apoptosis by binding to cell-surface receptors through which they transmit a signal that results in the activation of transcription of numerous genes within minutes of receptor occupancy (Darnell *et al*., 1994). IFN-α and IFN-β share the same cell-surface receptor to transmit their signal, whereas IFN-γ has its own receptor. The signal transduction is initiated by the binding of interferon to its receptor. In the case of IFN-α, binding to the IFN-α/β receptor results in the activation of two types of tyrosine kinases in the cytoplasm: the Janus kinase 1 (Jak 1) and the tyrosine kinase 2. These kinases phosphorylate cytoplasmic proteins that contain Src homology (SH2 and SH3) domains called Stats (signal transducers and activators of transcription). Stat 1α is a 91-kDa protein, which can also appear as a spliced form of 84 kDa (Stat 1β), and Stat 2 is a 113-kDa protein. The Stat proteins undergo phosphorylation and form a complex called the ISGF3 (interferon-stimulated gene factor 3) with a cytoplasmic protein of 48 kDa. The complex is translocated into the cell nucleus and there the ISGF3 complex binds via the 48-kDA protein to a specific DNA sequence (IFN-stimulated response element or ISRE) in the promoter region of IFN-stimulated genes

(ISGs) and activates their transcription (Darnell *et al.*, 1994). The binding of IFN-γ to its receptor activates JAK1 and JAK2, which phosphorylate Stat 1α. The activated Stat 1α then migrates into the cell nucleus and binds directly to IFN-γ activated site (GAS) and stimulates gene transcription (Darnell *et al.*, 1994). Thus, all three IFNs share Stat 1 and JAK1 in their signalling. The STAT signalling pathway can play essential roles in cell differentiation, cell cycle control and development. In addition, it has been shown that activation of the STAT signalling pathway by IFN-γ in cervical and breast cancer cells can induce apoptosis through the induction of caspase 1 gene expression.

IFN-γ was found to inhibit cell proliferation by blocking progression in mid-G_1 phase associated with a block in Rb phosphorylation, which may be responsible for the down-regulation of cdk2–6, cyclin A and c-Myc in Daudi Burkitt lymphoma cells and in normal epidermal, mammary and tracheobronchial epithelial cells. (See the chapter *Signalling by Cytokines*.)

Combinations of Agents

The same rationale that has led to the successful application of combination of cytotoxic agents with nonoverlapping toxicities and distinct mechanisms of action can be applied to differentiation-inducing agents. Extensive studies were carried out with the myeloid leukaemia cell line HL-60 to define the efficacy of combinations of retinoids and other agents (Breitman and He, 1990). Such studies have demonstrated additive or synergistic effects of retinoids and the following agents: dimethyl sulfoxide, dimethylformamide, hexamethylenebisacetamide, butyrate, tiazofurin, IFNs, TNF-α, 6-thioguanine, 5-aza-2′-deoxycytidine, actinomycin, granulocyte colony-stimulating factor and interleukin-6. Studies with embryonal carcinoma have shown that a combination of ATRA and cyclic AMP and cAMP elevating agents induces a distinct pathway of endodermal differentiation compared with that induced by RA alone. The combination of RA and tamoxifen was found to be more effective in suppressing the growth of human breast carcinoma cells than each agent alone.

ATRA Plus IFN

The efficacy of the combination of IFN-α or IFN-γ and ATRA in induction of differentiation and growth inhibition was detected initially *in vitro* in established leukaemia cell lines (e.g. HL60, U937). The effects of IFNs and retinoids were either additive or synergistic, depending on the cell system evaluated, and distinct differentiation pathways could be stimulated in different cell types. Further, IFN-α restored responsiveness to ATRA in ATRA-resistant HL60 cells. Likewise, IFN-γ restored responsiveness to ATRA in ATRA-resistant v-myc expressing U937 cells. Further studies with fresh leukaemic cells from AML patients treated in short-term cultures with IFNs and ATRA also demonstrated synergistic growth inhibition and differentiation induction. In addition, increased inhibition of clonal growth and differentiation was observed after *in vitro* RA treatment of cells from patients treated with IFN-α *in vivo*. Additive or synergistic growth inhibition of various human tumour cell lines *in vitro* was reported for the combination of IFN-α or IFN-γ and ATRA or other retinoids. These effects were sometimes also associated with differentiation induction, particularly in neuroblastoma cells. Tumour cell types in which the combination of the two agents showed greater growth inhibition than each agent alone included breast carcinoma, osteosarcoma, ovarian carcinoma, squamous cell carcinoma, neuroblastoma, non-small cell lung carcinoma and cervical carcinoma. Furthermore, the growth of human neuroblastoma tumours in nude mice was inhibited synergistically by the combination of IFN-γ and ATRA.

In various cell lines, including those derived from APL, ATRA induces directly the expression of two transcription factors, Stat 1 and IRF-1 which play central roles in the IFN signal transduction. In addition, ATRA induces IFN-α synthesis and enhances the IFN-induced Stat activation and increased the binding of ISGF-3 and the transcription of ISGs in embryonal carcinoma and breast cancer cells.

Retinoids Plus 1,25-Dihydroxyvitamin D_3

Combinations of $1\alpha,25(OH)_2D_3$ analogues and ATRA or 9-*cis*-RA exhibit additive or synergistic effects on differentiation of leukaemia cell lines (e.g. HL-60, U937 and NB4). Furthermore, vitamin D_3 compounds augment apoptosis induced by 9-*cis*-RA in HL-60 cells and the combination down-regulates the expression of the anti-apoptotic gene product Bcl-2 while increasing the pro-apoptotic gene product Bax. The molecular mechanism underlying the effects of the combination of these agents involves the interactions of the nuclear receptors for vitamin D (VDR) and 9-*cis*-RA (RXR), which are able to form heterodimeric complexes and transcriptionally activate or repress target gene expression.

ATRA and Histone Deacetylase Inhibitors

TSA acted synergistically with ATRA to enhance the differentiation of the leukaemia cell lines U937, HL60 and NB4 cells. Further, the combined treatment of TSA with ATRA induced differentiation even in ATRA-resistant HL60 and NB4 cells. The mechanism of this interaction has been elucidated recently and it is related to the finding that transcriptional regulation by RARs involves modifications of chromatin by HDACs, which are recruited to RA-target genes by nuclear corepressors. The differentiation of APL cells is blocked by the PML–RAR-α fusion protein, which recruits the nuclear corepressor–HDAC complex. However, this block can be reversed by high doses of ATRA, through the release of HDAC activity from PML–RAR-α. Interestingly, inhibitors of histone

deacetylase such as TSA can potentiate retinoid-induced differentiation of ATRA-sensitive and restore the retinoid responses of RA-resistant APL cell lines (Lin *et al*., 1998). Further, TSA was shown recently to restore ATRA-dependent transcriptional activation and triggered terminal differentiation of primary blasts from acute myeloid leukaemia M2 and M4 in which the AML–ETO fusion protein blocks ATRA signalling (Ferrara *et al*., 2001). Likewise, TSA enhanced the response to ATRA in APL cells with t(11;17) translocation, which leads to the production of the fusion protein PLZF–RAR-α and a poor response to ATRA. In these cells, TSA inhibits the HDAC enzyme that is associated with PLZF–RAR-α corepressor complex but cannot be dissociated after the addition of high-dose ATRA alone.

ATRA and Arsenic Trioxide

Arsenic trioxide induces nonterminal differentiation of malignant promyelocytes and promotes apoptosis. In contrast, ATRA induces terminal differentiation that eventually leads to apoptosis. Both agents enhance the degradation of the leukaemogenic chimaeric protein PML–RAR-α resulting from the t(15;17) translocation characteristic of APL by different pathways, suggesting a possible therapeutic synergism. Indeed, arsenic trioxide synergized with ATRA to induce differentiation in ATRA-resistant APL cells but the dose and the sequence of their combination were important. Severe combined immunodeficient mice bearing APL–NB4 cells showed an additive survival effect after sequential treatment, but a toxic effect was observed after simultaneous treatment with ATRA and arsenic trioxide. These data suggest that combined arsenic trioxide and ATRA treatment may be more effective than single agents in ATRA-resistant patients (Jing *et al*., 2001).

IFN Plus Mezerein

Treatment of human melanoma cells with a combination of recombinant human IFN-β and the antileukaemic compound mezerein (MEZ) results in irreversible growth arrest, suppression of tumorigenic properties and terminal cell differentiation by increasing the expression of numerous genes.

TNF-α Plus 1α,25-dihydroxyvitamin D_3

The combination of TNF-α and 1α,25$(OH)_2D_3$ showed efficacy in enhancing the differentiation of chemically transformed human mammary epithelial cell lines as evidenced by increased expression of pan cytokeratin, mucin 1 and the adhesion molecules CEA, ICAM-1 and CD44v6.

CLINICAL ASPECTS

The potential of agents that stimulate cell differentiation to serve for cancer therapy has been studied extensively *in vitro* and in animal models. Such agents can suppress growth and enhance differentiation, which may also lead to apoptosis. Only a few of the numerous differentiation inducing agents shown in Table 1 have been examined in clinical trials and even for those that have been investigated, most trials were phase I or II. The only definitive demonstration of the efficacy of differentiation therapy is the treatment of acute promyelocytic leukaemia patients with ATRA.

HMBA

A few clinical studies explored the potential of using HMBA in cancer therapy. Young *et al*. (1988) treated 33 patients with different types of advanced cancer with HMBA by 10-day continuous i.v. infusion courses. Patients with no disease progression were given repeat courses of HMBA at 28-day intervals. Thrombocytopenia with haemorrhage and disorientation and confusion were the dose-limiting toxic effects and the maximal tolerated dose was 28 $g\,m^{-2}\,day^{-1}$ for 10 days. HMBA was cleared from plasma with a mean half-life of 2.5 h. Mean plasma steady-state HMBA concentrations of 1 $mmol\,L^{-1}$ could be maintained for 10 days with acceptable patient tolerance; however, HMBA concentrations in excess of 1.4 $mmol\ L^{-1}$ for 10 days were toxic. Nonetheless, objective antitumour effects were observed in five patients, with one partial remission in a patient with non-small cell lung cancer. Transient regression of cutaneous metastases was observed in three patients with breast carcinoma and one patient with colorectal carcinoma. Andreeff *et al*. (1992) conducted a phase II clinical trial of HMBA in 41 patients with myelodysplastic syndrome (MDS) or acute myelogenous leukaemia (AML). HMBA was administered by continuous infusion for 10 days and repeated after an interval of 18–75 days. HMBA induced a complete remission (CR) in three patients and a partial remission (PR) in six patients. The median duration of CR was 6.8 months (range 1.3–16 months) and that of PR was 3.7 months (range 1–7 months). The mean HMBA plasma levels were 0.86 ± 0.04 and $0.87 \pm 0.12\,mmol\,L^{-1}$ in responders and nonresponders, respectively. In certain patients HMBA induced differentiation of transformed haematopoietic precursors as indicated by morphological and chromosome analyses. The most prominent toxicity, thrombocytopenia, was reversible. Thus, it appears that it was not possible to achieve *in vivo* HMBA concentrations of 3–5 $mmol\,L^{-1}$ that are optimal for inducing differentiation of a variety of transformed cell lines *in vitro*. This has led to a search for more active compounds and a second generation of hybrid polar compounds were synthesized and characterized. Among these compounds was SAHA, which exhibited a far greater potency in inducing murine erythroleukaemia cell differentiation than HMBA. SAHA has been evaluated in animal models for antitumour efficacy and toxicity phase I human clinical trials are

ongoing in which analysis of histone acetylation serves as an intermediate marker of HDAC inhibitory activity (Marks *et al*., 2000).

Butyrate

Only very few clinical trials have been conducted with butyrate or its derivatives. A case report described the results of parenteral administration of butyrate ($500\,mg\,kg^{-1}\,day^{-1}$) for 10 days to a child with AML in relapse, who was resistant to conventional therapy. This treatment resulted in elimination of myeloblasts from the peripheral blood. An indication for differentiation has come from an analysis of the bone marrow, which showed an increase in mature myeloid cells and a reduction in myeloblasts from 80 to 20%. No impairment of liver or renal function and no coagulation abnormalities were observed during butyrate treatment at the above dose. Arginine butyrate was given in a phase I trial together with interleukin-2 to six patients with advanced metastatic colon cancer delivered via continuous intravenous infusion on days 1–6 with escalating doses starting at $2\,g\,kg^{-1}\,day^{-1}$. A daily ArgB dose of $2\,g\,kg^{-1}$ was delivered for nine cycles. No clinical responses were seen. The dose-limiting toxicities were fatigue and liver function abnormalities.

A clinical trial with the prodrug tributyrin compared the toxicities of escalating doses from 50 to $400\,mg\,kg^{-1}\,day^{-1}$ given p.o. after an overnight fast, once daily for 3 weeks, followed by a 1-week rest. Thirteen patients with solid tumours were treated and showed variable toxicities. Grade 1 and 2 toxicities included diarrhoea, headache, abdominal cramping, nausea, anaemia, constipation, azotaemia, lightheadedness, fatigue, rash, alopecia, odour, dysphoria and clumsiness. Grade 3 toxicities consisted of nausea, vomiting and myalgia. Peak plasma butyrate concentrations ranged from 0 to $0.45\,mmol\,L^{-1}$ (between 0.25 and 3 h after dose), but butyrate disappeared from plasma by 5 h after dose. Butyrate pharmacokinetics were not different on days 1 and 15. Because peak plasma concentrations near those effective *in vitro* ($0.5–1\,mmol\,L^{-1}$) were achieved, the authors considered pursuing dose escalation with dosing three times daily, beginning at a dose of $450\,mg\,kg^{-1}\,day^{-1}$ (Conley *et al*., 1998). Thus, there is only one case of differentiation *in vivo* by butyrate.

Aromatic Fatty Acids

In a phase I trial, phenylacetate was given to 17 patients with advanced solid tumours as a single i.v. bolus dose followed by a 14-day continuous i.v. infusion. Twenty-one cycles of therapy were administered at four dose levels. Stable PSA levels were maintained for more than 2 months in three of nine patients with metastatic, hormone-refractory prostate cancer. Another patient experienced less bone pain. One of six patients with glioblastoma multiforme had functional improvement for more than 9 months. Continuous i.v. infusion rates resulting in serum phenylacetate concentrations exceeding the K_m ($105.1 \pm 44.5\,\mu g\,mL^{-1}$) often resulted in rapid drug accumulation and dose-limiting toxicity (emesis and reversible central nervous system depression). Because the goal is to maintain serum drug concentrations greater than $250\,\mu g\,mL^{-1}$, which are required for *in vitro* activity, a subsequent trial was designed. In the second trial patients received phenylacetate as a 1-h infusion twice daily at two dose levels, 125 and $150\,mg\,kg^{-1}$ for a 2-week period. The intermittent exposure to drug concentrations exceeding $250\,\mu g\,mL^{-1}$ allowed time for drug elimination to occur between doses to minimize accumulation. Patients who did not experience dose-limiting toxicity or disease progression received additional cycles of treatment at 4-week intervals. Eighteen patients received 27 cycles of therapy. Dose-limiting toxicity, consisting of reversible central nervous system depression, was observed in three patients at the higher dose level. One patient with refractory malignant glioma had a partial response, and one with hormone-independent prostate cancer achieved a 50% decline in PSA level, which was maintained for 1 month. These phase I studies have indicated that phenylacetate at $125\,mg\,kg^{-1}$ twice daily for two consecutive weeks is well tolerated and that high-grade gliomas and advanced prostate cancer are plausible targets for phase II clinical trials (Thibault *et al*., 1995).

Recent phase I studies with phenylbutyrate given by i.v. or oral routes have shown that this drug is well tolerated clinically at concentrations which effect acetylation of histones *in vitro*. However, higher doses lead to reversible CNS depression. Although these studies do not allow an assessment of efficacy, the current development of phenylbutyrate is for combination with other agents based on indications from *in vitro* studies (Gore and Carducci, 2000).

PPAR Ligands

The finding that PPAR-γ agonists can induce terminal differentiation of normal preadipocytes and human liposarcoma cells *in vitro* suggested that such ligands may be useful clinically because the differentiation status of liposarcoma is predictive of clinical outcomes. Thus, enhancement of the differentiation status of a liposarcoma may have an impact on its clinical behaviour. With this rationale in mind, the PPAR-γ ligand troglitazone was administered to three patients with intermediate- to high-grade liposarcomas. Troglitazone induced histological and biochemical differentiation *in vivo* as evidenced by extensive lipid accumulation in tumour cells and increases in tumour triglycerides compared with pretreatment biopsies. In addition, differentiation markers in the adipocyte lineage were induced and the proliferation marker Ki-67 was decreased. These results indicate that differentiation

can be induced pharmacologically in a human solid tumour (Demetri *et al*., 1999).

A larger phase II study was conducted in 41 patients with histologically confirmed prostate cancer and no symptomatic metastatic disease. These patients were treated orally with troglitazone. Treated patients have shown prolonged stabilization of PSA. In addition, one patient had a dramatic decrease in serum PSA to nearly undetectable levels. These data suggest that PPAR-γ has the potential to be used for differentiation therapy of human prostate cancer and this potential should be explored further.

Vitamin D_3 Analogues

The clinical use of 1α,25$(OH)_2D_3$ in cancer prevention or differentiation therapy is hampered by its hypercalcaemic activity. To overcome this limitation, new vitamin D analogues have been synthesized and screened for exerting potent cell regulatory effects, but with weaker effects on the calcium metabolism than those of 1α,25$(OH)_2D_3$. Several such analogues have been identified including 23-D, 16-23-D and Seocalcitol (EB 1089) that have the potential to be used clinically owing to their reduced action on calcium metabolism.

Analogues of 1α,25$(OH)_2D_3$ have been synthesized which exhibited at least 10-fold higher potency than the parental 1α,25$(OH)_2D_3$ in inhibition of clonal growth of HL-60 myeloid leukaemia cells. (22*R*)-1α,25-$(OH)_2$-16,22,23-triene-D_3 had an ED_{50} of approximately $6\times10^{-11}\,mol\,L^{-1}$, whereas 1α,25$(OH)_2D_3$ gave an ED_{50} of approximately $10^{-8}\,mol\,L^{-1}$ on clonogenic cells from patients with AML. It was suggested that clinical trials, especially in a setting of minimal residual disease, may be warranted after assessment of *in vivo* toxicity. *Ex vivo* purging of leukaemic cells from normal human peripheral blood mononuclear cells for stem cell transplantation has also been proposed with these analogues because they inhibited the growth of leukaemia cells without affecting normal bone marrow CFU-GM cells and only mild inhibiton of clonogenic cells from peripheral blood.

EB 1089 was found to be 50–200 times more potent than 1α,25$(OH)_2D_3$ with respect to regulation of cell growth and differentiation both *in vitro* and *in vivo*, yet it displayed a reduced calcaemic activity *in vivo* compared with that of 1α,25$(OH)_2D_3$. Recent clinical evaluation of EB 1089, which was focused mainly on establishing a maximum tolerated dose in cancer patients, has demonstrated that the low calcaemic activity observed in animals can be reproduced in cancer patients. Furthermore, EB 1089 induced regression of tumours (e.g. hepatocellular carcinoma). Although these findings suggest that the development of EB 1089 as an anticancer drug is warranted, such a decision must await the completion of ongoing controlled clinical trials (Hansen *et al*., 2000).

Retinoic Acid

Acute promyelocytic leukaemia (APL), which is designated M3 in the FAB classification of acute myeloid leukaemias, was found to respond to orally administered ATRA ($45\,mg\,m^{-2}$) *in vivo* (Huang *et al*., 1988). In contrast to classical anthracycline–cytarabine chemotherapy, ATRA induced complete remission in about 85% of APL patients by inducing differentiation of blast cells to granulocytes without inducing aplasia of the bone marrow. Although complete remissions occur in most APL patients, the response is transient, possibly because ATRA induces P-450 enzyme that leads to a progressive decline in plasma drug concentrations. Consequently, resistance develops and all patients treated with ATRA alone eventually relapse.

In an attempt to overcome this problem, ATRA was used in several randomized and nonrandomized trials for remission induction followed by several cycles of consolidation chemotherapy, usually with idarubicin and cytosine arabinoside. The results of such studies have shown that ATRA significantly increased event-free survival and survival and decreased the incidence of relapse by comparison to chemotherapy alone in newly diagnosed APL. This approach yielded an approximately 2.5-fold increase in the proportion of long-term survivors who have presumably been cured of this disease (Soignet *et al*., 1997). A recently completed randomized long-term European APL group study has confirmed the superiority of the combination of ATRA followed by chemotherapy over chemotherapy alone in newly diagnosed APL, and concluded that ATRA should be incorporated in the frontline treatment of APL (Fenaux *et al*., 2000).

Recently, APL patients resistant to ATRA and conventional chemotherapy were found to respond to arsenic trioxide. This agent was shown to induce apoptosis in APL cells by down-regulating the antiapoptotic protein Bcl-2, increasing the activity of caspases and enhancing the degradation of PML–RAR-α (Chen *et al*., 1996; Soignet *et al*., 1998).

ATRA caused in some APL patients a potentially fatal retinoic acid syndrome characterized by fever, respiratory disease, hypotension and increased leukocyte count (Tallman *et al*., 2000).

The mechanism of differentiation induction by ATRA in APL cells is unique and fascinating. APL is associated with a specific t(15:17) translocation. The breakpoint on chromosome 17 reciprocally rearranges RAR-α with a putative transcription factor on chromosome 15 designated PML for promyelocytic leukaemia. Consequently, the APL cells produce a fusion protein called PML–RAR-α (de The *et al*., 1990), which retains the DNA binding domain of RAR-α and the ability to form homodimers. The formation of this fusion protein is thought to be the reason for the block in differentiation that drives the leukaemogenesis (Randolph, 2000). The fusion protein can

form homodimers and act as a dominant negative inhibitor of retinoid signalling by recruiting a corepressor complex, which includes an HDAC (Lin and Evans, 2000). Paradoxically, among myeloid leukaemias, APL (AML-M3) is the most sensitive to induction of differentiation by pharmacological levels of ATRA, despite the disruption of retinoid signalling. At the higher doses, ATRA causes the dissociation of corepressor from the PML–RAR-α and enhances the degradation of the fusion protein. Consequently, the normal RAR-α can function in the absence of a dominant negative inhibitor to regulate the expression of numerous genes associated with differentiation and growth (Zhang *et al*., 2000).

Combination of 13-*cis*-Retinoic acid and IFN-α

Several clinical trials using combination of retinoic acid and interferons were conducted in patients with skin cancer (Lippman *et al*., 1997), cervical carcinoma (Lippman *et al*., 1997), renal cell carcinoma and prostatic carcinoma. In a few of these, some efficacy has been demonstrated. However, there was no evidence that the treatment resulted in differentiation induction, and therefore, they are not included in this chapter.

CONCLUSIONS

The concept of differentiation therapy is a very attractive one in that it is based on understanding the block in normal differentiation and the mechanisms to overcome the block or induce alternative pathways to reach the mature phenotype. *In vitro* studies provided ample support for the concept and the elegant mechanistic and clinical studies with APL have demonstrated the proof of principle that differentiation therapy can be effective *in vivo*. Unfortunately, the success observed with APL in 1988 has not been reproduced with solid tumours. However, the recent preliminary studies with PPAR-γ ligands and histone deacetylase inhibitors suggest that these agents alone or in combination with other differentiation agents or with cytotoxic agents might show efficacy in clinical trials with various malignancies.

REFERENCES

Andreeff, M., *et al*. (1992). Hexamethylene bisacetamide in myelodysplastic syndrome and acute myelogenous leukemia: a phase II clinical trial with a differentiation-inducing agent. *Blood*, **80**, 2604–2609.

Breitman, T. R. and He, R. Y. (1990). Combinations of retinoic acid with either sodium butyrate, dimethyl sulfoxide, or hexamethylene bisacetamide synergistically induce differentiation of the human myeloid leukemia cell line HL60. *Cancer Research*, **50**, 6268–6273.

Chambon, P. (1996). A decade of molecular biology of retinoic acid receptors. *FASEB Journal*, **10**, 940–954.

Chen, G. Q., *et al*. (1996). *In vitro* studies on cellular and molecular mechanisms of arsenic trioxide (As_2O_3) in the treatment of acute promyelocytic leukemia: As_2O_3 induces NB4 cell apoptosis with downregulation of Bcl-2 expression and modulation of PML–RAR-α/PML proteins. *Blood*, **88**, 1052–1061.

Conley, B. A., *et al*. (1998). Phase I study of the orally administered butyrate prodrug, tributyrin, in patients with solid tumors. *Clinical Cancer Research*, **4**, 629–634.

Darnell, J. E., *et al*. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, **264**, 1415–1421.

de The, H., *et al*. (1990). The t(15;17) translocation of acute promyelocytic leukemia fuses the retinoic acid receptor α gene to a novel transcribed locus. *Nature*, **347**, 358–361.

Demetri, G. D., *et al*. (1999). Induction of solid tumor differentiation by the peroxisome proliferator-activated receptor-gamma ligand troglitazone in patients with liposarcoma. *Proceedings of the National Academy of Sciences of the USA*, **96**, 3951–3956.

Fenaux, P., *et al*. (2000). Long-term follow-up confirms the benefit of all-*trans*-retinoic acid in acute promyelocytic leukemia. European APL group. *Leukemia*, **14**, 1371–1377.

Ferrara, F. F., *et al*. (2001). Histone deacetylase-targeted treatment restores retinoic acid signaling and differentiation in acute myeloid leukemia. *Cancer Research*, **61**, 2–7.

Gore, S. D. and Carducci, M. A. (2000). Modifying histones to tame cancer: clinical development of sodium phenylbutyrate and other histone deacetylase inhibitors. *Expert Opinions in Investigation of Drugs*, **9**, 2923–2934.

Hansen, C. M., *et al*. (2000). Seocalcitol (EB 1089): a vitamin D analogue of anti-cancer potential. Background, design, synthesis, pre-clinical and clinical evaluation. *Current Pharmaceutical Design*, **6**, 803–828.

Huang, M. E., *et al*. (1988). Use of all-*trans*-retinoic acid in the treatment of acute promyelocytic leukemia. *Blood*, **72**, 567–572.

Illmensee, K. and Mintz, B. (1976). Totipotency and normal differentiation of single teratocarcinoma cells cloned by injection into blastocyst. *Proceedings of the National Academy of Sciences of the USA*, **73**, 549–553.

Jing, Y., *et al*. (2001). Combined effect of all-*trans*-retinoic acid and arsenic trioxide in acute promyelocytic leukemia cells *in vitro* and *in vivo*. *Blood*, **97**, 264–269.

Lin, R. J. and Evans, R. M. (2000). Acquisition of oncogenic potential by RAR chimeras in acute promyelocytic leukemia through formation of homodimers. *Molecular Cell*, **5**, 821–830.

Lin, R. J., *et al*. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature*, **391**, 811–814.

Lippman, S. M., *et al*. (1997). Retinoid-interferon therapy of solid tumors. *International Journal of Cancer*, **70**, 481–483.

Marks, P. A., *et al*. (1995). Inducing differentiation of cancer cells with hybrid polar compounds involves modulation of cell cycle regulatory proteins. In: Waxman, S. (ed.), *Differentiation Therapy, Challenges of Modern Medicine.* f117–127 (Ares-Serono Symposia Publications, Rome).

Marks, P. A., *et al*. (2000). Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *Journal of the National Cancer Institute*, **92**, 1210–1216.

McCue, P. A., *et al*. (1988). Effects of dietary retinoids upon growth and differentiation of tumors derived from several murine embryonal carcinoma cell lines. *Cancer Research*, **48**, 3772–3779.

Newmark, H. L., *et al*. (1994). Butyrate as a differentiating agent: pharmacokinetics, analogues, and current status. *Cancer Letters*, **78**, 1–5.

Otaka, M., *et al*. (1989). Antibody-mediated targeting of differentiation inducers to tumor cells: inhibition of colonic cancer cell growth *in vitro* and *in vivo*: a preliminary note. *Biochemical and Biophysical Research Communications*, **158**, 202–208.

Pierce, G. B., *et al*. (1978). *Cancer: a Problem of Developmental Biology* (Prentice-Hall, Engelwood Cliffs, NJ).

Randolph, T. R. (2000). Acute promyelocytic leukemia (AML-M3)—Part 1: Pathophysiology, clinical diagnosis, and differentiation therapy. Part 2: Molecular defect, DNA diagnosis, and proposed models of leukemogenesis and differentiation therapy. *Clinical Laboratory Science*, **13**, 98–105, 106–116.

Roberts-Thomson, S. J. (2000). Peroxisome proliferator-activated receptors in tumorigenesis: targets of tumour promotion and treatment. *Immunological Cell Biology*, **78**, 436–441.

Samid, D., *et al*. (1997). Phenylacetate and phenylbutyrate as novel, nontoxic differentiation inducers. *Advances in Experimental Medicine and Biology*, **400A**, 501–505.

Soignet, S., *et al*. (1997). All-*trans*-retinoic acid significantly increases 5-year survival in patients with acute promyelocytic leukemia: long-term follow-up of the New York study. *Cancer Chemotherapy and Pharmacology*, **40** (Suppl.), S25–S29.

Soignet, S. L., *et al*. (1998). Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. *New England Journal of Medicine*, **339**, 1341–1348.

Speers, W. C. and Altmann, M. (1984). Chemically induced differentiation of murine embryonal carcinomas *in vivo* transplantation of differentiated tumors. *Cancer Research*, **44**, 2129–2135.

Tallman, M. S., *et al*. (2000). Clinical description of 44 patients with acute promyelocytic leukemia who developed the retinoic acid syndrome. *Blood*, **95**, 90–95.

Thibault, A., *et al*. (1995). Phase I study of phenylacetate administered twice daily to patients with cancer. *Cancer*, **75**, 2932–2938.

Young, C. W., *et al*. (1998). Phase I trial and clinical pharmacological evaluation of hexamethylene bisacetamide administration by ten-day continuous intravenous infusion at twenty-eight-day intervals. *Cancer Research*, **48**, 7304–7309.

Zhang, J. W., *et al*. (2000). Mechanisms of all-*trans*-retinoic acid-induced differentiation of acute promyelocytic leukemia cells. *Journal of Bioscience*, **25**, 275–284.

FURTHER READING

Carducci, M. A., *et al*. (1996). Phenylbutyrate induces apoptosis in human prostate cancer and is more potent than phenylacetate. *Clinical Cancer Research*, **2**, 379–387.

DeLuca, L. M. (1991). Retinoids and their receptors in differentiation, embryogenesis, and neoplasia. *FASEB Journal*, **5**, 2924–2933.

Elstner, E., *et al*. (1995). 20-epi-vitamin D_3 analogues: a novel class of potent inhibitors of proliferation and inducers of differentiation of human breast cancer cell lines. *Cancer Research*, **55**, 2822–2830.

Hu, J., *et al*. (1999). Long-term survival and prognostic study in acute promyelocytic leukemia treated with all-*trans*-retinoic acid, chemotherapy, and As_2O_3: an experience of 120 patients at a single institution. *International Journal of Hematology*, **70**, 248–260.

Kosugi, H., *et al*. (1999). Histone deacetylase inhibitors are the potent inducer/enhancer of differentiation in acute myeloid leukemia: a new approach to anti-leukemia therapy. *Leukemia*, **13**, 1316–1324.

Liu, T. X., *et al*. (2000). Gene expression networks underlying retinoic acid-induced differentiation of acute promyelocytic leukemia cells. *Blood*, **96**, 1496–1504.

Mueller, E., *et al*. (2000). Effects of ligand activation of peroxisome proliferator-activated receptor gamma in human prostate cancer. *Proceedings of the National Academy of Sciences of the USA*, **97**, 10990–10995.

Samid, D., *et al*. (2000). Peroxisome proliferator-activated receptor gamma as a novel target in cancer therapy: binding and activation by an aromatic fatty acid with clinical antitumor activity. *Clinical Cancer Research*, **6**, 933–941.

Tontonoz, P., *et al*. (1997). Terminal differentiation of human liposarcoma cells induced by ligands for peroxisome proliferator-activated receptor gamma and the retinoid X receptor. *Proceedings of the National Academy of Sciences of the USA*, **94**, 237–241.

Wang, J., *et al*. (1999). Inhibitors of histone deacetylase relieve ETO-mediated repression and induce differentiation of AML1-ETO leukemia cells. *Cancer Research*, **59**, 2766–2769.

Chapter 90

Chemoprevention

Jaye L. Viner, Ellen Richmond, and Ernest Hawk

National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Julia Lawrence

Louisiana State University Health Science Center, New Orleans, LA, USA

CONTENTS

- Carcinogenesis: The Basis of Risk, the Target of Prevention
- Levels of Prevention
- Biomarkers
- Cohorts
- Agents
- Goals of Chemoprevention
- Clinical Trials
- The Future of Cancer Chemoprevention

CARCINOGENESIS: THE BASIS OF RISK, THE TARGET OF PREVENTION

Cancer is a late, nonobligate stage of carcinogenesis, the long process during which cumulative alterations in DNA structure and function are expressed within and across all biological levels (e.g. RNA, protein, organelle, cell, tissue, organ, organism), as shown in **Figure 1**. Endogenous mechanisms, such as carcinogen detoxification, DNA repair, selective apoptosis, immunological surveillance and stromal responses, limit progression to dysregulated proliferation, clonal anarchy and cancer. Chemoprevention mimics or augments these natural defences with pharmacological agents that prevent, reverse or retard carcinogenesis during what is typically a decades-long latency to cancer (Sporn, 1976). For practical purposes, chemoprevention addresses carcinogenesis up to the point of invasion across the basement membrane, whereas clinically evident cancer is the domain of chemotherapeutics. In fact, the boundaries between cancer prevention and therapeutics are eroding – as are traditional notions of health and disease – with the advent of new technologies that allow us to detect and eradicate incipient disease long before it becomes clinically evident or lethal.

LEVELS OF PREVENTION

Terminology for cancer prevention derives from the more mature prevention sciences of cardiovascular and infectious diseases. As applied in preventive oncology, primary prevention refers to the identification and avoidance of carcinogens in unscreened, ostensibly healthy people presumed at average risk of cancer. Examples of primary prevention include tobacco cessation to lower the risk for lung cancer, dietary or lifestyle modifications to reduce the incidence of cardiovascular events and vaccination to reduce the incidence of hepatitis B infection (and also downstream consequences such as cirrhosis and hepatocellular carcinoma). Secondary prevention involves the identification of cohorts at greater than average risk of disease through screening for subclinical pathology followed by interventions to arrest disease progression (e.g. mammography followed by surgical resection, colonoscopy with adenomectomy). Analogies to secondary prevention in other fields of medicine include screening for persons with elevated serum lipid profiles followed by dietary modification and/or cholesterol-lowering drugs to reduce the incidence of cardiovascular events, and human immunodeficiency virus (HIV) testing of pregnant women followed by AZT administration to reduce viral infection of at-risk neonates. Finally, tertiary prevention involves patients with diagnosed disease in whom applied interventions may improve palliation and reduce complications. Examples of tertiary prevention include bisphosphonates to reduce the risk of skeletal fractures in breast cancer and multiple myeloma, aspirin to reduce cardiovascular events in patients with coronary artery disease, and glucose control to limit end-organ complications of diabetes.

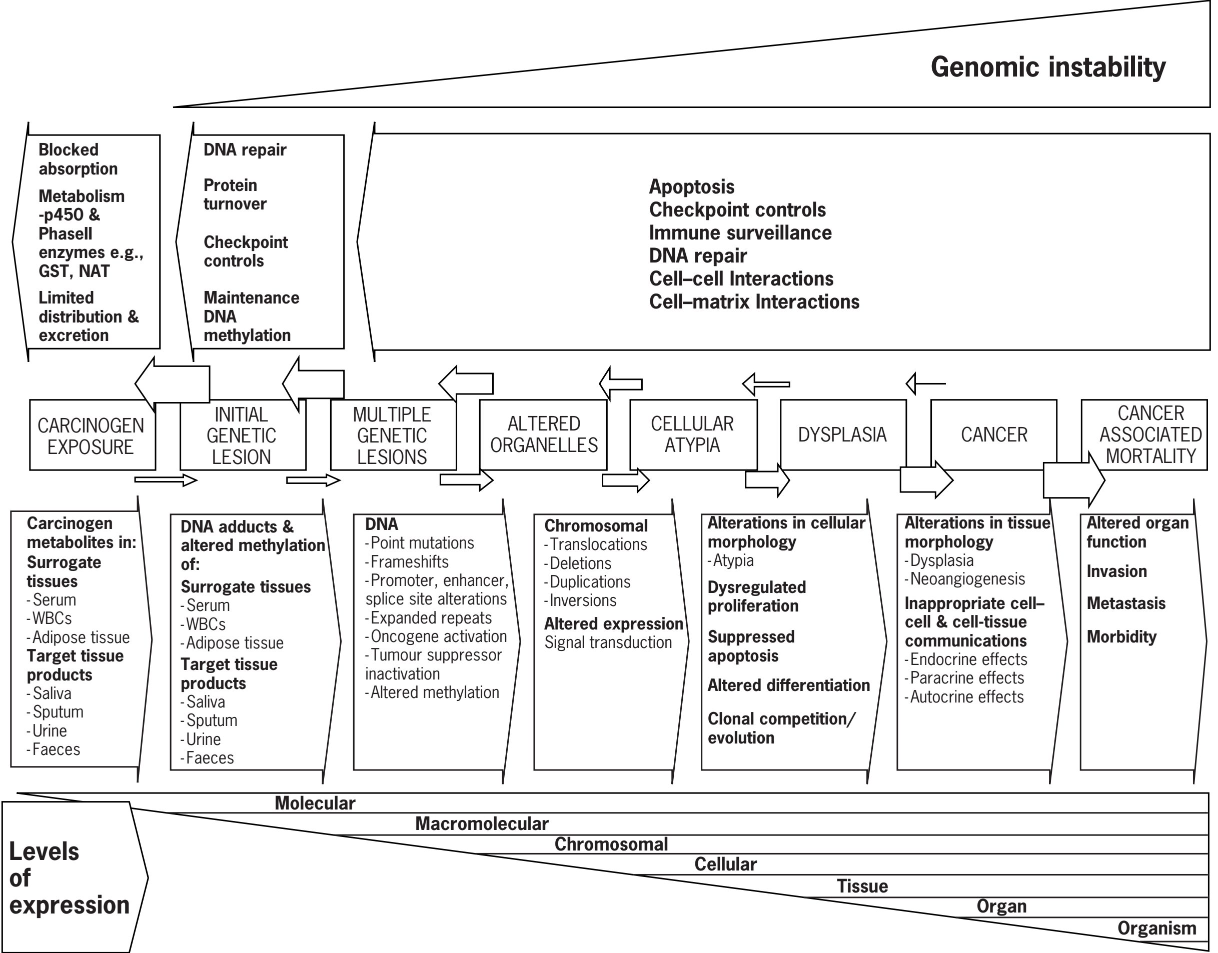


Figure 1 Mechanisms underlying carcinogenesis – a dynamic process. (Figure courtesy of Asad Umar, PhD, DVM.)

Context, rather than agent, dictates the grade of prevention. This context, however, is in flux owing to technological advances in biological imaging and molecular analyses that revolutionize our ability to visualize and characterize carcinogenesis (i.e. disease) at increasingly primitive levels of biological organisation. As a result, the demarcation between prevention and therapeutics is increasingly indistinct. Carcinogenesis is now recognized as a disease worthy of intervention, but it is a moving target.

BIOMARKERS

Progress in preventing cancer depends, in large measure, on the identification, development and validation of accurate, reliable and early biomarkers of carcinogenesis – events of structure or function on any level of biological organisation that convey insights into biology or pathology on a higher (and typically later) level, as shown in **Table 1**. Structural abnormalities are typically aberrations in number or component integrity within a biological level. Functional abnormalities may be quantitative or qualitative (i.e. inappropriate function relative to healthy biological need) and probably represent the more important criterion, but one that current technology measures with limited accuracy. As a result, functional dynamics are often deduced from structural abnormalities (e.g. loss of normal p53 activity inferred through overexpression of mutant p53). Biomarkers provide seminal insights into both the biological past and future. Molecular markers (e.g. DNA, RNA or protein) provide invaluable mechanistic insights into subclinical carcinogenesis, whereas tissue markers reflect imminent clinical pathology. Because of their proximity to clinical disease, tissue markers better justify the use of novel interventions in persons at risk of cancer.

Epithelial dysplasia (literally ‘disordered growth’) is architectural disarray at the tissue level spawned by clonal anarchy at lower biological levels, as described in **Table 2**. These neoplastic changes often occur adjacent to cancer foci, and are likely precursors to invasive disease. At a minimum, dysplasia reflects a biological milieu that is permissive for carcinogenesis. Although debate persists as to the absolute clinical risk associated with carcinogenesis identified at tissue (e.g. dysplasia), cellular (e.g. atypia) or even molecular (e.g. genetic mutation) levels, the potential reversibility of this process is the central premise of

Table 1 Hierarchy of biological aberrations (biomarkers) in carcinogenesis[a]

Level	Structural aberrations	Functional aberrations
DNA	Adducts Mutations Double minutes Translocations Hypo-/hypermethylation Hypo-/hyperploidy	Transcriptional over-/underexpression Transcriptional corruption Genomic instability
RNA	Inadequate copies Excessive copies Corrupted copies	Translational over-/underexpression Translational corruption
Protein	Inadequate Excessive Corrupted 1° structure 2° structure 3° structure 4° structure	Over-/ under-/ inactive Aberrant enzymatic function Impaired DNA repair Impaired DNA/ RNA replication Impaired post–translational modelling Impaired regulatory functions Impaired signal transduction Impaired maturation Impaired metabolism
Organelle	Aberrant nuclear morphology Enlarged nuclear–cytoplasmic ratio	Over-/ under-/ inactive
Cellular	Atypia	Over-/under-/inactive Functional immaturity Aberrant sensitivity to external growth stimuli
Tissue	Dysplasia Carcinoma *in situ* (CIS) Increased microvessel density Submucosal invasion Stromal reaction	Dysregulated population dynamics (increased proliferation and/or reduced apoptosis) Loss of intercellular adhesion Cell migration Neoangiogenesis Extracellular matrix degradation
Organ	Local invasion Metastasis	Functional compromise Mass effects

[a]Adapted from Hawk *et al.*, 2000.

Table 2 Histopathological components of epithelial atypia/dysplasia

Nuclear grade
Increased size
Irregular shape
Increased staining
Pleomorphism (variable size, shape and staining)
Nucleolar grade (glandular epithelia)
Increased size
Irregular shape
Increased staining
Cellular differentiation and maturation
Abnormal
Absent
Cellular mitoses
Increased number
Abnormal structure

chemoprevention. In organs that are more biopsy accessible (e.g. skin, bladder, oral mucosa, oesophagus, stomach, colorectum, cervix), preinvasive neoplasia [e.g. actinic keratosis (AK), dysplasia, adenomas] has become an accepted target for intervention. The current standard of care is to intervene against dysplastic lesions, with incisional biopsy for diagnosis and risk reduction, followed by serial surveillance, focal resection or even organ extirpation, depending on the risk of the lesion(s) and of the patients.

In addition to clinical response criteria (e.g. adenoma, dysplasia or plaque regression), valuable mechanistic insights into the genesis of preinvasive neoplasia are often gleaned from prevention trials that integrate evaluations of multiple response biomarkers at different levels of biological organisation (i.e. molecular level (*APC* mutations, genomic expression, cyclooxygenase activity) and cellular/tissue levels (proliferation, apoptosis, aberrant crypt foci (ACF) number/morphology, adenoma regression/prevention). Such concomitant assessments are critical, because reliance upon any biomarker in isolation is likely

to result in incomplete or flawed interpretations of pathobiological mechanisms underlying risk and response.

Biomarker Identification and Development

Candidate biomarkers may arise out of mechanistic or correlative insights derived from a variety of research settings – *in vitro*, *in vivo*, observational or clinical. Biomarkers may be categorized according to the intended application as (1) risk markers that estimate the probability of some later event in the neoplastic process, also referred to as exposure, susceptibility, diagnostic or prognostic biomarkers, or (2) response markers that measure response to an intervention, also referred to as drug, surgical, radiological or dietary effect, or outcome biomarkers (see **Figure 2**).

Although all cells within an epithelial sheet inherit a common susceptibility to neoplasia, variable exposure to genetic and epigenetic insults (i.e. carcinogens, hormones, ultraviolet radiation, etc.) results in nonuniform distribution and rates of neoplastic progression, particularly at more complex biological levels. Not surprisingly, this patchy distribution leads to critical sampling errors. In order to ensure measurement accuracy and reliability, the first step in developing a candidate biomarker involves standardization of tissue acquisition and laboratory assays for specimen analysis. Noninvasive imaging technologies, such as optical coherence tomography (OCT) and light-scattering spectroscopy (LSS), have improved clinical sampling, and laser-capture microdissection has reduced the noise-to-signal ratio in laboratory-based analyses. In addition, expression microarrays now enable concomitant assessment of multiple molecular endpoints, streamlining what has until recently entailed immensely time- and tissue-consuming evaluations of lesions within the field at risk.

Once standardized, candidate biomarkers are assessed to establish their natural history, and also their association with cancer risk cross-sectionally and over time. Often, owing to low rates of certain cancers, these interrogations are initiated through retrospective studies that compare individuals with cancer to matched control populations. Alternatively, the natural history of a promising marker may be derived from prospective studies of risk cohorts, or from the placebo arm of intervention studies that incorporate biomarkers as secondary measures of response. Risk biomarkers that correlate with the development of invasive cancer may be applied to (1) estimate risk for incident neoplasia, (2) monitor disease progression, (3) adjust interventions to optimize the therapeutic index (TI) (i.e. the benefit-to-risk ratio), (4) prioritize interventions based on mechanistic insights or (5) stratify individuals entering prevention trials to balance significant covariates.

In addition, risk markers merit further inspection as possible markers of response in prospective clinical

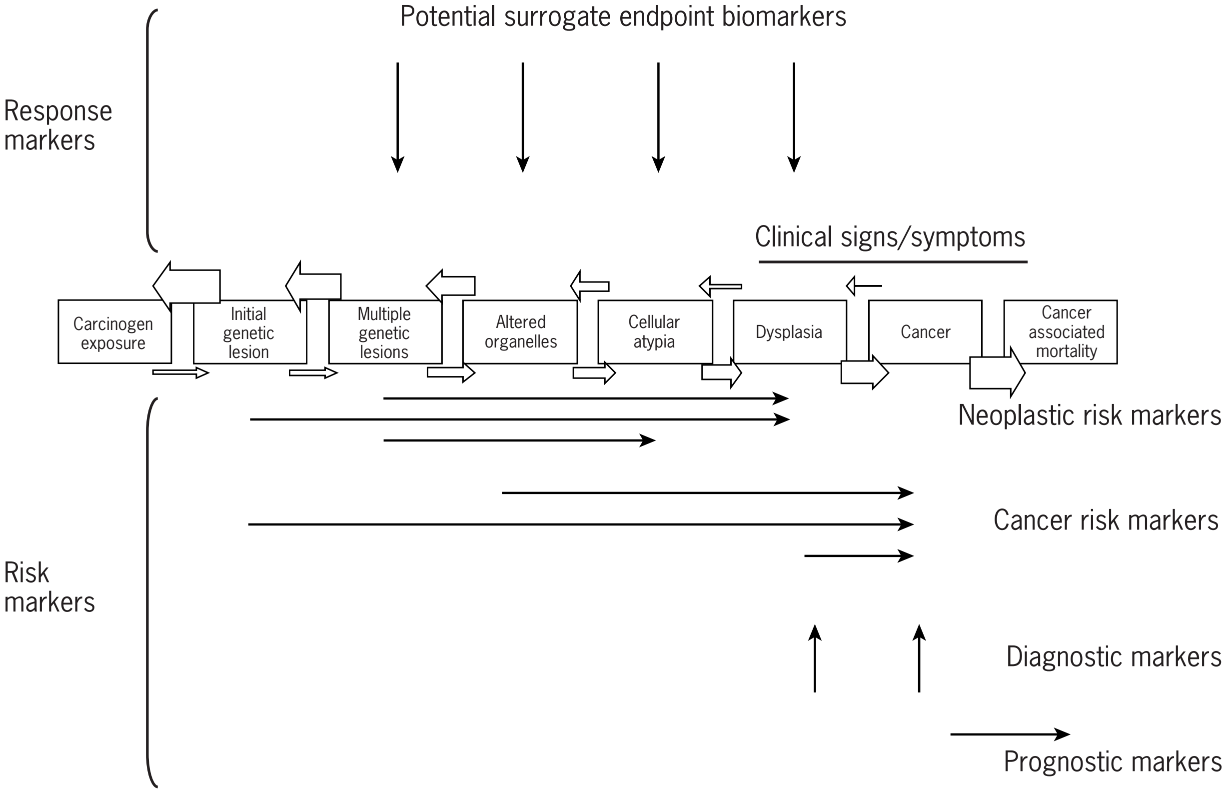


Figure 2 Application of biomarkers in carcinogenesis. (Adapted from Hawk *et al.*, 2000.)

Table 3 Surrogate endpoint biomarkers (SEB): ideal features[a]

Area	Features
Sampling	Standardized:
	Instruments
	Collection methods
	Preassay processing
	Sample transport
	Sample storage
Assay	Well-characterized neoplastic specimen as substrate
	Reproducible methods
	Limited inter-/intra-assay variance
	High sensitivity and specificity
	Inexpensive
	Adaptable to high-volume throughput
Tissue	Pathobiological context
	Differentially expressed in normal and neoplastic tissues
	Amenable to quantitative characterization
	Well-characterized natural history
	Easily measured with minimally invasive techniques
	Amenable to serial evaluations across time
	Validated as a marker of risk
	Response
	Modulable by preventive intervention of interest
	Specific for the effect of interest
	Short latency to invasive cancer
	Modulation correlates with reduction in invasive neoplasia
	Predictive validation
	Modulation predicts reduction in invasive neoplasia
	Relevant to a range of other interventions

[a]Adapted from Hawk *et al.*, 2000.

trials of investigational agents, as described in **Table 3**. Candidate response biomarkers with demonstrated biological relevance (often, established risk markers) are first evaluated in pilot studies to demonstrate their modulation following exposure to an intervention, then advanced into short- to intermediate-term trials that incorporate placebo controls. In these trials, the placebo arm confirms the marker's natural history and association with risk, whereas the active arm provides preliminary evidence of the agent's preventive efficacy. Once a response biomarker is established to be acceptably predictive of a later event of primary interest, such as cancer incidence or cancer-related mortality, it may be applied as a surrogate for the later event (i.e. surrogate endpoint biomarker (SEB)) with a reasonable measure of certainty (Temple, 1999; Hawk *et al.*, 2000). Technically, this credential is valid only with respect to the context in which it was tested; however, in a more liberal sense it might be generalized to other interventions and risk cohorts. The internal validity of the biomarker is based on its representativeness within a particular cohort. Relevance to a reference population and a broader range of interventions, in turn, affirms the biomarker's external validity. A conceptual algorithm for identifying, developing, and validating candidate SEBs is outlined in **Figure 3**.

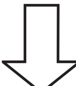
SEBs, Proven and Prospective

Validated SEBs, such as those presented in **Table 4**, facilitate preliminary efficacy testing of chemopreventive agents. This has been amply established with preinvasive lesions of the colorectum and breast. The 1993 National Polyp Study, for example, indirectly validated colorectal adenomas as SEBs for colorectal cancer in the context of surgical interventions, after showing up to 90% reductions in colorectal cancer incidence in subjects undergoing regular surveillance for adenomas, as compared with three historical control groups (Winawer *et al.*, 1993). Observational and experimental studies have also shown significant reductions in colorectal cancer mortality and, in some cases, incidence, in the context of colorectal cancer screening by stool blood-based tests or flexible sigmoidoscopy combined with colonoscopic adenomectomy. Finally, adenomas have achieved a measure of validation as SEBs in the context of medical interventions as well – over 20 observational studies confirm that nonsteroidal anti-inflammatory drugs (NSAIDs) reduce the risk of colorectal adenomas, carcinomas and cancer-associated mortality (Smalley and DuBois, 1997).

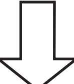
In breast cancer, preinvasive neoplastic lesions gained credibility as SEBs based on results from the National Surgical Adjuvant Breast and Bowel Project (NSABP) P-1 and B-24 trials (Fisher *et al.*, 1998, 1999). In these trials, women taking tamoxifen experienced reductions in ductal carcinoma *in situ* (DCIS) and/or invasive breast cancer. Proportional decreases in DCIS and invasive cancer provided a strong rationale for considering DCIS as an SEB for breast cancer. In response to data derived from the NSABP P-1 trial, the US Food and Drug Administration (FDA) approved the use of tamoxifen for risk reduction (i.e. prevention) in women at high risk of breast cancer. Although tamoxifen is the first agent specifically approved for cancer risk reduction, two other agents, intravesical Bacillus Calmette-Guerin (BCG) and topical fluorouracil, were approved years earlier to treat preinvasive neoplasia of the bladder and skin, respectively. Each of these approvals established dysplasia as an important target for medical interventions intended to control disease and ultimately prevent cancer. Given the brisk pace of biomarker research, dysplasia will likely achieve validation as a definitive, although probabilistic, SEB for cancer incidence in many other target organs over the next few years.

Select feasible/rational biomarkers
Based on mechanisms of disease and agent's hypothesized mechanism of action

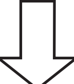
Standardize sampling
Establish target (neoplasia vs at-risk field), instruments and technique via standardized protocol



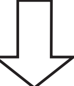
Standardize laboratory assay
Establish handling and processing methods for specimens including reagents, conditions and instruments; determine accuracy and reliability



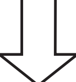
Establish status as cross-sectional risk marker for invasive neoplasia
Correlate cross-sectional risk within a population



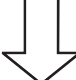
Establish status as a prospective risk marker for invasive neoplasia
Evaluate the natural history of the biomarker over times; correlate the marker with prospective risk either in a prospective study or via the placebo arm of a controlled intervention trial



Establish status as a modulable marker
Assess marker modulation between baseline and outcome in an intervention study (versus historical data on natural history or a concomitant placebo control)



Establish as a surrogate endpoint with one agent and in one model system
Correlate biomarker modulation at an intermediate time point with effects on clinical outcomes of interest (a clinical intervention, invasive cancer incidence, or another event)



Establish as a surrogate endpoint across agents and/or models
Test the predictiveness of biomarker modulation with different agents and/or in other model systems (against incident cancer or other clinical outcomes of interest); test the surrogate against combinations of agents

Figure 3 Development of tissue-based biomarkers as surrogate endpoints: a conceptual pathway. (Adapted from Hawk *et al.*, 2000.)

Table 4 Preinvasive neoplasia as markers of risk and surrogate efficacy[a]

Biomarker	Validated risk marker	Validated surrogate endpoint for surgical interventions[b]	Validated surrogate endpoint with medical interventions[c]
Ductal carcinoma *in situ* (DCIS)	✓	✓	Tamoxifen
Atypical ductal hyperplasia (ADH)	✓		Tamoxifen
Cervical intraepithelial neoplasia (CIN)	✓	✓	
Oral leukcoplakia	✓		
Oesophageal dysplasia (high grade)	✓	✓	
Colonic adenoma	✓	✓	Celecoxib
Bladder dysplasia	✓	✓	BCG, valrubicin
Prostatic intraepithelial neoplasia (PIN)	✓		
Actinic keratosis (AK)	✓	✓	5-Fluorouracil, diclofenac sodium, Aminolaevulinic acid with photodynamic therapy, masoprocol

[a]Adapted from Hawk *et al.*, 2000.
[b]Most of these dysplastic lesions are surgically removed upon identification, although only a few have been validated to demonstrate that doing so reduces the risk for subsequent cancers or cancer-associated mortality.
[c]Validation of these lesions as predictors of reduced cancer incidence is indirect in most instances, but the predictive association has been strong enough to justify approval for the treatment of a preinvasive neoplastic condition which may, therefore, reduce the risk of subsequent cancer incidence.

Table 5 Cohorts with an elevated cancer risk

Risk factor	Example
Endogenous	
Germ-line mutation	*APC* mutation (colorectal)
Genetic polymorphism	*GSTM1* −/− genotype (lung)
Family history	Two first-degree relatives with cancer and one with a precancer (colorectal, breast)
Exogenous	
Carcinogen exposure	
Chemical	Industrial benzene exposure (bladder)
Radiation	Therapeutic radiation exposure (breast)
Virus	Chronic hepatitis B infection (liver)
Somatic mutations	*p16* mutation, deletion or methylation (oesophagus)
Prevalent preinvasive neoplasia (dysplasia)	Actinic keratosis (skin) or adenoma (colorectal)
Personal history of past preinvasive neoplasia	Excised actinic keratosis (skin) or adenoma (colorectal)
Personal history of past cancer	Cancer survivor

Indeed, an expert advisory committee was recently convened under the auspices of the American Association for Cancer Research (AACR) to evaluate intraepithelial neoplasia as a target worthy of intervention.

The use of biomarkers to open up the ‘black box’ between early and late events of interest is not unique to cancer chemoprevention. Indeed, intermediate biomarkers are commonly applied in other areas of medical research and the management of chronic disease (e.g. diabetes, hypertension, HIV and cancer) to reduce morbidity and mortality through early diagnosis, and also to assess the necessity for, and responsiveness to, interventions applied with preventive or therapeutic intent. For example, serum lipid profiles and hypertension are biomarkers of risk and response that have expedited advances in cardiovascular disease research, and have had a profound impact on public health (Temple, 1999; Hawk *et al.*, 2000).

COHORTS

As noted above, carcinogenesis is fundamentally an unstable pathological process. However, even when it is expressed at the tissue level (i.e. as dysplasia), most disease regresses or remains stable, never advancing to invasive cancer. Certain populations and individuals are at increased risk of cancer due to carriage of highly penetrant susceptibility genes (e.g. *APC*, *MLH-1*, *MSH-2*, *BRCA1/2*), or environmental exposures (i.e. tobacco, radiation, exogenous hormones and hepatitis B, hepatitis C, Epstein–Barr and human papilloma viruses), as described in **Table 5**. These individuals not only have a greater likelihood of developing cancer than persons at average risk, but typically do so at an accelerated rate. As a result, trials conducted in high-risk populations may be sufficiently powered with relatively small cohorts followed for short

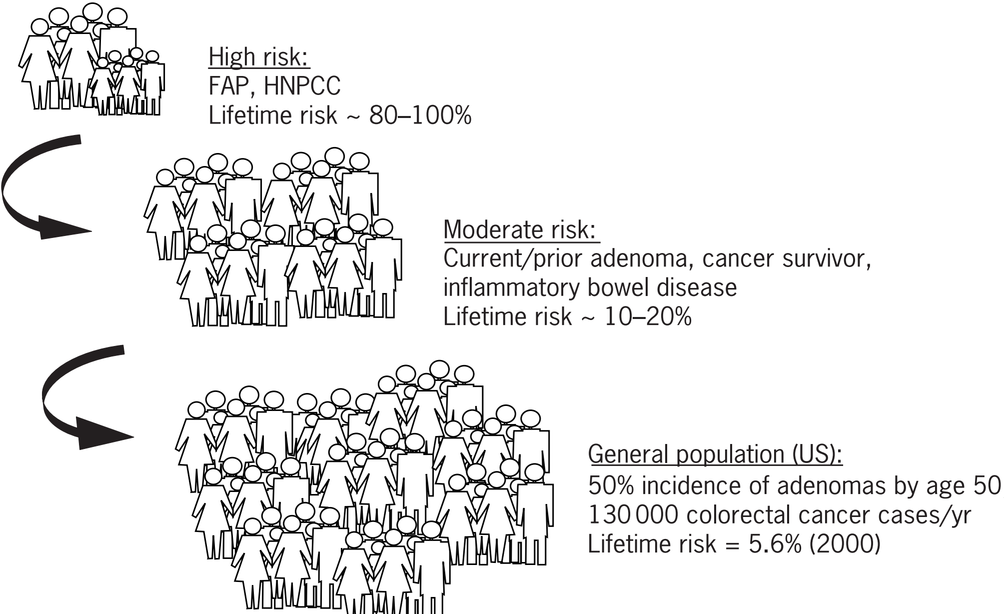


Figure 4 Agent testing along a risk gradient: colorectal neoplasia as a model.

periods of time. Agents confirmed as effective and safe in this context may then be tested for generalizability in progressively lower risk cohorts which collectively account for the overwhelming burden of cancer cases, provided that the diseases share a common aetiology (see **Figure 4**). Such strategies accelerate clinical agent development and the time to iterative testing in progressively lower risk cohorts, thereby conserving precious time and resources. Moreover, concerted testing along this risk gradient provides an ethical advantage – offsetting potential research risks with potential benefits for patients with the most to gain.

Table 6 Strategies to improve the therapeutic index (TI) (i.e. benefit:risk ratio) of chemopreventive agents

Reduce systemic exposure
Pharmacodynamic titration
Reduce dose
Reduce frequency
Reduce duration
Regional administration
Prophylax against or treat unintended toxicities
Improve mechanistic targeting
Combinatorial approaches/regimens

AGENTS

Toxicity and Safety Standards for Chemoprevention

In chemotherapeutics, iatrogenically induced toxicities may be tolerated in pursuit of cure. Cancer chemoprevention, however, may mandate long-term measures to prevent advanced disease that may not even develop. As a result, agents approved for preventive applications typically exceed safety and efficacy standards set for therapeutics. In an effort to improve the TI of promising chemopreventive agents, several strategies have been explored, including (1) enhanced mechanistic specificity (e.g. COX-2 selective vs COX nonselective inhibition), (2) agent combinations with additive or synergistic efficacy (e.g. COX + ODC inhibition; SERM + retinoid), (3) regional administration (i.e. topical formulations to maximize target:systemic exposure), (4) prophylaxis against worrisome toxicities (e.g. coadministration of antiulcer medication+NSAIDs) and (5) dose titration to biological efficacy, rather than maximally tolerated dose (see **Table 6**).

Enhanced mechanistic specificity has vast potential to improve the TI of promising agents, as demonstrated by the development of COX-2 selective inhibitors (derivatives of NSAIDs) for chemoprevention. Nonselective NSAIDs are a structurally diverse class of agents that have been shown to reduce cellular proliferation, slow cell cycle progression and stimulate apoptosis. These effects appear to be significantly linked to the inhibition of COX activity, which facilitates the conversion of arachidonic acid into bioactive lipids (e.g. prostaglandins, thromboxanes and prostacyclins) that are considered to be important promoters of carcinogenesis. The chemopreventive potential of NSAIDs as a class was established in a compelling body of

epidemiological, preclinical and clinical data accrued over the last two decades. However, safety concerns – largely due to gastrointestinal toxicities attributable to COX-1 inhibition – curbed enthusiasm for the development of NSAIDs as chemopreventives. COX-2 selective inhibitors have a markedly improved TI, as compared with nonselective COX inhibitors. Three complementary lines of research established COX-2 as a specific target for chemoprevention, based on the following: (1) COX inhibitors and *COX-2* gene deletions inhibit intestinal carcinogenesis in animal models; (2) nonselective COX inhibitors reduce the incidence of colorectal adenomas, cancer and cancer-associated mortality in human observational studies; and (3) COX inhibitors regress precancerous lesions (e.g. ACF and adenomas) in genetic and sporadic colorectal neoplasia cohorts. In 1999, the COX-2 selective inhibitor celecoxib was shown to reduce the number of colorectal adenomas in persons with the high-risk genetic condition known as familial adenomatous polyposis (FAP). This prompted the FDA to approve the provisional use of celecoxib in FAP patients, in conjunction with standard surveillance and surgical prophylaxis. Prospects for chemopreventive applications of NSAIDs have expanded markedly with the discovery of well-tolerated COX-2 selective derivatives such as celecoxib (Pharmacia) and rofecoxib (Merck). COX-2 inhibitors are now being aggressively developed as chemopreventives against an array of epithelial tumours, particularly those of the colorectum, oesophagus, skin and bladder.

Coadministration of several agents with chemopreventive activity may improve the TI by allowing downward dose adjustments of one or more of the applied agent(s). Additive/synergistic activity has been demonstrated with combinations of agents with different mechanisms of action, such as difluoromethylornithine (DFMO) or an epidermal growth factor receptor (EGFR) inhibitor plus NSAIDs in rat colonic tumours and retinoids plus anti-oestrogens in rat mammary cancer models. Combinatorial approaches are currently being tested in phase II clinical chemoprevention trials (e.g. DFMO plus sulindac in persons at risk of colorectal cancer, and fenretinide plus tamoxifen in women at risk of breast cancer).

Regional delivery systems are another promising area for agent development. Predictably, the most successful application of this strategy has been in skin cancer, where three topical agents have been approved by the FDA for regression of the nonmelanoma skin cancer precursor actinic keratosis (AK), i.e. 5-fluorouracil (5-FU), aminolaevulinic acid hydrochloride (Levulan Kerastick™), with photodynamic therapy, and sodium diclofenac (Solaraze™). Other effective or promising modes of direct delivery to target organs include aerosols to the lung, enemas or suppositories to the colorectum and instillations into the bladder.

Early neoplasia is characteristically more responsive and accessible to delivered agents than later stage disease, and therefore drugs may be titrated to a biologically active dose with negligible toxic potential. Such dose de-escalation models, although appropriate for population-based chemopreventive applications, may have less relevance to high-risk cohorts with limited medical prospects and viable options. Individuals with highly penetrant germ-line disease (FAP, hereditary non-polyposis colon cancer (HNPCC), basal cell naevus syndrome (BCNS), breast/ovarian syndromes, etc.), for example, may abide toxicity in order to improve the likelihood of disease control and/or organ preservation, increase surveillance intervals and latency to disease or alleviate concerns over developing cancer. Pharmacogenomics, the study of genetic polymorphisms that modulate drug metabolism, is one of many emerging fields that may advance cancer prevention by customizing agent and dose selection to individual patients, thereby enhancing the TI of chemopreventive agents.

Agent Development

Candidate chemopreventives are identified through scientific, medical and epidemiological literature review, preclinical drug screens and typically developed and tested through collaborations between academia, governmental agencies and the pharmaceutical industry. Chemoprevention research is intensely translational, and advances through bidirectional hypothesis generation and testing that swiftly applies laboratory findings in the clinic and vice versa. Lead compounds have been variously identified through (1) preclinical efficacy testing of promising agents with distinct molecular targets, (2) epidemiological data on drugs developed for other applications (e.g. oncological, endocrinological, rheumatological, gastroenterological) and (3) epidemiological data on dietary components. Once identified, mechanistic considerations prioritize agents for advancement into human trials. Although traditional testing is still required, this approach expedites agent development and conserves precious time and resources. Another strategy to shorten developmental timelines may be forthcoming. Owing to the polychronotopic (literally 'many times and places') nature of carcinogenesis, early pathological lesions (e.g. AKs, ACF) are a common incidental finding in persons diagnosed with more advanced disease (e.g. skin cancer or colorectal cancer). This phenomenon is consistent with a smouldering process (carcinogenesis) in which genetic and epigenetic events foster different rates and manifestations of neoplasia within the field at risk. Agent identification may be expedited by nesting 'preventive' clinical endpoints within compatible 'therapeutic' trials. This extraordinarily efficient strategy has potential to hasten the evaluation of promising agents with a chemopreventive potential based on their mechanism(s) of action, route of administration and safety profiles.

Preclinical Testing

Positive observational data and serendipitous results from nononcological clinical trials often prompt preclinical efficacy testing (e.g. in tissue culture and animal models) to define the mechanisms whereby agents inhibit carcinogenesis (e.g. inhibition of Ras farnesylation, EGFR, ornithine decarboxylase, matrix metalloproteinase, steroid aromatase and oncoproteins, or induction of antimutagenesis, apoptosis or antioxidation). Proven reductions in tumour incidence, multiplicity and/or overall burden in carcinogen-induced, transgenic or spontaneous animal cancer models may then justify human testing. Some of the most commonly used carcinogen-induced tumour models are azoxymethane (AOM)-induced rat colon tumours and ACF, dimethylbenz[*a*]anthracene (DMBA-) and *N*′-methyl-*N*-nitrosourea (MNU)-induced rat mammary tumours, benzo[*a*]pyrene (B(a)P)-induced rat lung tumours and *N*-butyl-*N*-(4-hydroxybutyl)nitrosamine (OH-BBN)-induced mouse urinary bladder tumours. Carcinogens and/or genetic lesions that reliably and reproducibly induce well-characterized tumours facilitate crude assessments of agent efficacy. Ideally, *in vivo* models approximate multistage genetic, cellular and histopathological changes that characterize human carcinogenesis; however, most models fall short of this ideal. In the case of carcinogen-induced tumour models, most commonly used toxins have no known role in human disease, severely limiting generalization to human carcinogenesis. Transgenic and gene knockout (i.e. gene-deleted) mice may mimic human disease more faithfully, although critical data are still falling into place. Among these, perhaps the most effectively applied model for chemoprevention is the multiple intestinal neoplasia (Min) mouse, that carries an *apc* mutation that predisposes them to intestinal adenomas. Min mice effectively simulate the clinical condition associated with *APC* mutations in human disease, FAP. The Min mouse model and similar ones based on mutations in the *APC* gene furnish dynamic systems in which modulation of colorectal adenoma formation by chemopreventive compounds may be tested. Another exciting model is the HPV-infected (K14-HPV16 heterozygote), oestradiol-treated mouse model, in which cervical intraepithelial neoplasia-like (CIN) lesions progress to squamous carcinoma of the cervix – potentially providing insights into another epithelial cancer. Technological advances arising out of ambitious international initiatives to develop mouse models for human cancers may generate and characterize robust new models for human carcinogenesis (http://www.nci.nih.gov/dcb/odhome.htm#MOUSE).

Several classes of pleiotropic agents (i.e. that target multiple signalling pathways) are under development for chemoprevention, including lipoxygenase (LOX) inhibitors, COX inhibitors, farnesyl-transferase inhibitors and corticosteroids. Agents such as antioestrogens (e.g. SERMs), anti-inflammatory drugs (e.g. aspirin, steroids and selective COX-2 inhibitors), polyamine synthesis inhibitors (e.g. DFMO), retinoid receptor selective ligands, agents that bind peroxisome proliferator-activated receptors (PPARs), and non-cytotoxic cancer chemotherapeutics (e.g. flutamide) are being comprehensively tested in clinical trials. Other promising prevention targets in the preliminary stages of investigation include cell cyclins, bile acids, telomerase and angiogenesis. Emerging structure–activity technologies may expedite the development of these and other promising leads.

Phased Clinical Testing of Agents for Chemoprevention

The efficacy and toxicity of candidate preventive agents are typically assessed in a series of clinical trials referred to as phase I, II or III studies, as summarized in **Table 7**.

Table 7 Common features of chemoprevention trials used in agent development

Phase	Agent dosing	Duration	Sample size and allocation	Control group	Goals
I	Escalation	Weeks–months	<25; nonrandomized or randomized	Occasionally	Pharmacokinetics; dose-finding based on short-term, mild to moderate toxicity
IIa	De-escalation	Months	<50; nonrandomized	Never	Dose-finding based on reliable biomarker modulation
IIb	Stable	Months–year	<100–200; randomized	Standard care[a]	Biomarker modulation (e.g. dysplasia regression) vs standard care[a]
III	Stable	Years	100–1000; randomized	Standard care[a]	Definitive efficacy to complement or replace standard care (e.g. reduce dysplasia/ cancer incidence)
IV	Stable	Unspecified	General marketing population	N/A	Long-term safety in target population

[a]In cancer chemoprevention, placebo may represent the standard of care.

Phase I trials quickly evaluate agent safety at several doses in a small number of subjects to establish an acceptable dose range. Phase II trials are larger and designed to assess preliminarily agent efficacy with regard to a specific precancerous condition (e.g. skin, breast or colorectal dysplasia). Compounds that have satisfied regulatory safety standards for other indications, such as NSAIDs for analgesia, obviate the need for phase I testing and may be ushered directly into phase II trials (referred to as phase IIa if no placebo group is included and phase IIb when one is). Phase III chemoprevention trials are usually randomized and placebo-controlled to test preventive agents or agent combinations against the standard of care for a specific condition. These studies typically involve hundreds to thousands of subjects observed for years, whereas phase II trials may be adequately powered with as few as 100 subjects observed over several months. Parenthetically, phase IV studies are postmarketing surveys conducted to assess agent side effects in the general population (Kelloff *et al*., 1995). (See the chapter *Translational Research*.)

GOALS OF CHEMOPREVENTION

Cardiovascular prevention research has shown that cholesterol-reducing statins effect 10–25% reductions in low-density lipoprotein (LDL) levels, yet account for 30–40% reductions in several major cardiovascular outcomes, clearly demonstrating the enormous public health implications of even small preventive effects. In terms of cancer prevention, Gaziano *et al*. estimated that an intervention that reduced all epithelial cancers by as little as 15% could conceivably prevent 73 000 cancer deaths annually in the USA, through reductions in cancer incidence, morbidity and mortality (Gaziano *et al*., 1996). Quality of life, latency to disease progression, frequency of screening and surveillance and organ preservation may also be improved by preventive interventions (see **Table 8**). In addition, although still controversial, diseases of ageing may share common initiating or promoting mechanisms, and certain classes of agents may prove beneficial in several chronic diseases (e.g. NSAIDs for arthritis/colorectal cancer, selective oestrogen receptor modulators (SERMs) such as tamoxifen or raloxifene for osteoporosis/breast cancer and statins for cardiovascular disease/colorectal cancer). As a specific example, NSAIDs exert analgesic, anti-inflammatory and cardiovascular effects; they may also inhibit cancer and neurodegenerative diseases, although these effects have not yet been confirmed in randomized trials.

Table 8 Goals of cancer chemoprevention

Short-range:
- Modulate validated mechanistic biomarkers
- Regress prevalent preinvasive neoplastic lesions
- Suppress recurrent preinvasive neoplastic lesions following definitive surgical resection
- Prevent incident preinvasive neoplastic lesions

Long-range:

Secondary prevention (i.e. screened cohorts at elevated risk)
- Improve efficacy of surveillance and its associated interventions
  - Address missed/complicated preinvasive neoplastic lesions
  - Preserve organs
- Reduce intensity of surveillance
- Improve safety of surveillance
- Reduce cancer risk of one (or multiple) organs
  - Reduce incidence
  - Delay onset

Primary prevention (i.e. general population)
- Provide an alternative/complement to broad-based cancer screening
- Reduce risk of several chronic diseases (e.g. cancer and cardiovascular disease)

CLINICAL TRIALS

Compelling mechanistic, observational and experimental evidence supports the development of several compounds for chemoprevention; however, many questions remain. As an example, for most chemopreventive applications, we do not know which agents are most effective or, more importantly, the optimal dose, duration, frequency or route of delivery. Nevertheless, the promise and feasibility of chemoprevention in persons at high risk of cancer has been amply demonstrated with tamoxifen for breast cancer and celecoxib for colorectal neoplasia (Fisher *et al*., 1998; Steinbach *et al*., 2000). These two trials provide a conceptual framework and set clinical standards for future studies. The following selected trials illustrate the accomplishments, challenges, and promise of clinical chemoprevention (summarized in **Tables 9–11**).

Upper Aerodigestive Tract Cancer

Head and neck cancer refers to a variety of cancers that develop in the mouth, throat, sinuses and nasal cavity. These tumours are most often associated with heavy smoking and/or alcohol consumption. The mucosa of the upper aerodigestive tract is readily visualized and biopsied, greatly facilitating identification of premalignant lesions, evaluation of agent efficacy and serial tissue sampling. Leucoplakia (literally 'white plaque') is a dysplastic precursor to oral cancer. With malignant transformation rates up to 36% depending on the degree of dysplasia, oral leucoplakia provides a clinical target for cancer chemoprevention trials. Retinoids are the most extensively studied chemopreventive agents for regression of oral leucoplakia and treatment of the field at risk. The first positive chemoprevention trial proved that 13-*cis*-retinoic acid (13*c*RA) administered to subjects over 3 months

Table 9 Chemoprevention clinical trials demonstrating prevention efficacy

Organ system	Endpoints	Cohort	Sample size	Control[a]	Intervention × Duration	Author, year
Upper aerodigestive tract (UADT)	Second primary incidence	Prior squamous cell carcinoma of the head and neck	103	✓	Isotretinoin (low dose) × 12 mo	Hong, 1990
	Leucoplakia incidence	Symptomatic oral mucosal hyperplasia/dysplasia	70 (59)[b]		Isotretinoin responders (× 3 mos) randomized to low-dose isotretinoin[e] vs. β-carotene × 9 mo	Lippman, 1993
	Leucoplakia incidence	Symptomatic oral mucosal hyperplasia/dysplasia	43		α-Tocopherol × 24 mo	Benner, 1993
	Leucoplakia regression	Prevalent oral leucoplakia	44	✓	Isotretinoin × 3 mo	Hong, 1986
Lung	Second primary incidence	Postresection of stage I non-small cell cancer	307	✓	Retinol palmitate × 12 mo	Pastorino, 1993
	Cancer incidence	History of skin cancer	974	✓	Selenium × 4.5 yr	Clark, 1998
	Metaplasia incidence	Current smokers	70		Etretinate × 6 mo	Gouveia, 1982
	Metaplasia incidence	Current smokers	40		Etretinate × 6 mo	Misset, 1986
	Metaplasia incidence	Cytologically abnormal sputum	26		13-*cis*-Retinoic acid × 6 mo	Saccomanno, 1982
	Cytological improvement	Current smokers with metaplasia	73	✓	Folate + hydroxycobalamin × 4 mo	Heimburger, 1988
	Cytological improvement	Heavy smokers with metaplasia	52		Aromatic retinoid (Et_1) × 6 mo	Mathe, 1982
Stomach	Precancerous lesion regression	Narino, Columbia residents with precancerous lesions	852	✓	Antibiotics[e], ascorbic acid[e], β-carotene[e] × 6 yr	Correa, 2000
Colon	Cancer incidence[c]	History of skin cancer	1312	✓	Selenium × 4.5 yr	Clark, 1998
	Adenoma recurrence	Resected sporadic adenoma	913	✓	Calcium carbonate × 4 yr	Baron, 1999
	Adenoma number and size	Familial adenomatous polyposis	22	✓	Sulindac × 9 mo	Giardello, 1993
	Adenoma number and size	Familial adenomatous polyposis	77	✓	Celecoxib × 6 mo	Steinbach, 2000
Liver	Cancer incidence	Hepatitis C-related cirrhosis	90	✓	Interferon-α × 12 wk	Nishiguchi, 1995
	Cancer incidence	Hepatitis B infection	101	✓	Interferon × 12 wk	Lin, 1999

Breast	Cancer incidence	Women at increased risk of breast cancer (Gail model)	13 388	✓	Tamoxifen × 5 yr	Fisher, 1998
Prostate	Cancer incidence[c]	Male smokers	29 133	✓	3 arms: α-Tocopherol[e] β-Carotene α-Tocopherol + β-carotene × 5–8 yr	ATBC, 1994
	Cancer incidence[c]	History of skin cancer	974	✓	Selenium × 4.5 yr	Clark, 1998
Bladder	Cancer incidence	Superficial transitional cell carcinoma following BCG	65		Megadose vitamins[e] vs RDA multivitamins × 10 mo	Lamm, 1994
	Tumour progression	Superficial transitional cell carcinoma	660	✓	Intravesical and percutaneous BCG – induction and maintenance over 3 yr	Lamm, 2000
Cervix	CIN[d] regression	CIN[d] II and III	301	✓	Cervical caps with β-*trans*-RA – periodically over 6 mo	Meyskens, 1994
Skin	AK regression	Moderate to severe actinic keratoses	48	✓	Topical DFMO × 6 mo	Alberts, 2000
	Time to cancer incidence	Prior actinic keratoses and/or skin cancers	2297	✓	Retinol × $\leq$5 yr	Moon, 1997
All sites	Cancer incidence and mortality	Linxian, China residents at high risk of gastro-oesophageal cancers	29 584	✓	4 arms: Retinol, zinc Riboflavin, niacin Vitamin C, molybdenum β-Carotene, vitamin E, selenium[e] × 1–5 yr	Blot, 1993

[a]Controlled trials include those with a placebo arm, no treatment control or standard of care.
[b]Responders were randomized.
[c]A secondary endpoint.
[d]CIN: cervical intraepithelial neoplasia.
[e]Agent with a significant preventive effect in a multi–arm trial.

Table 10 Cancer chemoprevention clinical trials demonstrating a null effect

System	Primary endpoint	Cohort	Size	Control[a]	Intervention × duration	Author, year
Upper aerodigestive tract	Second primary incidence (head and neck, or lung cancer) and overall survival	Past head and neck or lung cancer	2592	✓	Retinyl palmitate, *N*-acetylcysteine, both × 2 yr	van Zandwijk, 2000
	Second primary incidence	Resected head and neck cancer	316	✓	Etretinate × 24 mo	Bolla, 1994
Lung	Cancer incidence	Male US physicians	22 071	✓	Aspirin + β-carotene Aspirin β-Carotene × 12 yr	Hennekens, 1996
	Metaplasia incidence	Chronic smokers with prevalent metaplasia or dysplasia	152	✓	Isotretinoin × 6 mo	Lee, 1994
	Cytological atypia incidence	Former asbestos workers	755	✓	β-Carotene + retinol	McLarty, 1995
	Cytological atypia incidence	Current smokers	150	✓	Etretinate × 6 mo	Arnold, 1992
Oesophagus	Oesophageal/gastric cancer incidence	Residents of Linxian, China with oesophageal dysplasia	3318	✓	Supplementation with 14 vitamins and 12 minerals × 5.25 yr	Li, 1993
	Precancerous lesion incidence	Residents of Huixian, China	610	✓	Retinol, riboflavin, zinc × 13.5 mo	Munoz, 1985
	Chronic oesophagitis or oral leukoplakia regression	Uzbekistan males with oral leucoplakia and/or chronic oesophagitis	532	✓	Riboflavin Vitamins A and E + riboflavin β-carotene Vitamins A and E + β-carotene × 20 mo	Zaridze, 1993
Colon	Cancer incidence	US male physicians	22 071	✓	Low-dose aspirin vs β-carotene × 5 yr	Gann, 1993
	Adenoma incidence	Recent adenoma history	864	✓	β-Carotene, vitamins C and E vs β-carotene and vitamins C and E × 1 yr	Greenberg, 1994
	Adenoma incidence	Recent adenoma history	1429	✓	Wheat bran fibre × 5 yr	Alberts, 2000
	Adenoma incidence	Recent adenoma history	2079	✓	Intensive for low-fat, high-fibre, fruit and vegetable diet × 4 yr	Schatzkin, 2000
	Adenoma incidence	Recent adenoma history	665	✓	Calcium, fibre × 3 yr	Faivre, 2000
Liver	Cancer incidence	Hepatitis C-related cirrhosis	99	✓	Interferon-α2b × 48 wk	Valla, 1999
Breast	Cancer incidence	Stage I cancer or DCIS	2972	✓	Fenretinide × 5 yr	Veronesi, 1995
	Cancer incidence	Family history of breast cancer	2471	✓	Tamoxifen × 5.8 yr	Powles, 1998
Cervix	CIN I and II regression	Prevalent koilocytic atypia, CIN I or II	331	✓	Folic acid × 6 mo	Childers, 1995
Skin	Cancer incidence	History of nonmelanoma skin cancer	1805	✓	β-carotene × 5 yr	Greenberg, 1990
	Cancer incidence	Prior basal cell cancers	981	✓	Isotretinoin × 3 yr	Tangrea, 1993
	Cancer incidence	Prior basal or squamous cell cancers	525	✓	Retinol, isotretinoin × 3 yr	Levine, 1997

[a]Controlled trials include those with a placebo arm, no treatment control, or standard of care.

Table 11 Chemoprevention clinical trials demonstrating a harmful effect

System	Endpoint	Cohort	Size	Control	Intervention × duration	Author, year
Lung	Cancer incidence	Male smokers	29 133	✓	α-Tocopherol + β-carotene × 5–8 yr	ATBC, 1994
	Cancer incidence	Smokers, former smokers, and workers exposed to asbestos	18 314	✓	β-Carotene and retinol × 5 yr	Omenn, 1996
Colon	Adenoma incidence	Sporadic adenoma history	665	✓	Fibre × 3 yr	Faivre, 2000

dramatically transforms leucoplakia lesions into normal-appearing oral mucosa (Hong *et al*., 1986). However, these responses were neither durable nor complete. Although several retinoid doses, schedules and formulations have been evaluated, the TI remains dose limiting. Other candidate chemopreventives of considerable interest for this condition include NSAIDs, EGFR inhibitors and a pro-apoptotic adenovirus (ONYX-015).

Head and neck cancer is associated with a high rate of new primary cancers (up to 40%), many of which develop within apparently normal mucosa in the vicinity of the original cancer. It is postulated that chemopreventive interventions may reduce the rate of these 'second primary' cancers. Placebo-controlled secondary prevention trials of retinoids in patients with completely excised primary squamous cell carcinoma of the head and neck have shown reduced rates of second primaries in the treatment group, although tumour recurrence rates remain unchanged. The EUROSCAN Study, for example, was a randomized placebo-controlled trial designed to assess the chemopreventive efficacy of a retinoid (vitamin A) or an antioxidant (*N*-acetylcysteine) administered for 2 years to 2592 subjects with head and neck or lung cancers, almost all of whom were former (and many current) smokers. No statistically significant difference in second primary rates was observed with either agent. In aggregate, the data suggest a role for retinoids as secondary chemopreventives only in the context of early neoplasia. Because of their potential as a class, retinoids are being further explored in phase III trials.

Lung Cancer

Lung cancer is the leading cause of cancer death in both men and women. Smoking is the major preventable risk factor and accounts for up to 90% of all lung cancers. The lung presents formidable obstacles to observation and acquisition of tissue, thus hampering molecular characterization and mutational sequencing of early lesions. For this reason, manifestations of preinvasive lung neoplasia have traditionally been detected through random sampling of the bronchial tree (for metaplasia or dysplasia) or sputa (for cytological atypia). These screening methods, however, are inexact and inefficient. In one study, only 50% of heavy smokers screened by bronchoscopy met entry criteria for bronchial metaplasia, severely slowing study progress. Correlations between cigarette consumption and the risk of bronchial metaplasia may simplify selection criteria for future studies. A preliminary effort in this regard was a 6-month trial of the retinoid etretinate in persons with bronchial metaplasia, which showed a significant improvement in the metaplasia index. A less invasive approach to the detection of premalignant changes of the lung involves evaluation of sputum cytology. In one study, qualitative cytological changes in sputa correlated well with the use of another retinoid (13cRA). A slightly different efficacy assessment follows changes in sputum cytology serially collected from high-risk patients. In the course of these investigations, several potential confounders have come to light. For example, in a study of active smokers with bronchoscopic dysplasia or metaplasia treated with etretinate versus placebo, high rates of spontaneous regression (20%) and significant reductions in metaplasia upon smoking cessation were noted in the placebo group. As a result, dysplasia is now regarded as the more robust endpoint for prevention research. New directions in early phase lung cancer chemoprevention include the use of light-induced fluorescent (LIFE) bronchoscopy to detect prevalent dysplastic lesions with greater sensitivity and aerosolized agents to deliver agents directly to the respiratory epithelium.

An early trial of high-dose vitamin A in patients with resected stage I lung cancer showed preliminary efficacy in preventing the emergence of second primary cancers and recurrence of the original tumour. Although vitamin A significantly reduced new cancers, tumour recurrences were unaffected. The phase III EUROSCAN Study, however, did not confirm these positive results. NSAIDs, LOX inhibitors and the micronutrient selenium may have protective effects against lung cancer, and are currently being evaluated in clinical chemoprevention trials.

It is sobering to note that two large lung cancer prevention trials have conclusively demonstrated the harmful effects of β-carotene supplementation. The first reported trial was the Alpha-Tocopherol, Beta Carotene (ATBC) Prevention Study of 29 133 Finnish male smokers randomized to α-tocopherol (vitamin E), β-carotene, α-tocopherol plus β-carotene or placebo for 5–8 years (Alpha-Tocopherol, Beta Carotone Cancer Prevention Study Group, 1994). Surprisingly, men receiving β-carotene

experienced higher rates of lung cancer and lung-cancer associated mortality. These disturbing results were confirmed by the Carotene and Retinol Efficacy Trial (CARET) of 18 314 smokers and asbestos workers randomized to a combination of β-carotene and retinyl palmitate versus placebo for a median of 4 years. An interim analysis of CARET was conducted in the wake of the ATBC results, and showed a 28% increase in lung cancer incidence and a 17% increase in mortality rates in the β-carotene arm, forcing early termination of the trial.

Gastro-oesophageal Cancer

Gastric cancer is recognized as a major cause of cancer-associated mortality worldwide. Aetiological factors identified through observational studies include diets high in fats, pickled vegetables and salted or smoked fish, diets low in fruits and vegetables and gastric mucosal infection with *Helicobacter pylori*. The potential benefits of antioxidant supplements and/or *H. pylori* eradication are being evaluated in high-risk populations with prevalent precursor lesions. One study recently reported a marked histopathological improvement with ascorbic acid, β-carotene or antibiotic therapy administered over 6 years. Observational and laboratory data suggest that NSAIDs may also have preventive efficacy in this context.

Squamous oesophageal cancer is relatively infrequent in Western populations, but a major concern in Asia and Africa. In contrast, Barrett-associated adenocarcinoma is quickly rising in prevalence in the USA. Risk factors for the disease include tobacco, alcohol and dietary deficiencies in the case of squamous cell and gastro-oesophageal reflux in adenocarcinoma. Both cancers are characterized by dysplastic precursor lesions that can be identified by cytological or histopathological analyses.

Most work on oesophageal cancer chemoprevention has been conducted in areas with an extremely high prevalence of cancer, such as Northern China. Most of the trials have not reported improvements in the status of dysplasia with antioxidant vitamins or calcium supplementation. However, the largest trial to date demonstrated that treatment with β-carotene, α-tocopherol and selenium reduced all cancer-associated and overall mortality. Observational data suggest that NSAIDs may reduce the risk of cancer by 40–70%, though this hypothesis has yet to be tested in randomized clinical trials. Trials evaluating the merits of DFMO, COX-2 inhibitors, selenium, acid suppression, photodynamic therapy or combinations of these interventions in persons with prevalent dysplasia are currently under way.

Colon Cancer

Colorectal cancer is a leading cause of cancer morbidity and mortality worldwide. Despite continuing advances in therapy, long-term survival has not improved significantly over the last four decades. As a result of insights into the biologic potential of preinvasive neoplasia and effective early detection and interventional technologies, colorectal cancer has become a preventable disease. Because of these successes, colorectal cancer has become the model for chemopreventive agent development.

The most definitive NSAID chemoprevention trial to date was the Physicians' Health Study (PHS), in which 22 071 subjects were randomized to aspirin (325 mg QOD) versus placebo for an average of 5 years (Gann *et al*., 1993). No significant difference was seen between groups in the self-reported frequency of new large bowel malignancies. This surprising null effect has been variously attributed to the rarified cohort (all subjects were male physicians), the lack of uniform colorectal surveillance guidelines and the limited dose and duration of aspirin use. In the wake of the PHS, other large trials are testing the efficacy of the same or higher doses of aspirin for chemoprevention against colorectal adenoma/cancer in more heterogeneous cohorts.

Selective COX-2 inhibitors (e.g. celecoxib, rofecoxib) are under investigation in clinical trials of persons at elevated risk of colorectal cancer due to germ-line mutations or prior sporadic neoplasia. COX-2 is overexpressed in nearly 50% of adenomas and 80–85% of adenocarcinomas, whereas it is rarely expressed in normal colorectal epithelium. This overexpression appears to be functionally important for neoplastic progression, as evidenced by the lower incidence of intestinal polyps in mice without a functional *COX-2* gene. A phase II/III randomized, placebo-controlled trial administered celecoxib for 6 months to 83 FAP subjects and showed significant reductions in colorectal adenoma number and size, with a side effect profile comparable to placebo (Steinbach *et al*., 2000). These data served as the premise for a landmark FDA approval of celecoxib to complement standard care (i.e. surveillance and prophylactic surgery) for patients with FAP, a genetic condition which, in the absence of interventions, assures the development of colorectal cancer. Whether celecoxib as a single agent, or in combination with other drugs, can reduce the extraordinarily high cancer risk of this cohort will be established in future clinical trials.

Calcium has been studied in five adenoma recurrence trials. The most definitive of these studies randomized nearly 1000 persons with prior adenomas to calcium versus placebo and demonstrated a statistically significant 19% reduction in adenoma recurrence (Baron *et al*., 1999). The Women's Health Initiative (WHI) is currently randomizing 45 000 postmenopausal women to calcium and vitamin D versus placebo (among other agents) for a mean of 9 years and will assess incidence of diseases such as osteoporosis and colorectal cancer. Given its size and duration, the WHI is likely to generate authoritative data on the utility of

calcium supplementation for the prevention of colorectal cancer.

Two clinical trials have failed to demonstrate that dietary fibre modulates adenoma recurrence. One randomized trial of wheat bran fibre in 1429 subjects with prior adenomas showed no significant reduction in colorectal adenomas attributable to fibre. Another trial of persons who adopted a low-fat, high-fibre, fruit and vegetable diet also reported null results (Alberts *et al*., 2000; Schatzkin *et al*., 2000).

Clinical trials have shown that DFMO given over 12 months significantly suppresses rectal polyamine levels in patients at risk of colorectal neoplasia, and is well tolerated despite reports of mild gastrointestinal and audiological toxicities (Meyskens *et al*., 1998). Exciting data show that DFMO and NSAIDs in combination synergistically reduce intestinal neoplasia compared with placebo, even when the doses are reduced by 50%. Two clinical trials are testing this hypothesis, one in subjects with prior sporadic adenomas and the other in persons with FAP.

Liver Cancer

Hepatocellular carcinoma (HCC) (also referred to as hepatoma) is the most common type of liver cancer and one of the most common cancers (and causes of cancer death) globally, creating a strong impetus for cancer prevention. HCC aetiological risk factors include chronic infection with hepatitis B or C viruses (HBV, HCV) and cirrhosis, which may be considered a preneoplastic condition. Carcinogens such as aflatoxin and alcohol further potentiate risk.

Prevention of HCC has been accomplished with hepatitis B vaccine. For HCV-induced HCC, however, interventions are still being developed. The most promising chemopreventive agent for HCV-induced HCC is interferon, an immune modulator, that induces sustained virological responses (undetectable HCV RNA) and histological improvements in a limited number of patients. Interestingly, more than 10 observational studies suggest that interferon also reduces the risk of liver cancer. In addition, two of three small, randomized controlled trials conducted in persons harbouring HBV chronic infection or HCV-induced cirrhosis have reported significant reductions in liver cancer incidence in patients receiving interferon for limited periods. Clinical trials to confirm the preventive efficacy of interferon are ongoing.

Other promising approaches may involve induction of phase 2 enzymes to bolster aflatoxin detoxification, and retinoids for their antineoplastic activity. A trial of an acyclic retinoid (polyprenoic acid) reported significant reductions in the incidence of second primary HCCs, although no effect against tumour recurrence. COX-2 inhibitors or SERMs may be evaluated in future trials.

Breast Cancer

Breast cancer is the most common form of cancer among women in the USA. The incidence of breast cancer has been rising for the past two decades, although mortality has remained relatively stable. Much of the increase in incidence is associated with increased screening by physical examination and mammography. There are several well-established factors for breast cancer, including family history, nulliparity, early menarche, advanced age and a personal history of breast cancer.

The National Surgical Adjuvant Breast and Bowel Project (NSABP) P-1 Trial, completed in 1998, is the largest breast cancer chemoprevention study conducted to date (Fisher *et al*., 1998). A total of 13 388 women who met specific criteria for high risk of breast cancer (based on the Gail predictive model) were randomized to tamoxifen versus placebo (Gail *et al*., 1989). The rationale for this study evolved out of studies that showed a reduced rate of contralateral second primary cancers in women treated adjuvantly with tamoxifen. The study was prematurely terminated when an interim analysis showed a 49% reduced risk of breast cancer in women receiving tamoxifen. Importantly, risk reductions were noted with tamoxifen in all subsets.

The P-1 Trial demonstrated the protective potential of tamoxifen, but also exposed its profound limitations. Most striking was the finding that despite reductions in oestrogen receptor positive (ER +) disease, tamoxifen evidently has no impact on oestrogen receptor negative (ER−) breast cancer. In addition, two smaller trials of women at lower risk for breast cancer did not confirm P-1 Trial results. One of these studies, the Italian Tamoxifen Prevention Study, randomized 5408 hysterectomized women to tamoxifen versus placebo (Veronesi *et al*., 1998). Unlike the P-1 Trial, patient selection was not based on breast cancer risk. In fact, nearly half of the accrued cohort had prior ovarian ablation and was consequently at reduced risk of breast cancer, owing to low circulating oestrogen levels. Complicating interpretations further, the trial prematurely stopped accrual owing to concerns over a high drop-out rate. Under these circumstances, no difference in breast cancer incidence was observed between study groups. Even so, subset analysis suggested that women using hormone replacement therapy may have derived some preventive benefit from tamoxifen. In a separate study based at the Royal Marsden Hospital in the UK, 2494 women aged 30–70 years with a family history of breast cancer were randomized to tamoxifen versus placebo (Powles *et al*., 1998). As in the Italian trial, the premature drop-out rate was high and tamoxifen did not alter the incidence of breast cancer. Divergent outcomes between the European and American trials may relate, in part, to baseline differences between the study populations.

Tamoxifen was also tested as a secondary chemopreventive against breast cancer in the NSABP B-24 trial (Fisher *et al*., 1999); 1804 women with a prior diagnosis of DCIS were randomized to tamoxifen versus placebo for 5 years, and tamoxifen was shown to reduce the incidence of ipsilateral and contralateral breast cancers. Moreover, because effects of tamoxifen on DCIS were comparable to its effects on invasive breast cancer, DCIS is now regarded as a plausible SEB for breast cancer chemoprevention.

Despite FDA regulatory approval, tamoxifen has not been universally adopted as the standard of care for secondary prevention of breast cancer in the USA. Qualms over use of tamoxifen for this indication chiefly relate to the toxicity profile, which includes endometrial cancer, venous thromboembolism, hot flushes, cataracts and vaginal dryness. Negative reports from the European trials may also have dissuaded physicians from prescribing tamoxifen for chemoprevention. A follow-up study of 22 000 women at high risk of breast cancer, the Study of Tamoxifen and Raloxifene (STAR), will assess the relative efficacy and safety profiles of tamoxifen and a closely related compound, raloxifene.

Although premalignant breast lesions have long been associated with invasive cancer, prospective assessment of chemopreventive effects has been hampered by surgical removal of early breast neoplasia. Moreover, the organ's architecture limits options for reliably identifying early markers that may serve as SEBs. To determine tissue-based risk biomarkers, Fabian *et al*. recently published the results of 10 years of biomarker data derived from fine-needle breast aspirations performed in women with family histories of breast cancer, prior precancerous biopsies or invasive breast cancer (Fabian *et al*., 2000). Random peri-areolar fine-needle aspirations (up to 10 per breast) were obtained at baseline and after 6 months and evaluated for cytological and molecular markers (e.g. EGFR, p53, ER, HER2/NEU). At 45 months of follow-up, hyperplasia with atypia was associated with an increased cancer risk. Furthermore, although no single molecular marker was informative, overexpression of multiple molecular markers was associated with an increased cancer risk. This study establishes markers of risk in would-be candidates for chemoprevention, and provides an important clinical model for testing chemopreventive agents. A different investigational strategy assesses biomarkers in ductal carcinoma *in situ* (DCIS) specimens and adjacent, apparently unaffected epithelium. This 'presurgical model' exploits the interval between the time of diagnostic biopsy and definitive surgery in order to test whether test agents modulate biomarkers.

Prostate Cancer

Prostate cancer is the most common nondermatological malignancy worldwide. Risk factors include age, family history and black race. The only screening test for the detection of early prostate cancer is serum prostate-specific antigen (PSA), which may reduce cancer mortality in the context of surgical and/or radiation therapy directed at early disease. PSA biology, function, age-related expression and its various molecular forms are under intense investigation to determine whether PSA might be an SEB for prostate cancer. Two precursor lesions for prostate cancer (atypical adenomatous hyperplasia and prostatic intraepithelial neoplasia (PIN)) are also under investigation as candidate biomarkers for prostate cancer.

There are no mature data on agents for the prevention of prostate cancer, but several trials are under way. Finasteride (5α-reductase inhibitor), selenium and α-tocopherol (vitamin E) are undergoing clinical testing. The hypothesis to test the last two agents (selenium and α-tocopherol) arose from observations of reduced incidences of prostate cancer by secondary analysis in chemoprevention trials of other target organs.

Bladder Cancer

Bladder cancer is a common cause of cancer morbidity and mortality worldwide. In industrialized countries cigarette smoking is the greatest risk factor for bladder cancer, most of which are transitional cell carcinomas. In developing areas of the Middle East and parts of Africa, infection with schistosomiasis accounts for a high incidence of bladder cancer, the majority of which are squamous cell carcinomas.

Patients with newly diagnosed bladder cancer typically present with premalignant bladder lesions, most of which are curable. Management of these lesions, also referred to as superficial bladder cancer, includes surgical and non-surgical interventions. Superficial bladder cancer is associated with the highest annual rate of second primary tumours of any site. Although most of these new tumours are noninvasive, a significant percentage of patients have multiple recurrences, some of which will progress to invasive disease. Intravesical therapy with BCG vaccine after surgical resection decreases recurrence rates by 45%. Oral retinoids, NSAIDs and megadose vitamins are undergoing phase III clinical testing to prove efficacy as secondary preventives against superficial bladder cancer. The data from these studies will mature over the next several years.

Cervical Cancer

Cervical cancer, one of the most common cancers in women, is a promising area for chemoprevention because of the accessibility of the cervix to evaluation (i.e. Pap tests) and the relatively slow progression to cervical cancer from recognized precursor lesions, termed cervical intraepithelial neoplasia (CIN). High-grade CIN is usually treated by removing the abnormal tissue to reduce the risk of progression to invasive cancer. In fact, broad application of screening programmes to detect and treat preinvasive neoplasia have greatly reduced cervical cancer incidence

and mortality. As a result, CIN has achieved a high degree of validation as an SEB for cervical cancer incidence and mortality, at least in the context of screening coupled with surgery. To date, no medical intervention for cervical cancer has demonstrated an acceptable TI, however several agents are under investigation. Retinoids in particular have shown promise, although their associated toxicities may hinder compliance.

The major risk factor for cervical cancer is infection with human papillomavirus (HPV). Vaccines that eradicate or prophylax against the most common oncogenic viruses (primarily HPV-16 and -18) or their related proteins (E6 and E7) may prove valuable as chemopreventives against cervical cancer. Most cervical cancers (95%) contain HPV, and worldwide the most common type is HPV-16. A 6000 woman placebo-controlled trial of an HPV-16 vaccine will soon be initiated in Costa Rica, where the prevalence of HPV positivity in women with CIN approaches 89%. This vaccine is one of several under development that uses virus-like particles (VLPs) that resemble the authentic virus, but do not pose an infectious risk, to produce high titres of neutralizing antibodies. In the Costa Rica study, virological and pathological endpoints will be used to assess vaccine efficacy.

Skin Cancer

Skin cancer is classified as either nonmelanoma or melanoma, depending on the cell of origin. The incidence of non-melanoma skin cancer (NMSC) (i.e., basal and squamous cell carcinomas (BCC, SCC)) nearly equals that of all other cancers combined. Moreover, NMSC rates are increasing, which is consistent with dose-dependent effects of photocarcinogenesis on ageing populations. NMSC may arise from actinic keratoses (AKs), lesions that are associated with cumulative exposure to UV irradiation, immunosuppression and certain chemicals. AKs are typically cured by destructive techniques (cryotherapy, chemical peels, carbon dioxide laser, shave excision or dermabrasion) or the application of topical agents (5-FU, trichloroacetic acid, phenol or retinoids). However, the malignant potential of AKs, coupled with the cost and cosmetic deformity of standard management, mandate the identification of well-tolerated agents that eradicate AKs, and potentially reverse or retard progression to NMSC. Certain cohorts are at exceptionally high risk of developing NMSC and provide accelerated models for skin carcinogenesis due to chronic immune-suppression (as with solid organ transplant recipients), ultraviolet hypersensitivity syndromes or certain rare genetic syndromes. Several agents have been tested experimentally for the chemoprevention of NMSC, including retinoids, NSAIDs, DFMO and green tea compounds (e.g. polyphenols and epigallocatechin-3-gallate (EGCG)).

Chemoprevention trials of oral vitamin A (retinol) for nonmelanoma skin cancer have shown mixed results. One phase III study showed efficacy against moderate to severe AKs and associated reductions in SCC. However, in another study of subjects with higher risk disease (i.e. a history of skin cancer), the results were unremarkable. Topical formulations of DFMO and colchicine have shown preliminary efficacy in the regression of AKs. Recent US regulatory approvals for regression of AKs include aminolaevulinic acid hydrochloride (Levulan Kerastick™) with photodynamic therapy (PDT) and sodium diclofenac (Solaraze™) for AKs (http://www.fda.gov/cder/approval/index.htm).

Melanoma skin cancer accounts for a much smaller fraction of skin cancers than does NMSC, but it is a substantially more lethal disease with few good therapeutic options. Despite a robust melanocyte progression model with well-characterized intermediate lesions, clinical experience in the chemoprevention of melanoma is minimal. Topical retinoids have been preliminarily tested in subjects with dysplastic naevi (DN), precursor lesions with a high rate of malignant transformation to melanoma. Although clinical and histological improvements were observed in DN treated with topical tretinoin, the responses were not durable. In addition, the degree of skin irritation prompted concerns over compliance with the regimen. A follow-up study will assess whether improved efficacy is achieved through coadministration of systemic retinoids with a lower dose of the topical formulation.

Melanoma is an immunogenic tumour, and may be a good target for vaccine-based prevention. Typically the response to immunization correlates inversely with disease stage and tumour burden, thus vaccines are likely to have a greater impact on subclinical disease (i.e. prevention) than advanced cancer (i.e. therapeutics). This provides a strong rationale for testing promising melanoma vaccines in a preventive context. (See the chapter *Genetic and Cellular Vaccines*.)

THE FUTURE OF CANCER CHEMOPREVENTION

As recently as 1990, there were no accepted SEBs or approved agents for cancer prevention. Since then, preinvasive neoplastic lesions have come to be regarded as excellent markers of risk and chemopreventive response, and several preventive agents have been approved. Our ability to characterize neoplasia on the molecular level is likely to accelerate advances over the next decade (see **Table 12**).

Worldwide, carcinogenesis and atherogenesis impart heavy societal burdens. Reductions in competing cardiovascular deaths – achieved through prevention – ensure that cancer will become the dominant public health issue for ageing populations. With this recognition comes increased responsibility for professional and public implementation of prevention as the foundation of clinical care and health policy. Indeed, we have already witnessed the emergence of a new cancer paradigm. Formerly the

Table 12 Promising directions in cancer chemoprevention

Interventions with improved mechanistic targeting
Genomics, proteomics, etc., to define molecular targets better
Combinatorial screening to identify agents with greater specificity
Broader consideration of agents with preventive potential
Antibiotics
Antiviral agents
Biologics/vaccines
Gene-directed interventions
Nutraceuticals
Cytostatic cancer 'chemotherapeutics' – optimized for risk cohorts
Improved agent administration
Biologically active dose
Targeted agent delivery – topical/regional
Optimized dosing schedules – frequency, duration
Agent combinations
Improved agent development pathways
New *in vitro/in vivo* models with high throughput and greater biological relevance
Better defined risk cohorts
Better validated biomarkers and animal models
Cross-sectional, multiplex efficacy evaluations, e.g. expression arrays
Sensitive and specific superficial imaging, e.g. high-resolution/magnifying endoscopy, 'optical biopsies,' functional imaging technologies
Nested prevention evaluations in therapeutic trials

simultaneous declaration of illness, disability and death, cancer is now regarded as one late stage in a prolonged disease process that affords both time and targets for intervention. As a result, our focus must shift from downstream events (cancer) to midstream processes (carcinogenesis). The challenges that this presents, and how we respond to them, will determine the pace of our success, although not success itself, which is inevitable.

REFERENCES

Alberts, D. S., *et al*. (2000). Lack of effect of a high-fiber cereal supplement on the recurrence of colorectal adenomas. Phoenix Colon Cancer Prevention Physicians' Network. *New England Journal of Medicine*, **342**, 1156–1162.

Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group (1994). The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *New England Journal of Medicine*, **330**, 1029–1035.

Baron, J. A., *et al*. (1999). Calcium supplements for the prevention of colorectal adenomas. Calcium Polyp Prevention Study Group. *New England Journal of Medicine*, **340**, 101–107.

Fabian, C. J., *et al*. (2000). Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. *Journal of the National Cancer Institute*, **92**, 1217–1227.

Fisher, B., *et al*. (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute*, **90**, 1371–1388.

Fisher, B., *et al*. (1999). Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet*, **353**, 1993–2000.

Gail, M. H., *et al*. (1989). Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute*, **81**, 1879–1886.

Gann, P. H., *et al*. (1993). Low-dose aspirin and incidence of colorectal tumors in a randomized trial. *Journal of the National Cancer Institute*, **85**, 1220–1224.

Gaziano, J. M., *et al*. (1996). Methodological issues in the design of primary prevention trials. *IARC Scientific Publications*, **136**, 7–12.

Hawk, E., *et al*. (2000). Biomarkers as surrogates for cancer development. *Current Oncology Reports*, **2**, 242–250.

Hong, W. K., *et al*. (1986). 13-*cis*-retinoic acid in the treatment of oral leukoplakia. *New England Journal of Medicine*, **315**, 1501–1505.

Kelloff, G. J., *et al*. (1995). Approaches to the development and marketing approval of drugs that prevent cancer. *Cancer Epidemiology Biomarkers and Prevention*, **4**, 1–10.

Meyskens, F. L., *et al*. (1998). Effect of alpha-difluoromethylornithine on rectal mucosal levels of polyamines in a randomized, double-blinded trial for colon cancer prevention. *Journal of the National Cancer Institute*, **90**, 1212–1218.

Powles, T., *et al*. (1998). Interim analysis of the incidence of breast cancer in the Royal Marsden Hospital tamoxifen randomised chemoprevention trial. *Lancet*, **352**, 98–101.

Schatzkin, A., *et al*. (2000). Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. Polyp Prevention Trial Study Group. *New England Journal of Medicine*, **342**, 1149–1155.

Smalley, W. E. and DuBois, R. N. (1997). Colorectal cancer and nonsteroidal anti-inflammatory drugs. *Advances in Pharmacology*, **39**, 1–20.

Sporn, M. B. (1976). Approaches to prevention of epithelial cancer during the preneoplastic period. *Cancer Research*, **36**, 2699–2702.

Steinbach, G., *et al*. (2000). The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *New England Journal of Medicine*, **342**, 1946–1952.

Temple, R. (1999). Are surrogate markers adequate to assess cardiovascular disease drugs? *Journal of the American Medical Association*, **282**, 790–795.

Veronesi, U., *et al*. (1998). Prevention of breast cancer with tamoxifen: preliminary findings from the Italian randomised trial among hysterectomised women. Italian Tamoxifen Prevention Study. *Lancet*, **352**, 93–97.

Winawer, S. J., *et al*. (1993). Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *New England Journal of Medicine*, **329**, 1977–1981.

FURTHER READING

Boone, C. W., *et al*. (1997). Properties of intraepithelial neoplasia relevant to cancer chemoprevention and to the development of surrogate end points for clinical trials. *Proceedings of the Society for Experimental Biology and Medicine*, **216**, 151–165.

Pappalardo, P. A., *et al*. (1998). Microdissection, microchip arrays, and molecular analysis of tumor cells (primary and metastases). *Seminars in Radiation Oncology*, **8**, 217–223.

Sporn, M. B. (1980). Combination chemoprevention of cancer. *Nature*, **287**, 107–108.

Sporn, M. B. and Suh, N. (2000). Chemoprevention of cancer. *Carcinogenesis*, **21**, 525–530.

Wallace, M. A. and VanDam, J. (2000). Enhanced gastrointestinal diagnosis: light-scattering spectroscopy and optical coherence tomography. *Gastrointestinal and Endoscopy Clinics of North America*, **10**, 71–80.

Wattenberg, L. W. (1985). Chemoprevention of cancer. *Cancer Research*, **45**, 1–8.

Wattenberg, L. W. (1997). An overview of chemoprevention: current status and future prospects. *Proceedings of the Society for Experimental Biology and Medicine*, **216**, 133–141.

Chapter 91

Antisense and Ribozyme Therapy

Mohammed Kashani-Sabet
University of California, San Francisco, CA, USA

Kevin J. Scanlon
Keck Graduate Institute, Claremont, CA, USA

CONTENTS

- Introduction
- Scientific Background
- Design, Formulation and Delivery of Antisense ODNs and Ribozymes
- Preclinical Models and Results
- Clinical Trials of Antisense and Ribozymes
- New Uses for Antisense and Ribozymes in Cancer
- Conclusion

INTRODUCTION

The fundamental advances in molecular genetics in the last two decades have introduced a number of new tools in our armamentarium for the manipulation of gene expression. Chief among these are the use of antisense molecules initially, and ribozymes subsequently, for the manipulation of mammalian gene expression. Not surprisingly, some of the most important demonstrations of the utility of antisense and ribozyme targeting have occurred in the realm of cancer. Moreover, recent advances in understanding the molecular basis of cancer have defined a plethora of new targets for antisense and ribozyme-based inhibition. More recently, the maturation of this technology to the clinical oncological arena is being achieved with the completion of a number of clinical trials using a broad array of targets in cancer. In short, the biological and therapeutic potential of antisense and ribozyme technology is being explored. We therefore take the opportunity in this chapter to review the scientific basis for this technology and discuss its broad utility in cancer-related applications, including the progress made and problems encountered in this field in its march toward the clinical arena.

SCIENTIFIC BACKGROUND

Antisense

The principal appeal of antisense therapy is its inherent simplicity. With the knowledge that a particular stretch of DNA encodes an mRNA molecule which is then the template for protein synthesis, it was hypothesized that an antisense DNA or RNA strand would hybridize to the mRNA in complementary Watson–Crick fashion, thereby resulting in blockage of translation and sequence-specific inhibition of expression. Antisense molecules are commonly composed of single-stranded DNA, where they are termed antisense oligodeoxynucleotides (ODNs). Alternatively, DNA encoding the antisense can be expressed as antisense RNA by an expression cassette. With antisense ODNs, the presumed mechanisms of action include blockage of translation, RNA transport and splicing. Most importantly, perhaps, the RNA–DNA hybrid can be recognized and cleaved by RNase H, yielding an additional mechanism of repression of gene expression. Finally, antisense RNA generated by expression cassettes or viral vectors can also block target gene expression through activation of double-stranded RNase. The first demonstration of the ability of antisense molecules to inhibit target gene expression (Zamecnik and Stephenson, 1978) raised considerable interest in the potential of antisense molecules not just as a molecular tool in biomedical research, but also as a novel class of therapeutic agent.

With the further testing of antisense oligonucleotides in preclinical models, a number of problems surfaced which threatened the theoretical appeal and presumed scientific mechanisms of action of these compounds. Over time, however, a clearer picture has emerged regarding the practical realities of antisense targeting. Moreover, the issues of nonsequence-specific and/or non-antisense activity need to be fully addressed before a particular antisense molecule is to enter the clinical arena in order to maximize the potential of antisense compounds.

Ribozymes

Catalytic RNAs (known as RNA enzymes or ribozymes) were discovered initially as self-cleaving entities in two disparate systems, the ribonuclease P of *Escherichia coli* (Guerrier-Takada *et al*., 1983), and the intervening sequence of *Tetrahymena thermophila* (Zaug and Cech, 1986). The hammerhead ribozyme was identified in the virusoid from lucerne transient streak virus, and the hairpin ribozyme from the 'minus' strand of satellite RNA of tobacco ringspot virus. The initial demonstration of *trans*-acting catalytic activity and *in vitro* mutagenesis studies of the hammerhead ribozyme (Haseloff and Gerlach, 1988) paved the way for use of ribozymes as tools to manipulate the expression of potentially any gene. The therapeutic potential of hammerhead ribozymes was subsequently examined in the setting of human immunodeficiency virus (HIV) infection (Sarver *et al*., 1990) and cancer (Scanlon *et al*., 1991). **Figure 1** depicts the secondary structure of the hammerhead ribozyme, including conserved sequences required for catalytic activity, as well as the proposed interaction with its target mRNA. The hammerhead ribozyme is composed of two regions: (1) the catalytic core, effecting the cleavage reaction, and (2) three hybridizing regions. Stems I and III represent the hammerhead flanking sequences that hybridize to the target RNA in complementary fashion. Hammerhead ribozymes have been shown to recognize and cleave target sequences containing XUN, with X being any nucleotide and N being A, C or U. In general, targets containing GUC, GUA, GUU, CUC and UUC sequences are favoured. Modifications in the hammerhead have been shown to yield the ability to cleave AUA sequences in *trans*. Recently, *in vitro* selection techniques were utilized to expand further the range of cleavable triplets by the hammerhead ribozyme to include AUG and even AAG, rendering the central U nonessential. Therefore, it is likely that further advances will continue to increase the range of targets amenable to hammerhead ribozyme cleavage.

The hairpin ribozyme has also undergone a similar path of development and progress. The hairpin ribozyme (depicted in **Figure 2**) consists of 50 bases and cleaves corresponding 14-base RNA substrates. The ribozyme consists of four helices separated by two internal loop sequences. The hairpin binds to the substrate through helices 1 and 2 and cleaves 5′ to a guanosine within loop A separating helices 1 and 2. Mutagenesis studies have defined the B^4, N^5, G^6, U^7, C^8 sequence in the target RNA as optimal for recognition and cleavage by the hairpin ribozyme. In this scheme, cleavage occurs between N^5 G^6, and B represents C, U or G, whereas N represents C, U, A or G. Following the identification of transacting activity by hairpin ribozymes, their potential therapeutic utility

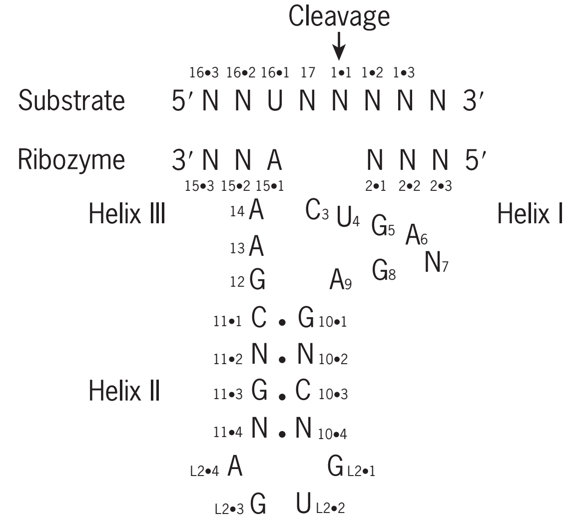


Figure 1 Conserved sequences of the hammerhead ribozyme and position relative to its target RNA.

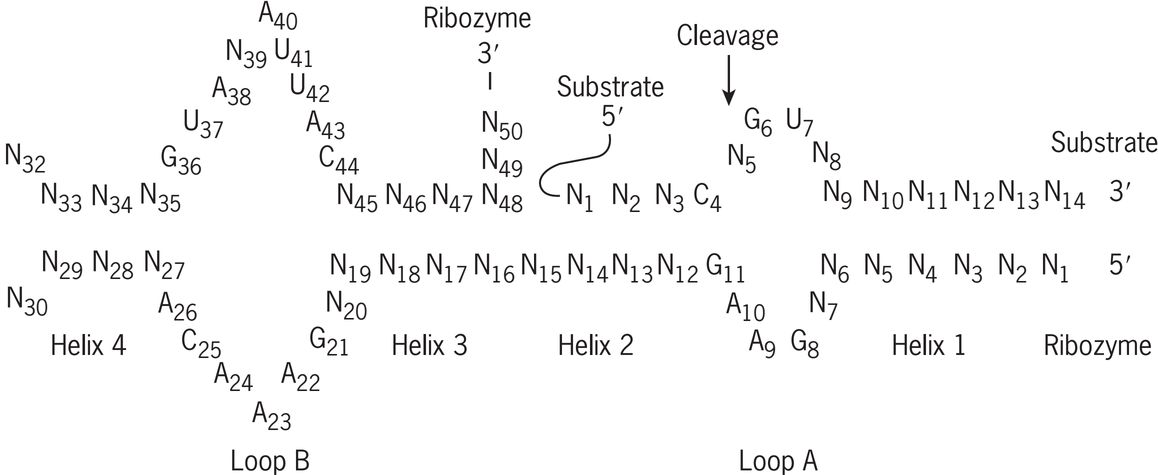


Figure 2 Consensus sequences of the hairpin ribozyme and position relative to its substrate.

was examined initially in the setting of HIV (Ojwang *et al.*, 1992).

Finally, in addition to cleavage activity, ribozymes can result in inhibition of target gene expression through antisense mechanisms, namely blockage of translational machinery or through activation of RNase recognizing double-stranded RNA.

DESIGN, FORMULATION AND DELIVERY OF ANTISENSE ODNs AND RIBOZYMES

Since the initial reports demonstrating the utility of antisense and ribozymes in suppressing gene expression, a number of issues have emerged which deserve further attention in the optimization of antigene strategies. These issues relate to (1) optimal design, (2) optimal formulation and (3) optimal delivery of antisense ODNs and ribozymes.

Design of Antisense ODNs and Ribozymes

Although numerous studies have demonstrated the successful implementation of antisense- and ribozyme-based strategies, with certain exceptions the optimal design of such molecules still remains largely empirical. Thus, proof of target gene inhibition, catalytic activity (in the case of ribozymes), sequence-specific effects (in the case of ODNs), biological activity and the inclusion of convincing control sequences is required for each new molecule that is being examined. Nevertheless, certain patterns have emerged, attention to which will probably increase the possibility of a successful endeavour. In the design of antigene strategies, issues regarding both the target and the antigene must be considered.

With respect to the target gene to be inhibited, several factors are of paramount importance. First, selection of a target that plays a key role in a cellular pathway, or in a biological phenotype, is crucial. In short, the more important the target gene, the more likely it is that antigene strategies will be successful (see the next section on optimal targets tested thus far in cancer). Moreover, the half-life of the target RNA and protein may play a key role in the kinetics of the desired inhibition, especially in studies that allow for transient inhibition of the target gene. Therefore, targeting an immediate–early gene (with a short protein half-life) such as c-*fos* was shown to reverse drug resistance more rapidly than targeting the *mdr*-1 gene, whose protein is more stable. Thus, the stability of the target protein may dictate the relative utility of transient studies with ODNs as opposed to achieving prolonged inhibition with the use of stable transformants.

The next critical issue concerns the selection of the appropriate stretch of mRNA within the target gene. This issue differs from target to target and must be tested with the design of each new molecule. Both antisense and ribozyme strategies have been shown to be successful whether they target the 5′ end of the mRNA, the 3′ untranslated region or regions in between. More important than the actual site is the accessibility of the target RNA region to antigene sequences. In this regard, the formation of stem-loops or lariats within the target mRNA as well as protein–RNA interactions may severely limit the accessibility of the target RNA to antisense ODNs or ribozymes. In fact, double-stranded regions of folded single-stranded RNA and sites of interaction with RNA-binding proteins have been shown to be less accessible to ribozymes. In one study, a ribozyme expression library demonstrated that effective target sites for ribozymes were limited to single-stranded gaps of mRNA. As an illustration, screening the human acetylcholinesterase transcript for GUC and CUC triplets revealed that only five of the 55 potential sites were accessible to hammerhead ribozyme-mediated cleavage. To address this concern, several groups have utilized RNA folding predictions prior to the selection of sites for ribozyme cleavage. In the case of ribozymes, several groups have utilized *in vitro* cleavage of target RNAs to screen ribozymes recognizing different areas of the same gene. However, the kinetics observed with *in vitro* studies is not always predictive of ribozyme activity *in vivo*, thus underscoring the need for empirical testing.

With respect to the design of the antisense or ribozyme molecule itself, several additional factors must be considered. The first is the optimal length of antigene sequences. Again, as with the selection of the inhibition site, no universal rule can be stated. In general, most studies have used ODNs (or flanking sequences in the case of ribozymes) as small as 12–20 bases in length. However, longer antisense molecules, in certain cases including the entire gene, have been shown to retain activity. This issue may be especially important in the case of ribozymes where the trade-off between specificity (necessitating longer sequences) and rapid dissociation that allows multiple turnover (favoured by shorter sequences) may help determine success. A more recent study examining this issue found a minimum length of 51 nucleotides in the 3′ antisense arm to be necessary for inhibition of HIV replication. Conversely, short-chain ribozymes were superior when microinjected into the cytoplasm or when transcribed *in vitro*. Therefore, there may be different requirements for ribozyme length dependent on the site of inhibition (nucleus vs cytoplasm).

In the case of antisense ODNs, recent studies have shed light on additional considerations in the inclusion of specific sequences within the antisense moiety. The most troubling initial finding was the observation of biological activity by nonantisense mechanisms with the use of phosphorothioate ODNs. First, ODNs with four contiguous guanine residues (also known as G quartets) were shown to display a high affinity for proteins (Krieg *et al.*, 1995) and

to be capable of mediating anti-proliferative effects through a nonantisense mechanism (Burgess *et al*., 1995). Subsequently, several reports indicated significant interactions between antisense ODNs and heparin-binding proteins. Mac-1, also known as CD11b/CD18, a member of the integrin family, was shown to act as a cell-surface receptor for ODNs, capable of mediating their internalization (Benimetskaya *et al*., 1997). Moreover, ODNs were shown to interact nonsequence specifically with two other heparin-binding proteins, the vascular endothelial growth factor (VEGF) receptor known as flk-1 and the epidermal growth factor receptor. In the case of flk-1, ODN binding resulted in perturbations of ligand-induced activation and receptor phosphorylation. To address these concerns, guidelines for the appropriate testing and development of specific antisense ODNs have been established (Stein and Krieg, 1994).

More recently, so-called 'second generation' antisense ODNs have been developed in order to improve activity and reduce toxicity. These have included modifications in the base, the sugar or the phosphate backbone. A prominent example is the 2′-alkyl substitution in the ribose moiety. Therefore, the prototypical ODN has been defined to consist of a 20-mer with 2′-alkyl-modified bases and phosphorothioate linkages on the 5′ and 3′ ends of the ODN with a phosphodiester backbone.

Formulation of Antisense/Ribozymes

Given the inherent instability of RNA and the susceptibility of DNA to serum nucleases, the utility of several modifications has been examined in order to optimize the half-life of antigene molecules. In general, these have fallen into two major categories: (1) chemical modifications of oligonucleotides for exogenous delivery and (2) use of expression cassettes to generate antisense or ribozyme RNA.

The most commonly used modification to the phosphodiester ODN is the phosphorothioate ODN, in which a single sulphur substitutes for oxygen at a nonbridging position at each phosphorus atom. This resulted in improved serum half-life but was compounded by the nonantisense effects discussed in the preceding section. In the case of ribozymes, several other modifications have been examined to improve exogenous delivery studies. Briefly, these include 2′-fluoro, 2′-amino, 2′-*O*-allyl and 2′-*O*-methyl substitutions. Another modification is the so-called chimaeric ribozyme, which includes both DNA and RNA moieties. Despite encouraging results *in vitro*, extensive *in vivo* testing of these modifications has been lacking in appropriate model systems. Therefore, for *in vivo* ribozyme studies, expression of the ribozyme by plasmid DNA or viral vector still represents the best-tested method.

In *in vitro* systems, long-term expression of ribozymes has been achieved by stably transfecting ribozyme DNA into the desired cell line. In this manner, transformant clones can be selected with high-level ribozyme expression, thereby increasing the chances of observing an altered phenotype (reduced cell growth, angiogenesis or resistance to chemotherapeutic agents). One of the major factors in modulating ribozyme expression is the selection of an appropriate promoter. As is common for many gene therapy applications, viral promoters represent a popular approach for use in expression cassettes. Ribozyme expression has been achieved using several viral promoters, including the dexamethasone-inducible mouse mammary tumour virus promoter, the SV40 promoter, the herpes simplex virus thymidine kinase promoter, the HIV and Rous sarcoma virus long terminal repeats and the cytomegalovirus (CMV) promoter–enhancer. Cellular promoters utilized to drive ribozyme expression have included the β-actin promoter and the melanocyte-specific tyrosinase promoter. Ribozymes have also been incorporated into tRNA genes to take advantage of RNA polymerase III-mediated transcription. Finally, ribozymes have been embedded in the 3′ untranslated region of other genes. Interestingly, even though the ribozyme RNA sequences transcribed intracellularly are typically enhanced by the addition of poly(A) tails, these noncomplementary sequences do not appear to alter significantly substrate cleavage *in vivo*. Nevertheless, a common approach to circumvent this potential problem is the use of *cis*-acting ribozymes to cleave and liberate the desired *trans*-acting ribozyme upon transcription.

Delivery of Antisense ODNs and Ribozymes

The development of appropriate delivery vehicles represents the single largest impediment to the successful application of gene therapy. Not surprisingly, limitations in vector technology have also slowed the progress of antisense and ribozyme approaches to the clinic. While a detailed overview of gene therapy vectors currently in use is beyond the scope of this chapter and is undertaken elsewhere, this section will discuss the strategies used to date to deliver potential antisense and ribozyme therapeutics.

With respect to antisense ODNs, given that the oligonucleotide is the active moiety itself (without a requirement for gene expression), there may not be a formal requirement for a delivery vector. Thus, these compounds have been more actively studied and have matured most rapidly to the clinical arena (see Clinical Trials section). Antisense ODNs are usually tested either alone or complexed with cationic liposomes. Optimizing antisense ODN delivery both to the appropriate tissues and in order to facilitate efficient intracellular (and intranuclear) uptake is of paramount importance. Uptake in nonspecific tissues can lead to increased toxicity, whereas poor cellular uptake

will result in diminished activity. Given that antisense ODNs are polyanionic structures, they diffuse poorly across lipophilic cellular membranes. Therefore, their cellular uptake occurs mainly through absorptive and fluid-phase endocytosis. Following internalization, ODNs enter the endosomal compartment, where they are susceptible to enzymatic degradation.

Several strategies have been tested to circumvent these potential obstacles. To improve intracellular delivery, multiple cations have been utilized to balance the net negative charge of ODNs. These have included cationic lipids, polylysine, polyethylenime and transferrin polylysine complexes. A recent report suggested that a new transfection reagent, cytofectin, was capable of destabilizing the endosomal compartment, thereby facilitating ODN release to the cytoplasm. Second, specific cell or tissue targeting has been attempted to direct ODN uptake. These approaches have included the use of receptors such as folate or transferrin. More recently, immunoliposomes were developed by coupling monoclonal antibodies directed against the disialoganglioside GD_2 to cationic liposomes. Phosphorothioate antisense ODNs targeting the c-*myb* oncogene were shown to bind selectively to neuroblastoma cells when complexed to these immunoliposomes, resulting in antiproliferative effects *in vitro* (Pagnan *et al*., 2000). In the case of ribozymes, when used as oligonucleotides, similar strategies have been employed as described for antisense ODNs. In addition, a recent report demonstrated efficient *in vitro* transport of ribozymes to the cytoplasm and nucleus when complexed to liposomes fused with haemagglutinating virus of Japan.

In the case of ribozymes expressed intracellularly, however, other delivery techniques have been examined. These have largely included the use of viral vectors. Paralleling the progress in the gene therapy field, early studies utilized retroviral vectors for ribozyme delivery. However, given the inability to transduce non-dividing cells efficiently and to generate high-titre virus particles, the ribozyme field has quickly moved to second-generation vectors. These have largely included recombinant adenoviruses and, more recently, adeno-associated viruses. Examples of viral targeting of ribozymes will be provided in the following section. Finally, in recent studies, cationic liposome–DNA complexes have been utilized to deliver ribozymes to tumour-bearing animals.

PRECLINICAL MODELS AND RESULTS

Antisense

It is beyond the scope of this chapter to discuss every study utilizing antisense technology with potential utility to cancer. Therefore, we will limit the discussion to those targets that have been shown to have significant therapeutic promise through vigorous *in vitro* and/or *in vivo* testing and thus may be closest to clinical application. In general, the type of cellular genes targeted can be categorized into the following: (1) oncogenes; (2) cell cycle regulatory genes; (3) drug-resistance genes; (4) angiogenic genes; (5) growth factor receptor genes; and (6) genes in cell signalling pathways. The targets representing these groups are summarized in **Table 1**.

With respect to oncogenes, some of the earliest studies utilizing antisense technology targeted oncogenes with known defects in particular tumours, thus potentially representing tumour-specific agents. A prominent example of this approach included antisense targeting the *BCR-ABL* gene formed by the chromosome 9 to 22 translocation (the Philadelphia chromosome) in chronic myelogenous leukaemia (CML). Several studies have examined the utility of targeting *BCR-ABL* with antisense ODN (Szczylik *et al*., 1991), such that this approach has reached the clinic (see the following section). A second targeted approach to specific tumour types is the use of antisense targeting of the mutant *ras* gene family. Specifically, antisense RNA against the entire K-*ras* gene was delivered by a retroviral vector to nude mice bearing human lung carcinoma cells, resulting in inhibition of tumour growth (Georges *et al*., 1993). Antisense ODNs have also been used to target mutated H-*ras*, a strategy which is under clinical investigation.

A second strategy for selecting oncogenes for antisense targeting concerns those that are activated by virtue of overexpression in multiple cancers without translocation or mutation. The most prominent examples of this approach have included antisense ODNs targeting c-*myb* in CML (Calabretta *et al*., 1991) and melanoma and targeting the c-*raf*-1 kinase in human bladder and breast carcinoma xenograft models (Monia *et al*., 1996).

A somewhat related strategy has concerned the use of antisense approaches to inhibit expression of growth factor receptors in different tumours. Examples of targets have included the insulin-like growth factor I receptor, the EGFR and the HER-2/*neu* receptor overexpressed in breast cancers (see **Table 1**).

More recently, the identification of cyclins and their role in normal and neoplastic cell growth has generated a new array of targets for cancer. Antisense targeting of cyclin D1 has been extensively studied in different preclinical models of carcinomas, and has been shown to enhance cytotoxicity to cisplatin. A recent study utilizing antisense to $p21^{WAF1/CIP1}$ showed increased sensitivity of colon cancer cell lines to radiation therapy.

Finally, other studies have examined and demonstrated the utility of targeting signalling molecules such as Nuclear Factor *Kappa* B and protein kinase Cα (Dean and McKay, 1994). More recent uses of antisense technology have targeted VEGF to inhibit angiogenesis and growth of melanoma cells *in vivo* and *bcl-2* to chemosensitize melanomas in SCID mice (Jansen *et al*., 1998).

Table 1 Use of antisense therapy in preclinical models of human cancer

Target gene	Sequence	Results	Reference
BCR-ABL	ODN	Inhibition of leukaemic but not normal haematopoietic colonies *in vitro* and *in vivo*	Szczylik *et al.* (1991)
K-*ras*	Retrovirus; antisense RNA	Intratracheal installation decreases lung carcinoma growth in nude mice	Georges *et al.* (1993)
NF-κB	ODN	Regression of HTLV-1 Tax-transformed fibrosarcomas in mice	
c-*myb*	ODN; immunoliposomes	Reduced growth of CML and neuroblastoma *in vitro* and of melanoma *in vivo*	Calabretta *et al.* (1991) Pagnan *et al.* (2000)
c-*raf*	ODN	Reduced growth of human bladder and breast tumour xenografts	Monia *et al.* (1996)
ILGF-R	Stable transformant	Inhibition of tumorigenicity of He La cells	
EGF-R	ODN; folate–PEG liposomes	Inhibition of keratinocyte proliferation *in vitro*	
HER-2/*neu*	ODN	Inhibition of breast cancer cell growth; induction of apoptosis	
Cyclin D1	Stable transformant	Decreased cell growth and tumorigenicity; sensitization to cisplatin	
VEGF	Stable transformant	Diminished melanoma cell growth, angiogenesis and metastasis	
Protein kinase A, Cα	ODN	Reduced tumour cell growth *in vivo*; synergy with chemotherapeutic agents	Dean and McKay (1994)
bcl-2	ODN	Reduced growth of lymphomas in SCID mice; improved response of melanoma to dacarbazine *in vivo*	Jansen *et al.* (1998)

Table 2 Use of ribozyme therapy in preclinical models of human cancer

Target gene	Ribozyme/delivery method	Results	Reference
c-*fos*	HH; stable transformant; inducible promoter	Reversal of cisplatin resistance; elucidation of downstream DNA repair pathway	Scanlon *et al.* (1991)
H-*ras*/N-*ras*	HH; transformants, retroviral and adenoviral delivery	Decreased *ras* gene expression in tumours; reversed invasive phenotype; decreased growth of xenografts, melanoma and transformed 3T3 cells	Kashani-Sabet *et al.* (1994)
BCR-ABL	HH; exogenous delivery	Diminished expression and *in vitro* cell growth	Snyder *et al.* (1993)
Mdr-1	HH; stable transformant	Reversal of MDR phenotype	
HER-2/*neu*	HH; adenoviral delivery; stable transformant	Reduced protein, *in vitro* and *in vivo* cell growth	Suzuki *et al.* (2000)
CAPL/mts 1	HH; stable transformant	Proliferation-independent inhibition of metastatic activity	
EGFR	HH; stable transformant	Targeting splice site mutation reduces tumour growth in nude mice	
FGFBP	HH; stable transformant	Reduced angiogenesis and growth of xenograft tumours	
Pleiotrophin	HH; stable transformant	Reduced angiogenesis and metastasis in nude mice	Czubayko *et al.* (1996)
MMP-9	HH; stable transformant	Decreased lung metastasis; cell growth unaffected	
α6 integrin	HH; stable transformant	Reduced adherance, invasion and metastasis	
EBNA	HH; adenoviral	Blocked virus-induced B cell proliferation	

HH: hammerhead.

Ribozymes

Ribozyme targeting of genes associated with cancer shares some common features with antisense, but interesting differences also exist (see **Table 2**). The most obvious area of overlap is the use of ribozymes targeting oncogenes. The most extensively studied system includes ribozymes targeting the activated H-*ras* gene at codon 12, in which the GUC sequence in the tumour (encoding valine) is recognized by ribozyme cleavage, whereas the wild-type GGC sequence (encoding glycine) is not. Ribozymes targeting mutated *ras* genes have been investigated in a number of

preclinical model systems (Kashani-Sabet *et al.*, 1994). Delivery of anti-*ras* ribozymes has been accomplished by both retroviruses and adenoviruses.

BCR-ABL is another gene that has been extensively analysed in ribozyme studies. Several investigators have examined the exogenous delivery of modified ribozymes targeting *BCR-ABL* for CML (Snyder *et al.*, 1993). Finally, other oncogenes targeted by ribozymes are listed in **Table 2**.

The second area of overlap with antisense and ribozyme targeting is in reversal of drug resistance, although the targets utilized have differed. The first demonstration of ribozyme targeting in cancer showed the reversal of resistance to cisplatin in ovarian carcinoma cells (Scanlon *et al.*, 1991). The use of the anti-*fos* ribozyme was subsequently broadened to cover the multidrug-resistant (MDR) phenotype also. Along these lines, several investigators demonstrated the utility of hammerhead ribozymes targeting the *mdr-1* gene to reverse resistance to a variety of chemotherapeutic agents associated with the MDR phenotype.

The third common category concerns targeting angiogenic genes. Once again, the genes selected for ribozyme targeting differ from those discussed in the antisense section. Hammerhead ribozymes have been shown to reduce angiogenesis by targeting the pleiotrophin gene (Czubayko *et al.*, 1996) and the gene encoding the basic fibroblast growth factor-binding protein. In the case of pleiotrophin, melanoma cell lines expressing the ribozyme had a reduced capacity for metastasis. Recently, continuous infusion of a synthetic hammerhead ribozyme oligonucleotide targeting the VEGF receptor flt-1 (termed ANGIOZYME), was shown to produce anti-metastatic effects in murine models (Pavco *et al.*, 2000).

Antimetastatic activity is another well-documented use of hammerhead ribozymes. This potential has been amply demonstrated in studies using ribozymes to target the *CAPL/Mts 1* gene, the integrin gene $\alpha 6$ and the matrix metalloproteinase *MMP-9* gene.

Finally, two additional targets of ribozyme therapy not previously discussed with antisense include cancer-associated viral genes and telomerase. Ribozymes have been shown to suppress the transformed phenotype associated with the human papillomavirus *E6* and *E7* genes and the Epstein–Barr virus nuclear antigen-1. Finally, several groups have reported on the use of ribozymes targeting telomerase sequences. However, to date, no effects on the neoplastic phenotype have been reported.

CLINICAL TRIALS OF ANTISENSE AND RIBOZYMES

As alluded to in the preceding section, with the completion of preclinical studies using antisense ODNs and ribozymes in a number of model systems, the stage was set for these compounds to enter the clinical arena. In fact, a number of studies using antisense ODNs in cancer have been completed, and many others are currently under way (see **Table 3**). Intravitreal administration of an antisense ODN targeting the cytomegalovirus genome (CMV) has been approved by the US Food and Drug Administration for CMV-induced retinitis associated with HIV infection. Ribozymes are also entering the clinical arena, as a clinical trial of hairpin ribozyme is under way in the treatment of HIV infection. In this section, the results of antisense studies described thus far will be reviewed, and the potential of ribozymes as anticancer therapeutics will be further explored.

The first report of a clinical trial of antisense therapy in cancer involved the use of an antisense ODN targeting *p53* (Bishop *et al.*, 1996). In this phase I trial, 16 patients with advanced haematological malignancies (either refractory acute myelogenous leukaemia or advanced myelodysplastic syndrome) were treated by continuous intravenous infusion of a 20-mer phosphorothioate oligonucleotide complementary to *p53* mRNA for 10 days. Five dose levels, ranging from 0.05 to $0.25\,\mathrm{mg\,kg^{-1}\,h^{-1}}$ were examined. No specific toxicity was found to be directly related to the administration of the antisense ODN. Pharmacokinetic studies revealed that plasma drug concentrations increased with dose, and that 36% of the administered dose was recovered in the urine. While there was diminished *in vitro* expansion of leukaemic cells following ODN therapy, no clinical complete responses were observed. A second study examined the pharmacokinetic profile

Table 3 Antisense ribozyme compounds in clinical trials of cancer

Target gene	Target cancer	Phase	Reference/status
c-*raf*	Multiple	I; II	Stevenson *et al.* (1999)
p53	Haematological malignancies	I	Bishop *et al.* (1996)
bcl-2	Non-Hodgkin lymphoma; prostate cancer; melanoma	I; II	Webb *et al.* (1997) Ongoing
c-*myb*	CML	I	
PKC-α	Multiple	I	Ongoing
	Ovarian	II	
H-*ras*	Solid tumours	I	Ongoing
flt-1	Breast cancer, melanoma	II	Ongoing

of an antisense ODN targeting intercellular adhesion molecule-1. A phase I dose escalation study was conducted in volunteers with single or multiple intravenous infusions of ISIS 2302. Laboratory abnormalities included dose-related increases in activated partial thromboplastin time and insignificant increases in C3a. The plasma half-life was found to be 53–54 min for intact drug and 67–74 min for total oligonucleotide. No adverse effects were related to drug treatment.

Subsequently, the results of two phase I trials of antisense ODNs targeting *bcl-2* and c-*raf*-1 have been published at the time of this writing. The preliminary results of several other trials have been reported in abstract form. In the case of *bcl-2* antisense ODN, a phase I study examining daily subcutaneous infusion of an 18-mer phosphorothioate oligonucleotide for 2 weeks was performed in five patients with bcl-2-positive non-Hodgkin lymphoma (Webb *et al*., 1997). This dose escalation study investigated doses ranging from 4.6 to 73.6 mg/m^2 . With the exception of local inflammation at the injection sites, no treatment-related adverse events were noted. One patient had a complete response of nodal disease lasting 4 months. Another patient had improved tumour burden that did not qualify as a partial response. Two patients had improvements in their constitutional symptoms with antisense ODN treatment, and several patients had a reduction in the number of circulating lymphoma cells or serum concentration of lactate dehydrogenase. Importantly, Bcl-2 expression was assayed using flow cytometry, and was reduced in the peripheral blood and bone marrow aspirates of one patient and in the lymph node aspirates of another patient (Webb *et al*., 1997).

More recently, the results of a phase I clinical, pharmacokinetic and pharmacodynamic trial of ISIS 5132 targeting c-*raf*-1 kinase have been reported (Stevenson *et al*., 1999). Thirty-one patients with advanced malignancies underwent a dose-escalation trial of a 20-base phosphorothioate antisense ODN complementary to 3′ untranslated sequences of c-*raf*-1 mRNA with nine dose levels ranging from 0.5–6.0 $mg\,kg^{-1}$ as a 2-h intravenous infusion three times weekly for three consecutive weeks. In this study, a maximum tolerated dose was not achieved, and toxicities included fever and fatigue. Laboratory toxicities included elevations in the activated partial thromboplastin time and levels of C3a. The potential for complement activation resulted in closure of the study at a dose of 6.0 $mg\,kg^{-1}$. One patient with metastatic colorectal carcinoma was observed to have minor response in a liver metastasis. Another patient with metastatic renal cell carcinoma had stable disease during 10 courses of treatment prior to progressive disease.

Pharmacokinetic studies, similar to prior studies, revealed a linear increase in mean plasma concentrations with increasing dose. The mean half-life of ISIS 5132 was 59.8 min. In this study, patients' circulating peripheral blood mononuclear cells were used as a surrogate to study expression of c-*raf*-1 during ISIS 5132 administration. Reductions of c-*raf*-1 expression were noted in 13 of 14 patients by day three of treatment, with a median decrease of 42% of initial values. Interestingly, higher doses of antisense ODN did not result in more prolonged gene expression. In the two patients with possible clinical benefit noted previously, c-*raf*-1 mRNA levels were reduced during the period of disease stability, but became elevated with disease progression.

A few reports have described the preliminary experience with antisense ODNs in other oncological settings. A study of antisense c-*myb* ODN in 18 patients with refractory leukaemia showed minimal toxicity and one response in a patient with CML in blast crisis who reverted to chronic phase. Two trials of antisense ODN targeting protein kinase Cα to patients with refractory cancer have been performed. Early analysis of these trials shows tolerability of the antisense therapy, and clinical responses were observed in ovarian cancer patients. This has resulted in a phase II trial in ovarian carcinoma. More recently, a trial of G3139, an antisense ODN targeting *bcl-2* mRNA combined with dacarbazine in patients with metastatic melanoma showed responses in three of 14 evaluable patients. Study treatment was associated with transient fever, flushing, rash and liver function abnormalities. Reduction of Bcl-2 protein in tumour biopsy samples and induction of apoptosis of melanoma cells was observed.

Taken together, the studies performed to date allow a number of conclusions to be drawn. First, antisense ODN therapy appears to be remarkably safe, even in the setting of advanced malignancies. Several dosing and administration schedules have been analysed, and further optimization is required. Antisense ODN therapy appears to have a short half-life and a fairly predictable pharmacokinetic profile. Demonstration of inhibition of gene expression has been provided in a few instances, but needs to be examined more closely. To date, the presence of nonantisense effects in humans is unknown. Transient clinical benefits and/or responses have been observed. These studies suggest the readiness of antisense ODNs for phase II trials, in which a more accurate assessment of clinical response is possible. Moreover, no conclusions can yet be drawn regarding the specific genes targeted. Furthermore, the exact nature of the malignancies that would represent the most suitable targets for antisense therapy has yet to be defined. Nevertheless, the results amassed to date are encouraging to permit the continued clinical maturation of antisense technology.

With respect to ribozymes, further optimizations in gene delivery techniques will probably be required prior to initiation of widespread clinical testing. Recent studies using systemic delivery and expression of ribozymes using cationic lipid–DNA complexes with vectors that permit sustained transgene expression to tumour-bearing animals offer a possible improvement to currently existing ribozyme therapeutics for cancer.

NEW USES FOR ANTISENSE AND RIBOZYMES IN CANCER

In addition to the use of antisense/ribozymes to down-regulate gene expression for therapeutic purposes, novel uses of these compounds are emerging, with potential applications to human cancer. One intriguing use of these compounds is in the functional genomics of cancer. With the completion of the human genome project, a myriad of candidate genes will emerge with possible roles in the initiation, progression, prevention, and/or treatment of human cancer. The use of antisense and/or ribozymes to suppress a target gene of interest in tumour cell lines or in tumour-bearing animals may represent a powerful strategy to define the role of previously uncharacterized genes in cancer. Second, antisense/ribozyme technology has been adapted to achieve repair of mutated genes. While this represents an evolving technology at the present time, the theoretical appeal to the treatment of the primary genetic defects of human disease (including cancer) are obvious.

CONCLUSION

In this chapter we have reviewed the evolution of antisense and ribozymes from tools in molecular biology to tools in the clinical management of cancer. It is clear that there is enormous potential in antisense/ribozyme therapy, but also that there are significant pitfalls and obstacles yet to be overcome. However, these technologies are being increasingly tested in clinical trials of human cancer. The coming decade is sure to see the continued progress and development of antisense and ribozymes as therapeutics.

REFERENCES

Benimetskaya, L., *et al*. (1997). Mac-1 (CD11b/CD18) is an oligodeoxynucleotide-binding protein. *Nature Medicine*, **3**, 414–420.

Bishop, M. R., *et al*. (1996). Phase I trial of an antisense oligonucleotide OL (1) p53 in hematologic malignancies. *Journal of Clinical Oncology*, **14**, 1320–1326.

Burgess, T. L., *et al.* (1995). The antiproliferative activity of c-*myb* and c-*myc* antisense oligonucleotides in smooth muscle cells is caused by a nonantisense mechanism. *Proceedings of the National Academy of Sciences of the USA*, **92**, 4051–4055.

Calabretta, B., *et al*. (1991). Normal and leukemic hematopoietic cells manifest differential sensitivity to inhibitory effects of c-*myb* antisense oligodeoxynucleotides: An *in vitro* study relevant to bone marrow purging. *Proceedings of the National Academy of Sciences of the USA*, **88**, 1251–2355.

Czubayko, F., *et al*. (1996). Melanoma angiogenesis and metastasis modulated by ribozyme targeting of the secreted growth factor pleiotrophin. *Proceedings of the National Academy of Sciences of the USA*, **93**, 14753–14758.

Dean, N. M. and McKay, R. (1994). Inhibition of protein kinase C-alpha expression in mice after systemic administration of phosphorothioate antisense oligodeoxynucleotides. *Proceedings of the National Academy of Sciences of the USA*, **91**, 11762–11766.

Georges, R. N., *et al*. (1993). Prevention of orthotopic human lung cancer growth by intratracheal instillation of retroviral antisense K-*ras* construct. *Cancer Research*, **53**, 1743–1746.

Guerrier-Takada, C., *et al*. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.

Haseloff, J. and Gerlach W. L. (1988). Simple RNA enzymes with new and highly specific endoribonuclease activity. *Nature*, **334**, 585–591.

Jansen, B., *et al*. (1998). *bcl-2* antisense therapy chemosensitizes human melanoma in SCID mice. *Nature Medicine*, **4**, 232–234.

Kashani-Sabet, M., *et al*. (1994). Suppression of the neoplastic phenotype *in vivo* by an anti-*ras* ribozyme. *Cancer Research*, **54**, 900–902.

Krieg, A. M., *et al*. (1995). CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature*, **374**, 546–549.

Monia, B. P., *et al*. (1996). Antitumor activity of a phosphorothioate antisense oligodeoxynucleotide targeted against c-*raf* kinase. *Nature Medicine*, **2**, 668–675.

Ojwang, J. O., *et al*. (1992). Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme. *Proceedings of the National Academy of Sciences of the USA*, **89**, 10802–10806.

Pagnan, G., *et al*. (2000). Delivery of c-*myb* antisense oligodeoxynucleotides to human neuroblastoma cells via disialoganglioside GD2-targeted immunoliposomes: antitumor effects. *Journal of the National Cancer Institute*, **92**, 253–261.

Pavco, P. A., *et al*. (2000). Antitumor and antimetastatic activity of ribozymes targeting the messenger RNA of vascular endothelial growth factor receptors. *Cancer Research*, **6**, 2094–2103.

Sarver, N., *et al*. (1990). Ribozymes as potential anti-HIV-1 therapeutic agents. *Science*, **247**, 1222–1225.

Scanlon, K. J., *et al*. (1991). Ribozyme-mediated cleavage of c-*fos* mRNA reduces gene expression of DNA synthesis enzymes and metallothionein. *Proceedings of the National Academy of Sciences of the USA*, **88**, 10591–10595.

Snyder, D. S., *et al*. (1993). Ribozyme-mediated inhibition of bcr-abl gene expression in a Philadelphia chromosome-positive cell line. *Blood*, **82**, 600–605.

Stein, C. A. and Krieg, A. M. (1994). Problems in interpretation of data derived from *in vitro* and *in vivo* use of antisense oligodeoxynucleotides (Editorial). *Antisense Research and Development*, **4**, 67–69.

Stevenson, J. P., *et al*. (1999). Phase I clinical/pharmacokinetic and pharmacodynamic trial of the c-*raf*-1 antisense oligonucleotide ISIS 5132 (CGP 69846A). *Journal of Clinical Oncology*, **17**, 2227–2236.

Suzuki, T., *et al*. (2000). Adenovirus-mediated ribozyme targeting of HER-2/*neu* inhibits *in vivo* growth of breast cancer cells. *Gene Therapy*, **7**, 241–248.

Szczylik, C., *et al*. (1991). Selective inhibition of leukemia cell proliferation by BCR-ABL antisense oligonucleotides. *Science*, **253**, 562–565.

Webb, A., *et al*. (1997). *BCL-2* antisense therapy in patients with non-Hodgkin lymphoma. *Lancet*, **349**, 1137–1141.

Zamecnik, P. C. and Stephenson, M. L. (1978). Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proceedings of the National Academy of Sciences of the USA*, **75**, 280–284.

Zaug, A. J. and Cech, T. R. (1986). The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science*, **231**, 470–475.

FURTHER READING

Bratty, J., *et al*. (1993). The hammerhead RNA domain, a model ribozyme. *Biochimica Biophysica Acta*, **1216**, 345–359.

Flanagan, W. M. (1998). Antisense come of age. *Cancer Metastasis Reviews*, **17**, 169–176.

Gewirtz, A. M., *et al*. (1996). Facilitating oligonucleotide delivery: helping antisense deliver on its promise. *Proceedings of the National Academy of Sciences of the USA*, **93**, 3161–3163.

Herschlag, D. (1991). Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules *in vivo:* more isn't always better. *Proceedings of the National Academy of Sciences of the USA*, **88**, 6921–6925.

Kashani-Sabet, M. and Scanlon, K. J. (1995). Applications of ribozymes to cancer gene therapy. *Cancer Gene Therapy*, **2**, 213–223.

Scanlon, K. J. and Kashani-Sabet, M. (eds), (1998). *Ribozymes in the Gene Therapy of Cancer.* (R. G. Landes, Austin, TX).

Scanlon, K. J. and Kashani-Sabet, M. (1998). Ribozymes as therapeutic agents: are we getting closer? *Journal of the National Cancer Institute*, **90**, 558–559.

Stein, C. A. (1995). Does antisense exist? *Nature Medicine*, **1**, 1119–1121.

Stein, C. and Cheng, Y. (1993). Antisense oligonucleotides as therapeutic agents – is the bullet really magical? *Science*, **261**, 1004–1012.

Symonds, R. H. (1992). Small catalytic RNAs. *Annual Review of Biochemistry*, **61**, 641–667.

Wagner, R. W. (1994). Gene inhibition using antisense oligodeoxynucleotides. *Nature*, **372**, 333–335.

Chapter 92

Hormonal Therapy for Breast Cancer

V. Craig Jordan and William J. Gradishar

Robert H. Lurie Comprehensive Cancer Center, Northwestern University Medical School, Chicago, IL, USA

CONTENTS

- Introduction
- Breast Cancer: An Introduction
- Hormones and Breast Cancer
- ERSα and -β as Targets for Therapy
- Current Strategies for the Endocrine Treatment of Breast Cancer
- Inhibition of Oestrogen Synthesis
- Blocking Oestrogen Action
- Selective Oestrogen Receptor Modulation
- Conclusions

INTRODUCTION

Remarkable advances have been made during the past 40 years in the treatment of childhood leukaemia. However, the application of combination chemotherapy for the treatment of solid tumours has met with only modest success. The one exception is the treatment of breast and prostate cancer. These diseases are major killers of women and men, respectively.

Advances in the treatment of breast and prostate cancer have occurred primarily because of changes in the way treatment strategies have evolved. The treatment of both diseases depends on surgery, staging, radiation therapy and combination chemotherapy. Nevertheless, unlike other solid tumours in humans, breast and prostate cancer are dependent upon oestrogens or androgens, respectively for growth. The tumour cells contain the relevant steroid hormone receptor that can harness circulating steroids to cause uncontrolled cell replication and tumour growth. As a result of knowledge of the link between steroid hormones and breast and prostate tumour growth, treatment strategies to deny access of steroids to the tumour have been investigated and shown to provide benefits for the patient. In 1966, Professor Charles B. Huggins received the Nobel Prize for Medicine for his work on the hormone therapy of prostate cancer.

In its simplest form, hormonal therapy initially was the removal of the ovaries or the testes. Endocrine ablation caused tumour regression or stabilization of disease progression in a majority of patients. However, the discovery and understanding of steroid hormone receptor mechanisms provided an invaluable target for drug discovery and rational therapeutics. Although the principles for the treatment of breast and prostate cancer are similar, breast cancer research has led the way with discoveries and innovative therapeutic strategies. In this chapter, we will consider breast cancer treatment as the model for hormonal therapy for two major reasons. First, the application of the antioestrogenic drug tamoxifen has been shown to benefit millions of women with all stages of breast cancer, and it is calculated that 400 000 women are alive today because of the use of the drug following surgery. Second, tamoxifen is a pioneering medicine as it is the first drug to be available to prevent breast cancer in high-risk women. There is no other therapeutic intervention that is approved to prevent any other cancer in people at risk.

BREAST CANCER: AN INTRODUCTION

Breast cancer is diagnosed in approximately 176 000 patients in the USA each year. Even more striking is the prevalence of breast cancer in the population, which exceeds other solid tumours such as lung cancer, colon cancer and prostate cancer. Not only are there a large number of new diagnoses of breast cancer each year, but there are also an enormous number of breast cancer survivors who remain at risk for recurrence of the disease or the development of a new primary breast cancer.

Breast cancer can be broadly divided into various stages. Advanced disease presents with clearly identified metastases at distant sites. In node-positive breast cancer,

tumour is identified by the pathologist in the axial lymph nodes sampled during surgery. A patient who is node negative at presentation has a good prognosis with a $>80\%$ chance of survival depending on the size of the original primary tumour. However, the routine use of regular mammography has resulted in an increased diagnosis of ductal carcinoma *in situ* (DCIS). This is non-invasive breast cancer that is currently treated with surgery, radiation and the antiestrogen tamoxifen (Fisher *et al.*, 1999).

During the past 30 years, breast cancer therapy has evolved from the treatment of advanced disease to the administration of adjuvant therapy for the eradication of microscopic disease. Systemic treatment options that are available to breast cancer patients include chemotherapy, hormonal therapy and biological therapy. However, results from chemoprevention studies now offer risk reductions for women with elevated risk factors for the disease (Fisher *et al.*, 1998). Indeed, the widespread availability of agents with the potential to prevent disease (e.g. chemoprevention) has introduced a new dimension into women's health care. (See the chapters on *Breast* and *Chemoprevention*.)

HORMONES AND BREAST CANCER

The relationship between endocrine manipulation and breast cancer was first established over 100 years ago. In 1896, it was first shown that the removal of the ovaries from a premenopausal woman with advanced breast cancer resulted in tumour regression. However, surgical oophorectomy was not found to be beneficial for all premenopausal patients. Only about one-third of premenopausal women with metastatic breast cancer showed objective evidence of tumour regression, and these women gained approximately 1 year of additional life. A similar effect on breast cancer was found to occur by irradiating the ovaries. It was later found that the effect of radiation on ovarian function correlated with decreases in oestrogen production that reached basal levels several months following treatment.

Once synthetic corticosteroids became available in the 1950s, other forms of endocrine ablation became feasible, including surgical adrenalectomy and hypophysectomy. Tumour responses were reported in 30–40% of patients undergoing these procedures, but the associated morbidity and mortality were unacceptably high. Eventually, medical therapies for inducing an adrenalectomy were identified with drugs such as aminoglutethimide (see later under Aromatase Inhibitors). Although much of the early work focused on ways of achieving oestrogen deprivation, paradoxically high-dose oestrogen therapy resulted in tumour regression in postmenopausal women with advanced breast cancer. This became the treatment of choice for postmenopausal women during the 1950s and early 1970s before the antioestrogen tamoxifen became available (see later under Tamoxifen for Treatment). The link between ovarian hormones and breast cancer growth was established in the first half of the twentieth century, but only one patient in three with advanced breast cancer responded. The problem was to determine prospectively who would benefit from endocrine ablation.

Through a series of elegant experiments, Jensen and Jacobsen hypothesized that an oestrogen receptor (ER) must be present in oestrogen-target tissues, such as the uterus, vagina and pituitary (Jensen and Jacobsen, 1962) that would control the response of the tissue selectively. Ultimately, the ER was isolated and characterized as an extractable protein from oestrogen-target tissues (**Figure 1**). These findings led to the development of ER assays that could be used to predict which of the patients with metastatic breast cancer would respond to endocrine manipulation. In a random population of patients with metastatic breast cancer, 30–35% will respond to endocrine manipulation, whereas 55–60% of patients with tumours known to be ER-positive respond to endocrine manipulation. Several studies have shown that older patients diagnosed with breast cancer are more likely than younger patients to express ER. About 44% of breast cancer patients in the age group 40–49 years express ER, whereas 69% of patients 70 years or older express ER.These data suggest that most patients with metastatic

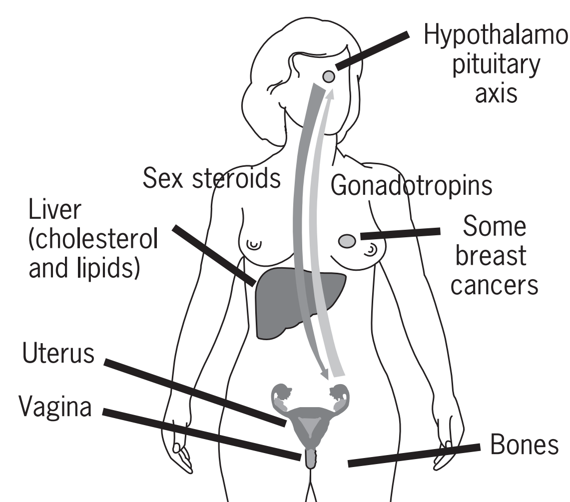


Figure 1 The oestrogen target tissues identified around a woman's body. The menstrual cycle is controlled by the negative feedback system that operates between ovarian steroid synthesis and gonadotropin release through the hypothalamopituitary axis. The pulsatile release of luteinizing hormone-releasing hormone (LHRH) from the hypothalamus causes the release of gonadotropins from the pituitary gland. Oestrogen receptors are found in a majority of breast cancers.

Mer 25

Tamoxifen

Figure 2 The formulae of nonsteroidal antioestrogens that block the actions of oestrogen in its target tissues. Antioestrogens are competitive inhibitors of oestradiol binding at the ER.

breast cancer are eligible for endocrine manipulation at some point in the course of the disease.

In 1958, MER-25 (**Figure 2**), the first nonsteroidal antioestrogen, was discovered, but because of toxicity in the central nervous system and relatively low potency, further development was discontinued. This class of drugs, however, block the binding of oestradiol to the ER so they had potential in treatment for breast cancer. Tamoxifen, the antioestrogenic, pure *trans* isomer of a substituted triphenylethylene, was initially developed as a potential postcoital contraceptive, but was eventually shown to induce ovulation in subfertile women. Fortunately, preliminary clinical studies of tamoxifen in patients with metastatic breast cancer showed antitumour activity, leading to its development as a treatment of breast cancer (Lerne and Jordan, 1990).

ERSα AND -β AS TARGETS FOR THERAPY

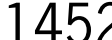
The biological effects of oestrogen are now known to be mediated by two receptors, referred to as ERα and ERβ. ERα was first identified and isolated in the mid-1960s. The cloning of ERα in 1986 focused research efforts on the existence of only one ER protein (ERα) identical in all target tissues. However, almost three decades after the initial discovery of ERα, ERβ was identified in the rat, human and mouse. The discovery of ERβ has already advanced our understanding of oestrogen signalling and may explain the responses to oestrogen in tissues in which ERα was not detectable. The role of ERα and -β in physiology is currently being deciphered through the use of 'knockout' transgenic mice that have ERα, ERβ or both receptors knocked out (Couse and Korach, 1999).

The two ERs share functionally conserved structure (A–F domains) (**Figure 3**) with a variable N-terminal region involved in transactivation function (A/B), a centrally located, well conserved DNA-binding domain (C) and C-terminal region involved in dimerization and in binding to heat shock protein (HSP) 90 (D), ligand-binding domain (LBD) (E) with transactivation functions and F region. Region F appears to play a role in modulating transcriptional activation by ERα. ERβ is homologous to ERα, at the ligand-binding (58%) and DNA-binding (96%) domains, whereas the A/B region, hinge domain and F region are not well conserved.

ERα contains two activation functions (AFs) which contribute to the ER's transcriptional activity. ERα contains an activating function only in the ligand-binding domain but has an AF-1 site that is functionally very different to ERα. AF-1 of ERα is located in the N-terminal region within the A/B region and was originally believed to be constitutively active and ligand independent. It has been demonstrated that ligand-independent activation of ER via the AF-1 domain is closely related to the phosphorylation status of the receptor. In particular, Ser118 in the AB region of ERα is important for the activation through the Ras-MAP kinase (MAPK) signalling cascade and Ser106 and Ser124 are two phosphorylation sites in the A/B region of ERβ which are essential for ligand-independent activation of the ERβ via the MAPK cascade. In addition, both receptors contain a second activation domain, AF-2, which is present at the C-terminus and is ligand dependent. Mutational analyses demonstrate the importance of this region for ER transactivation because AF-2 can interact with a number of putative transcriptional coactivators in a ligand-dependent manner.

AF-1 and AF-2 can activate transcription independently but in most cases they synergize with one another in a promoter- and cell context-specific manner. It is believed that ER activates gene expression by binding to oestrogen response elements (EREs) in responsive genes through the synergistic action of AF-1 and AF-2. ERβ was also shown to activate transcription of target genes through EREs, but there is evidence that ERβ may reduce the actions of low levels of oestrogen. However, it has been demonstrated recently that although oestrogen can induce an AP-1 site in a reporter construct through ERα, it is inactive via ERβ. Interestingly, ER antagonists activate ERβ to induce activity through an AP-1 site and through the human retinoic acid receptor α-1 promoter.

The two receptors ERα and ERβ may form functional heterodimers on DNA that could bind the steroid receptor coactivator SRC-1 and stimulate transcription of a target gene. The ability of ERα and ERβ to form heterodimers suggests that ER may function through different dimeric states, and it is possible that the dimers could be activated by selective ligands.



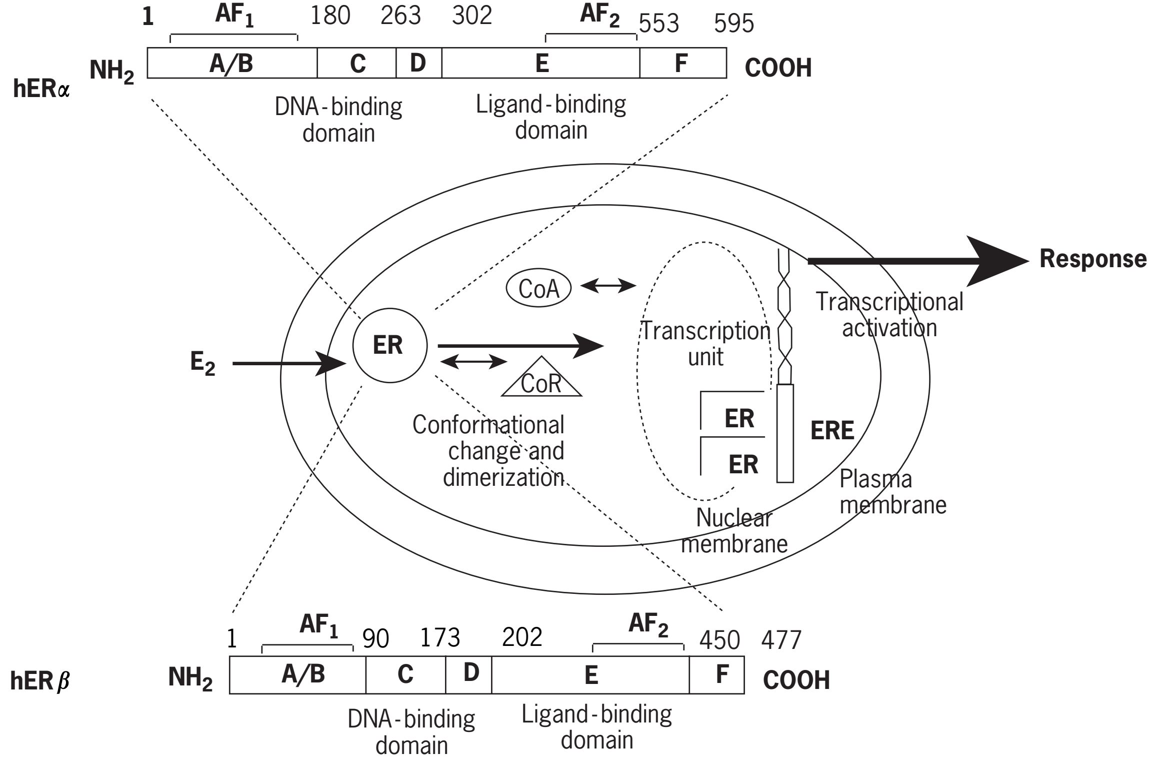


Figure 3 The subcellular mechanism of action of oestrogen in a target tissue, i.e. ER-positive breast cancer. Oestrogen action is regulated by two oestrogen receptors: the classical ER now referred to as ERα and the newly described oestrogen receptor ERβ. Once oestrogen binds to the receptor, the complex changes shape and dimerizes before binding to oestrogen-response elements (EREs) in the promoter region of an oestrogen responsive gene. Gene translation occurs by the binding of coactivator (CoA) molecules to the ER complex that contribute to the transcription complex. RNA polymerase and DNA unwinding enzymes are recruited to ensure gene transcription. Antioestrogen ER complexes could incorporate novel corepressor (CoR) molecules to block gene transcription. (From Levenson and Jordan, 1999, *European Journal of Cancer*, **35**, 1628–1629.)

A Molecular Mechanism for Oestrogen Action

The existence of two rather than one ER indicates that the mechanism of action of oestrogen and antioestrogens is even more complex than previously thought. Oestrogen, upon binding to its high-affinity receptor (or receptors), triggers the expression of multiple genes involved in the regulation of cell proliferation and differentiation. The binding to the agonist allows the ER to dissociate from heat-shock protein (HSP 90), dimerize, bind to specific DNA sequences and stimulate transcription of responsive genes. Now it is well known that ER itself is far from being the direct controller of transcription: it requires an interaction with a complex of coregulatory proteins (coactivators or corepressors), which act as signalling intermediates between the ER and the general transcriptional machinery (**Figure 3**). Recently, the crystal structure of the LBD of ERα was determined with oestradiol and the synthetic nonsteroidal oestrogen diethylstilboestrol (Brzozowsky *et al*., 1997; Shiau *et al*., 1998). Both high-affinity ligands interact with the same amino acids to locate the oestrogens correctly (**Figure 4**). A key feature of an oestrogen is the ability to be enveloped in an hydrophobic pocket that is closed by helix 12 (an essential site for AF-2 activation) in the ligand-binding site of ERα (**Figure 4**). Apparently, helix 12 is critical for the recruitment of coactivators to the AF-2 site and subsequent initiation of RNA polymerase activity. The repositioning of the helix 12 after ligand binding has been proposed as an important mechanism for full oestrogen action at ERα (Brzozowsky *et al*., 1997; Shiau *et al*., 1998).

Coactivators

There is considerable evidence that nuclear steroid receptors must associate with other nuclear proteins to form a transcription complex. For example, it has been shown that the AF-1 and AF-2 domains of ERα bind to TATA-binding protein (TBP) *in vitro*. Moreover, ER can interact with components of the transcription factor IID (TFIID) complex and TFIIB. Interactions between ER and

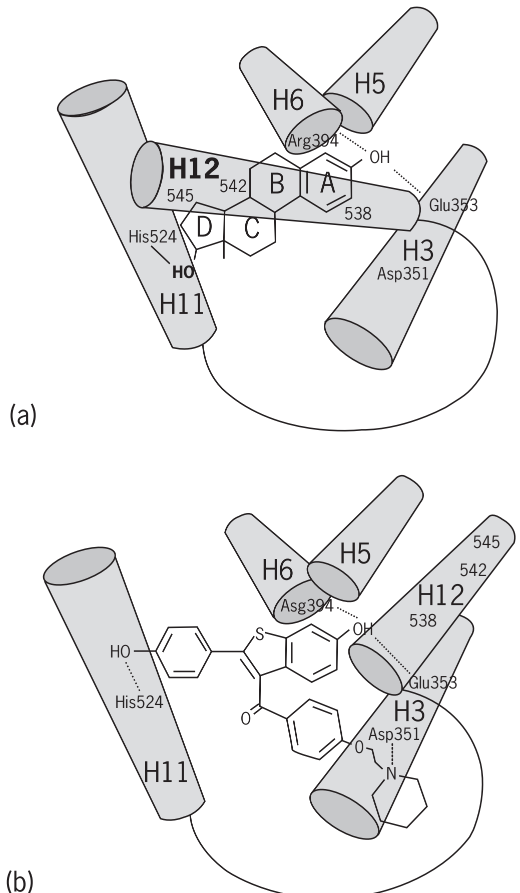


Figure 4 (a) The locking of oestradiol in the ligand-binding domain of ERα. Helix 12 seals the steroid into hydrophobic pocket and exposes sites on the complex that can bind to coactivators. (b) The wedging of raloxifene in the ligand binding domain of ERα. Helix 12 cannot seal the SERM into the hydrophobic pocket because the antioestrogenic side chain interacts with AA 351, which acts as a pivot for the helix. Helix 12 blocks the binding sites so coactivators cannot bind. (From Jordan and Morrow, 1999, *Endocrine Reviews*, **20**, 253–278.)

TFIID-associated factor (TAF), $TAF_{II}30$, have also been described. However, although all these interactions are necessary, they are not sufficient to mediate transcription. The multiprotein complex must not only provide the machinery to transcribe the appropriate gene but also facilitate a mechanism for exposing the quiescent gene. This process must destabilize the DNA–histone complex to allow RNA polymerase access to complete transcription.

Several coactivators have been described for ER based upon their ability to interact with agonist-bound receptor but not with an antagonist-bound receptor. ERAP 160 and RIP 140 were identified using the LBD of ER as bait. Interestingly, RIP 140 does not interact with basal factors of the transcription machinery, suggesting a mechanism other than a bridging function for a coactivator. SRC-1, on the other hand, was isolated from a human cDNA library using the LBD of hPgR as bait. However, studies indicate that SRC-1 and ERAP 160 are variants of the same family of proteins. The majority of receptor-interacting factors have been identified by using LBD of nuclear receptor as a bait. This, therefore, explains why the AF-2 site is believed to be the most important site for coactivators recruitment. However, It has been demonstrated that SRC family members are not only transcriptional mediators for the ligand-dependent AF-2 and AF-1 of ERα but are also involved in ligand-independent interaction with ERβ through phosphorylation of AF-1 via MAP-kinase cascade. Functional interaction of ER and SRC-1 occurs in the absence of exogenous ligand through cyclin D bridging. In addition, SRA (steroid receptor RNA activator), which is part of the SRC-1 complex, mediates transactivation via AF-1. Moreover, ERα/ERβ heterodimers were able to bind the SRC-1 and stimulate transcription of a reporter gene.

TIF2 belongs to the family of transcriptional intermediary factors (TIFs) which mediate AF activity to the transcription machinery. The SRC-1-related TIF2 also efficiently stimulates AF-2 activity of ER *in vivo*. Another protein involved in coactivator complex with SRC-1 is p300/CBP, a 300-kDa protein related to the cAMP response element-binding protein, CBP. It has been demonstrated that SRC-1 and p300/CBP contain intrinsic acetyltransferase activity and can interact with other histone acetyltransferases (HATs). Acetylation by SRC-1 of histones bound at specific promoters could be a mechanism by which the AFs of ER and associated coactivators activate transcription of specific genes by enhancing formation of a stable preinitiation complex.

A Molecular Mechanism for Antioestrogen Action

Antioestrogens are competitive inhibitors of oestrogen action. The shape of the ligand, with a strategically placed alkylaminoethoxy side chain, is essential to convert an oestrogenic complex to an antioestrogenic complex with reduced oestrogenic efficacy. Earlier models of antioestrogen action (proposed before the ER was cloned and sequenced) suggested that the antioestrogenic side chain prevents the receptor from closing so that the activation of the complex would be incomplete. The partial agonism of receptor complexes would occur because of an equilibrium mixture of activated and unactivated receptor complexes.

There is now convincing evidence that the shapes of different antioestrogen–ER complexes are not all the same. A mutated ER assay methodology in human liver cancer cells was used to classify new drugs based on their interaction with an engineered C_3 promoter target. However, it is the realization that a third component of the signal transduction system is involved that has provided the most excitement for understanding why some antioestrogens are

more oestrogen-like than others. Coactivators and possibly corepressors are key components that can modulate gene activation. Clearly, the shape of an antioestrogenic ER complex will dictate how or if any other protein in a transcription unit will bind.

The crystallization of the LBD of the ER with oestrogens and antioestrogens has provided enormous insight into the change in protein shape that prevents antioestrogens from building a transcription complex at AF-2. It is known that 4-hydroxytamoxifen silences AF-2 whereas AF-1 remains constitutively activated.

Both raloxifene and 4-hydroxytamoxifen fit into the hydrophobic pocket of the ligand-binding domain but the antioestrogenic side chain prevents the reorientation of helix 12 that must seal the ligand into the receptor before coactivators can bind and produce a transcription complex (Brzozowsky *et al*., 1997; Shiau *et al*., 1998). The high affinity ligands both interact, through phenolic hydroxyls, with Glu353 and Arg394 to locate the ligands correctly in the binding domain (**Figure 4**). However, the tertiary nitrogen side chain with a bulky phenyl group is critical for antioestrogenic activity because of an interaction with Asp351. This interaction is the key to oestrogen-like properties of the ER complex which may or may not subsequently bind coactivators (MacGregor Schafer *et al*., 2000). Although the X-ray crystallography of the raloxifene–and tamoxifen–ER complexes are similar (Brzozowsky *et al*., 1997; Shiau *et al*., 1998), there are significant differences in the interaction of the antioestrogenic side chains of raloxifene and 4-hydroxytamoxifen with amino acid 351. These differences have the potential to alter the degree of oestrogenic activity of the different antiestrogen–ER complexes (MacGregor and Jordan, 1998).

Corepressors

Although the experimental evidence suggests that activating functions can be prevented in the antioestrogen–ER complex by stopping coactivators from binding, an alternative hypothesis is an increased binding of proteins that interact only with the antiestrogen–ER complex. Corepressors would therefore neutralize an antiestrogen–ER complex.

CURRENT STRATEGIES FOR THE ENDOCRINE TREATMENT OF BREAST CANCER

The description of the target site specificity of oestrogen and the regulation of steroidogenesis by the hypothalamopituitary axis have been fundamental to devise current strategies that restrict the access of oestrogens to the tumour. It is now appropriate to summarize the principal therapeutic strategies that have proved to be successful endocrine approaches to control tumour growth.

Most of the work was originally focused on postmenopausal patients with metastatic breast cancer (**Figure 5**). Antioestrogens bind specifically to the ER in the breast tumour so the action of the group was a direct application of the emerging knowledge of oestrogen action in its target tissue during the 1960s (Jenson and Jacobson, 1962). In contrast, aminoglutethimide restricted the availability of circulating oestrogen in postmenopausal women. Aminoglutethimide blocks the biosynthesis of steroids in the adrenal glands and blocks the conversion of androstenedione to oestrone by aromatase enzymes in peripheral body fat (**Figure 5**). Although the approach of using aminoglutethimide was clinically successful, the incidence of side effects focused efforts to design more specific agents. The result is a variety of aromatase inhibitors that are specific for the enzyme at peripheral sites.

The antioestrogen tamoxifen has been proved to be effective in premenopausal patients (**Figure 6**) despite increases in circulatory oestrogen caused by interruption of the hypothalamopituitary axis feedback system. The sensitivity of the hypothalamopituitary axis to an *apparent* reduction in circulating oestrogen levels from the ovary causes a reflex rise in gonadotrophins. This is also the reason why aromatase inhibitors are not used in premenopausal women because the powerful action of

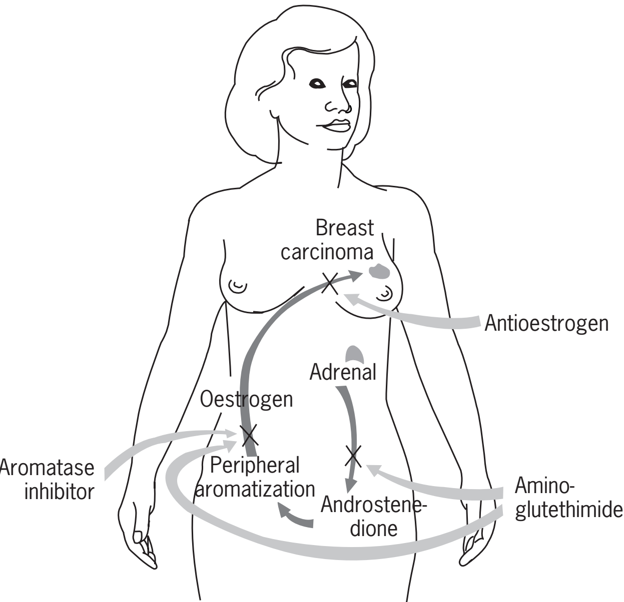


Figure 5 The evolution of strategies for the endocrine treatment of advanced breast cancer in postmenopausal patients. Antioestrogens block oestrogen action in the tumour. Aminoglutethimide blocks steroidogenesis in the adrenal glands and the conversion of androstenedione by aromatase enzymes in peripheral body fat. Aromatase inhibitors are now available that block the peripheral enzyme specifically.

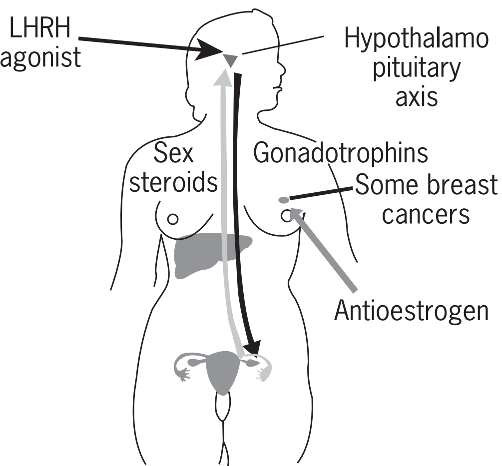


Figure 6 Premenopausal women with advanced breast cancer can be successfully treated with luteinizing hormone-releasing hormone (LHRH) agonists that causes a chemical oophorectomy by desensitizing the pituitary gland, thereby blocking gonadotrophin release. Additionally, tamoxifen blocks oestrogen binding to tumour ER. Combination therapy of the antihormones is proving to be a valuable new treatment strategy.

gonadotrophins would reverse ovarian aromatase blockade. Another strategy, that is proving to be successful in premenopausal patients, is the use of sustained release preparations of luteinizing hormone-releasing hormone (LHRH) to cause desensitization of the pituitary gland. As a result, the reduction in gonadotrophins causes a chemical oophorectomy. A combination of a sustained release preparation of an LHRH agonist and an antioestrogen will effectively decrease the availability of oestrogen to the tumour. This treatment approach is being evaluated at present.

INHIBITION OF OESTROGEN SYNTHESIS

Premenopausal Women – Luteinizing Hormone-Releasing Hormone (LHRH) Agonists

Castration was first reported as an effective therapy for metastatic breast cancer over 100 years ago, at a time when hormones had not been characterized. Several different means of causing ovarian ablation have been studied in the interim: surgical oophorectomy, radiation-induced ablation of the ovaries and, more recently, medical therapy using LHRH agonists.

LHRH is produced by the hypothalamus and stimulates the release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) from the anterior pituitary, which in turn stimulate the ovaries to synthesize oestrogen and progesterone. Under normal conditions, LHRH is released in a pulsatile manner, resulting in pulsatile release of LH and FSH. The continuous administration of an LHRH agonist overstimulates the LHRH receptors, causing an initial rise in LH and FSH. However, after 1–2 weeks the LHRH receptors become desensitized, leading to decreased release of FSH and LH. The ovaries respond by decreasing the synthesis of oestrogen and within 4 weeks of starting therapy with an LHRH agonist, circulating oestrogen levels are at a postmenopausal level. Unlike other ways of causing castration, serum oestradiol levels return to normal within 4 weeks of discontinuing the LHRH agonist.

Goserelin acetate (Zoladex®) is the only LHRH agonist approved by the US Food and Drug Administration (FDA) for the treatment of metastatic breast cancer in pre- and perimenopausal women. It is administered as a monthly, subcutaneous injectable implant (3.6 mg) that provides continuous release of the drug. Response rates ranging from 28 to 60% have been reported from phase II clinical trials of goserelin acetate in premenopausal women with metastatic breast cancer. In a randomized trial of premenopausal women with metastatic disease comparing goserelin acetate with surgical oophorectomy, with or without tamoxifen, no significant differences in response rates or survival were detected between treatment arms. Toxicities reported in women receiving goserelin acetate are those consistent with a menopausal state including hot flushes, loss of libido, vaginal dryness and decreased bone mineral density. Other LHRH agonists under investigation include buserelin and tryptorelin. The role of LHRH agonists as an adjuvant therapy for early-stage breast cancer is the subject of ongoing clinical trials.

Postmenopausal Women – Aromatase Inhibitors

The aromatase enzyme is the rate-limiting step in the conversion of androgens to oestrogens. Aromatase is an enzyme complex containing a cytochrome P-450 haemoprotein and a flavoprotein, nicotinamide adenine dinucleotide phosphate (NADPH) P-450 reductase. The cytochrome P-450 catalyses a series of three hydroxylations of the androgen substrates, androstenedione and testosterone. Since the aromatase cytochrome P-450 has little homology to other P-450 enzymes, it has been assigned to a separate gene family designated CYP19.

The major substrates for the aromatase enzyme are androstenedione (**Figure 7**) and testosterone, both produced principally in the adrenal glands, and to a lesser degree in the ovaries. Aromatase expression has also been demonstrated in the stromal compartment of some human breast tumours. This finding suggests that some breast tumours may have an innate ability to synthesize oestrogen, thereby stimulating growth locally. Whether this finding is clinically relevant is unknown. The aromatase

Suicide inhibitors

Exemestane

Formestane

Androgen substrate

Androstenedione

Nonsteroidal competitive inhibitors

Aminoglutethimide

Letrozole

Anastrazole

Figure 7 Classification of aromatase inhibitors.

enzyme selectively catalyses only the production of oestrogens, whereas other similar enzymes govern the production of other steroid hormones including glucocorticoids, androgens or mineralocorticoids. The optimum way of decreasing oestrogen production is to target *selectively* the aromatase enzyme without inhibiting the activity of the other enzyme systems.

Aromatase inhibitors are of two general types, competitive inhibitors and suicide inhibitors (**Figure 7**). Once a suicide inhibitor binds the aromatase enzyme, the sequence of hydroxylations is initiated, but the hydroxylations produce a covalent bond between the enzyme and inhibitor that irreversibly blocks the activity of the enzyme. Following exposure to a suicide inhibitor, the aromatase enzyme activity can only be restored with new enzyme synthesis. Suicide inhibitors include formestane and exemestane. Competitive inhibitors *reversibly* bind to the active enzyme site and either no enzyme activity is triggered or it has no effect. The inhibitor can dissociate from the binding site, allowing renewed competition between the inhibitor and substrate (i.e. androstenedione) for binding to the active binding site. For the effect of a competitive inhibitor to persist, a constant concentration of the inhibitor, in excess of the substrate, must be present. Competitive inhibitors include aminoglutethimide (nonselective), anastrazole and letrozole.

Competitive Aromatase Inhibitors

Aminoglutethimide

Aminoglutethimide (**Figure 7**) was the first widely used aromatase inhibitor for the treatment of advanced breast cancer. It caused a 'medical adrenalectomy' by blocking cholesterol side-chain cleavage, leading to a decrease in adrenal steroid synthesis. In addition, aminoglutethimide inhibits the peripheral aromatase system, resulting in decreased conversion of androstenedione to oestrogens. Adrenal synthesis of cortisol is inhibited by the action of aminoglutethimide, which can lead to clinical features of Addison disease including skin rash, fatigue, lethargy, ataxia and orthostatic hypotension. As a result, aminoglutethimide is usually administered with a supplemental corticosteroid, such as hydrocortisone.

The results from numerous clinical trials involving the combination of aminoglutethimide and replacement steroids show response rates of 20–56% in previously untreated patients with advanced breast cancer. The median duration of response is 11–12 months. Although described as previously untreated, many of the patients in these trials had received prior androgens, oestrogens or oophorectomy. Patients with ER–positive tumours have higher response rates (56%) compared with patients with ER-negative tumours (12%). Additionally, patients with sites of disease predominantly in the skin, soft tissues and bone were most likely to respond to aminoglutethimide.

Anastrazole

Anastrazole (Arimidex®) (**Figure 7**) is a potent, oral selective aromatase inhibitor that is approved as treatment for metastatic breast cancer in postmenopausal women who have failed tamoxifen. In animal studies, anastrazole was selective for the aromatase enzyme, while not inhibiting enzymes in the adrenal gland that control other

pathways such as mineralocorticoid or glucocorticoid synthesis. For all doses of anastrazole evaluated in phase I studies, the drug is well tolerated and does not effect cortisol or aldosterone secretion, at baseline or in response to ACTH. Consequently, glucocorticoid or mineralocorticoid replacement therapy is not required with anastrazole treatment.

The results from two mature phase III clinical trials have been reported. Anastrazole (1 or 10 mg per day) was compared to oral megesterol acetate (160 mg per day) in 764 postmenopausal patients with metastatic breast cancer that developed progressive disease following treatment with tamoxifen. Patients who received the 1 mg dose of anastrazole had a statistically significant survival advantage over patients receiving megestrol acetate (22% relative reduction in mortality). Patients receiving 1-mg of anastrazole had a longer median time to death (26.7 months) than patients receiving megesterol acetate (22.5 months). A nonsignificant survival advantage was observed for patients receiving the 10-mg dose of anastrazole compared with patients receiving megesterol acetate. Anastrazole was well tolerated at both dose levels and was not associated with the weight gain or dyspnoea observed in patients treated with megesterol acetate. The recommended daily dose of anastrazole is 1 mg. Anastrazole is as good as, if not better than, tamoxifen as a first-line therapy for metastatic breast cancer and is also under investigation as an adjuvant therapy.

Letrozole

Letrozole (Femara®) (**Figure 7**) is a potent, oral selective aromatase inhibitor, recently approved for the treatment of tamoxifen-resistant, advanced breast cancer in postmenopausal women. Phase I/II clinical trials showed that both 0.5- and 2.5-mg doses of letrozole suppressed circulating levels of oestrogen by over 90% and both doses were clinically effective. In a randomized, phase III clinical trial comparing letrozole (0.5 or 2.5 mg per day) with megesterol acetate (160 mg per day), 551 patients were evaluated. Patients treated with 2.5 mg of letrozole attained a 24% overall response rate compared with 16 and 13% in patients treated with megesterol acetate and 0.5 mg of letrozole, respectively. Patients treated with 2.5 mg per day of letrozole also had a longer time to treatment failure than patients in the other two treatment groups, but there was not a statistically significant survival advantage for any treatment group. Similarly to the anastrazole trials, letrozole was well tolerated and was associated with less weight gain and dyspnea compared with megesterol acetate.

In a large trial, 555 postmenopausal women with tamoxifen-resistant, advanced breast cancer were randomized to letrozole (0.5 mg or 2.5 mg per day) or aminoglutethimide (250 mg twice daily). Patients treated with letrozole achieved a higher response rate than patients treated with aminoglutethimide: letrozole (2.5 mg per day) 19.5%, letrozole (0.5 mg per day) 16.7% and aminoglutethimide 12.4%. Patients treated with the higher dose of letrozole attained a small, but statistically significant improvement in time to disease progression and overall survival (28 versus 20 months) compared with patients treated with aminoglutethimide. The recommended daily dose of letrozole is 2.5 mg.

In the USA, only anastrazole and letrozole are FDA approved for second-line therapy of metastatic breast cancer in postmenopausal women following treatment with antioestrogens. Ongoing clinical trials are evaluating the efficacy of selective aromatase inhibitors as first-line therapy in postmenopausal patients with ER-positive, metastatic breast cancer. In addition, these agents are being evaluated in the adjuvant setting, either as primary therapy versus tamoxifen or following 5 years of treatment with tamoxifen.

Suicide Aromatase Inhibitors

Formestane

Formestane (Lentaron®) or 4-hydroxyandrostenedione (**Figure 7**) is a steroidal inhibitor that is approved outside the USA for the treatment of advanced breast cancer in postmenopausal women. Formestane has poor oral bioavailability and is generally administered as an intramuscular (IM) injection. A schedule of 250 mg, IM, every 2 weeks inhibits peripheral aromatization by 85% and suppresses oestradiol levels by 65%. Several reports indicate that this administration schedule results in objective tumour responses of up to 40% and stable disease status in an additional 14–29% of patients. Formestane has been compared with tamoxifen as first-line therapy of metastatic disease in postmenopausal patients who have received prior adjuvant therapy. Objective response rates were similar for the two agents. IM administration of the drug can be associated with injection abscesses and inadvertent intravenous injection has been associated with anaphylactoid reactions.

Exemestane

Exemestane (Aromasin®) (**Figure 7**) is a steroidal aromatase inhibitor developed for oral use. Following a 25 mg per day oral dose of exemestane, plasma levels of oestrone, oestradiol and oestrone sulphate were found to be suppressed by 94.5, 92.2 and 93.2%, respectively. A large phase II, multicentre trial evaluated exemestane (25 mg per day) in patients with progressive metastatic disease following treatment with nonsteroidal aromatase inhibitors. A total of 241 patients were enrolled and an objective response rate of 7% was observed and an additional 25% of patients maintained stable disease for >6 months. Nonvisceral sites of disease were most likely to respond. Exemestane produced an overall response rate of 40% (CR + PR + SD) in 78 patients who were refractory to, or who developed progressive disease while receiving aminoglutethimide. An objective response rate of 9% was

observed in 87 patients who developed progressive disease following treatment with megesterol acetate and tamoxifen. An additional 20% of patients maintained stable disease for >6 months. Patients who had become refractory to exemestane, 25 mg per day, did not respond to 100 mg per day. A large, international trial, comparing exemestane with megesterol acetate in tamoxifen-refractory metastatic disease has just been completed. In general, these clinical trials have shown exemestane to be well tolerated. Nausea, fatigue and hot flushes are the most commonly reported toxicities.

BLOCKING OESTROGEN ACTION

Tamoxifen for Treatment in Pre- or Postmenopausal Patients

The antitumour effect of tamoxifen (Nolvadex®) (**Figure 2**) is mediated primarily through the ER. Tamoxifen is considered the first-line endocrine therapy for all stages of breast cancer and was recently reported to reduce the risk of developing breast cancer in high-risk populations of women. Tamoxifen is currently approved for the treatment of metastatic breast cancer in postmenopausal women, ER-positive, metastatic breast cancer in premenopausal women and metastatic breast cancer in men. Tamoxifen is also approved as an adjuvant therapy, either alone or following chemotherapy for early-stage, hormone receptor-positive, breast cancer in pre- and postmenopausal women.

Clinical trials evaluating tamoxifen in postmenopausal women with advanced breast cancer have reported response rates of 15–53%, with a median duration of response of 20 months. The patient characteristics that predict a response to tamoxifen are similar to other endocrine therapies. Tumours expressing ER have a 50% probability of responding to tamoxifen compared with $<10\%$ in patients with ER-negative tumours. Tamoxifen has been shown to be equally effective in inducing tumour responses compared with other forms of endocrine therapy; however, tamoxifen is associated with less toxicity. The few published phase II clinical trials of tamoxifen treatment in premenopausal women with ER-positive, metastatic breast cancer report response rates between 20 and 45% with a median duration of response of 2.5–36 months. Small, randomized clinical trials comparing tamoxifen with oophorectomy in premenopausal patients with metastatic disease show no statistically significant difference in overall response rate between the two treatments.

The Oxford Overview Analysis of systemic adjuvant therapy for early-stage breast cancer confirmed that tamoxifen reduced the odds of breast cancer recurrence by 25% and the risk of death due to breast cancer by 16%, compared with patients not treated with tamoxifen (EBCTCG, 1998). The benefit produced by tamoxifen was more significant in postmenopausal women, in whom the annual odds of breast cancer recurrence and death were reduced by 29 and 20%, respectively. The relative benefit of tamoxifen was the same whether patients were axillary node positive or negative, but the absolute benefit was of greater magnitude in patients with the greatest risk of recurrence (e.g. axillary node positive). The magnitude of benefit derived from adjuvant tamoxifen therapy also depended on whether a tumour was ER positive or negative. The annual odds of recurrence were reduced by 13% in patients with ER-poor tumours and by 32% in patients with ER-positive tumours. Similarly, the reduction in the annual odds of death was 11 and 21% in patients with ER-poor and ER-positive tumours, respectively. In postmenopausal women with ER-positive tumours, tamoxifen reduced the annual odds of recurrence and death by 36 and 23%, respectively, compared with patients with ER-negative tumours, in whom the annual odds of recurrence and death were reduced by only 16%. Reanalysis of the data in the Oxford Overview Analysis 5 years later demonstrated little or no benefit for patients with ER-negative tumours taking tamoxifen (EBCTCG, 1998). In ER-positive patients under the age of 50 years treated with adjuvant tamoxifen, there is a significant reduction in the odds of death that is comparable to patients over the age of 50 years, treated with adjuvant tamoxifen, and similar to the reduction in the annual odds of death detected in patients under the age of 50 years who are treated with adjuvant chemotherapy.

It is important to point out that the Oxford Overview Analysis (EBCTCG, 1998) of breast cancer adjuvant clinical trials confirmed several principles defined in the laboratory: (1) tamoxifen blocks oestrogen binding to the ER so a patient with an ER-positive tumour is more likely to respond; (2) long-term early treatment in animal models produces a more complete antitumour effect than short-term treatment, thus long-term adjuvant therapy should be superior to short-term adjuvant therapy; and (3) tamoxifen prevents rat mammary carcinogenesis so the drug should reduce primary breast cancer incidence. It is now established that tamoxifen is effective in pre- and postmenopausal women, 5 years of adjuvant treatment is superior to 1–2 years of treatment and 5 years of adjuvant treatment reduces the incidence of contralateral breast cancer by 47%.

The optimum duration of tamoxifen therapy in the adjuvant setting has been addressed by several clinical trials and considered in the Oxford Overview Analysis. The Oxford Overview Analysis stratified patients based on tamoxifen therapy administered for <2, 2 and >2 years. Although the Oxford Overview Analysis does not answer the question regarding the optimum duration of tamoxifen treatment, the greatest reduction in the risk of recurrence and death was seen in patients treated with tamoxifen for >2 years. It is now clear that 5 years of therapy is superior

to shorter durations of therapy in terms of reducing the risk of recurrence and death (**Figure 8**). This is also true for the reductions in incidence of contralateral breast cancer (**Figure 9**).

The benefit of tamoxifen therapy for longer than 5 years has been addressed in three studies, comparing 5 years of tamoxifen therapy with 10 years or indefinite therapy. The results from these trials do not show additional clinical benefit from adjuvant tamoxifen treatment with durations of therapy longer than 5 years. The National Cancer Institute (NCI) recently recommended that tamoxifen should be discontinued after 5 years of adjuvant treatment in patients with node-negative disease (Fisher *et al*., 1996). An important observation, bearing on the duration of treatment, is the fact that patients who stop tamoxifen at 5 years continue to accrue survival benefit (EBCTCG, 1998). Obviously, the long-term effects of tamoxifen merit continued investigation. Indeed, there is a possibility that the duration (i.e. 5 years) of tamoxifen currently used for chemoprevention is not the optimum. Only additional clinical trials of longer duration can address this issue, but decades of patient follow-up will be necessary before sufficient data are available. The major problem is that long-term therapy tends to increase the possibility of developing drug resistance. In the case of tamoxifen, drug resistance can manifest itself in the selection of ER-positive tumours that grow in response to tamoxifen (Gottardis and Jordan, 1988). As a result, an antioestrogen that is devoid of oestrogen-like properties would be an advance for the treatment of breast cancer.

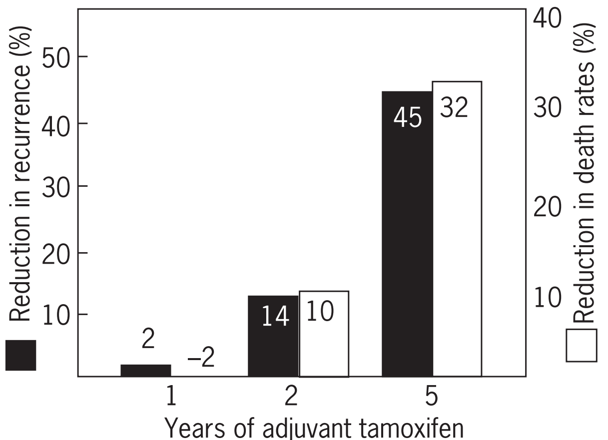


Figure 8 The relationship between the duration of adjuvant tamoxifen therapy in ER-positive premenopausal patients and the reduction in recurrence and death rate. A longer duration of treatment has a dramatic effective on patient survival. (From Jordan and Morrow, 1999, *Endocrine Reviews*, **20**, 253–278.)

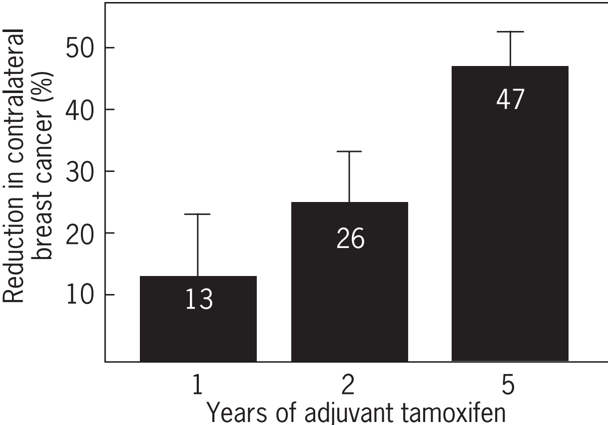


Figure 9 The relationship between the duration of adjuvant tamoxifen and the reduction in contralateral breast cancer. A longer duration is clearly superior, and the 5-year result that produces a 47% reduction in contralateral breast cancer is equivalent to the result observed in the NSABP tamoxifen prevention trial (see **Figure 11**). (From Jordan and Morrow, 1999, *Endocrine Reviews*, **20**, 253–278.)

Pure Antioestrogens for Treatment in Postmenopausal Patients

Pure antioestrogens, such as ICI 182,780 (Faslodex®) (**Figure 10**) are being developed to avoid the weak oestrogen-like effects of tamoxifen that are associated with drug resistance and changes in the uterus. Pure

OH
HO
17β-oestradiol

OH
HO
7
O
S
CF_2CF_3
ICI 182,780

Figure 10 Comparison of the structures of oestradiol and the pure antioestrogen ICI 182,780. The antioestrogen is administered by depot injection each month and is completing clinical trials for the treatment of breast cancer.

antioestrogens, unlike mixed agonist–antagonist compounds such as tamoxifen, have a different molecular mode of action. ICI 182,780 binds to the ER with affinity similar to oestrogen, but it causes destabilization of the ER dimer and enhanced ER degradation. The drug has proved effective against tamoxifen-resistant and tamoxifen-stimulated tumours *in vitro* and *in vivo*. Animal experiments also show that ICI 182,780 causes complete block of endometrial and myometrial proliferation. The potential advantages for complete oestrogen blockade in patients include more complete and/or longer lasting clinical responses, activity in tumours that have become resistant to tamoxifen therapy and lack of uterine complications. The potential disadvantages of complete oestrogen blockade include exacerbation of menopausal symptoms, reduced bone mineral density, vaginal dryness and lipid changes. Large randomized clinical trials are now under way that compare pure antioestrogens with selective aromatase inhibitors as second-line therapy for metastatic breast cancer. Clinical trials are also evaluating the effect of pure antioestrogens compared with antioestrogens as first-line therapy for metastatic breast cancer (Howell *et al*., 2000).

Tamoxifen for Prevention in Pre- or Postmenopausal Women

Tamoxifen has 30 years of experience in clinical trials and is used to treat all stages of breast cancer. Naturally, since tamoxifen prevents mammary carcinogenesis in the laboratory, it was a reasonable extension of the use of tamoxifen to test its worth in women at high risk. The National Surgical Adjuvant Breast and Bowel Project (NSABP) Tamoxifen Prevention Trial (P-1) recruited a total of 13 388 women of ≥ 35 years of age who were at increased risk of developing breast cancer to be randomized to receive tamoxifen or placebo for 5 years. All women with a history of lobular carcinoma *in situ* (LCIS) or who were over 60 years of age were eligible to participate in the trial. For all other women, a modified Gail algorithm was used to estimate their risk over time and those women with a 5-year risk of breast cancer of at least 1.67% were eligible to participate. The Gail model (Gail *et al*., 1989) incorporated the following variables: number of affected first degree relatives, number of previous breast biopsies, atypical hyperplasia in a biopsy, age at menarche, nulliparity or age at first live birth and race. The Gail model is available from the NCI as a Risk Disc.

With a median follow-up of 54.6 months, a 49% reduction in the risk of invasive breast cancer and a 50% reduction in the risk of noninvasive breast cancer were detected in women receiving tamoxifen. The reduction in risk was observed in women of all ages and at all levels of breast cancer risk (e.g. LCIS, atypical hyperplasia, etc.). Tamoxifen reduced the occurrence of ER-positive tumours by 69%, but there was no difference in the occurrence of ER-negative tumours. The number of vascular events (e.g. pulmonary embolism, deep venous thrombosis and stroke) was higher in the tamoxifen group than the placebo group. Furthermore, these events tended to occur in women over the age of 50 years. In the tamoxifen-treated women there was also a 45% reduction in hip fractures, a 39% reduction in Colles wrist fractures and a 26% reduction in spine fractures compared with the placebo group. This last finding with an 'antioestrogen' was not unexpected as tamoxifen maintains bone density in the ovariectomized rat and postmenopausal women. Two other chemoprevention studies have been published, but they are too small to be classified as chemoprevention trials. However, the results are valuable to assess the toxicity in well women. The results of the studies are compared in **Figure 11**.

Side Effects of Tamoxifen in Pre- or Postmenopausal Women

The widespread use of tamoxifen as a treatment and now as a chemopreventive has focused attention on its side effects. The most common adverse effect reported in relation to tamoxifen therapy is vasomotor instability, manifested by hot flushes, facial flushing, tachycardia and sweating. These symptoms are reported in 15–20% of pre- and postmenopausal patients and typically decrease in intensity and frequency over time. Amenorrhoea can occur in up to 40% of premenopausal women. Visual disturbances caused by cataract formation, optic neuritis, retinopathy or macular oedema rarely occur. Thromboembolic events can occur in 1–2% of patients treated with tamoxifen (Fisher *et al*., 1998).

During the past 10 years, a number of case reports and clinical trial results have associated tamoxifen therapy with an increased incidence of endometrial carcinoma. Several analyses of the data have concluded that there is a modest increase in the incidence of endometrial carcinoma as a result of receiving tamoxifen (4/1000 tamoxifen-treated women per year versus 1/1000 women per year for the normal postmenopausal population). Premenopausal women do not have an elevation of their risk of endometrial cancer (Fisher *et al*., 1998). This potential toxicity is more relevant to patients receiving tamoxifen as a chemopreventive agent.

SELECTIVE OESTROGEN RECEPTOR MODULATION

Nonsteroidal antioestrogens were originally defined as compounds that would inhibit oestradiol-stimulated rat uterine weight. The nonsteroidal antioestrogens are partial oestrogen agonists in the uterus, but also inhibit carcinogen-induced rat mammary tumour growth and the growth

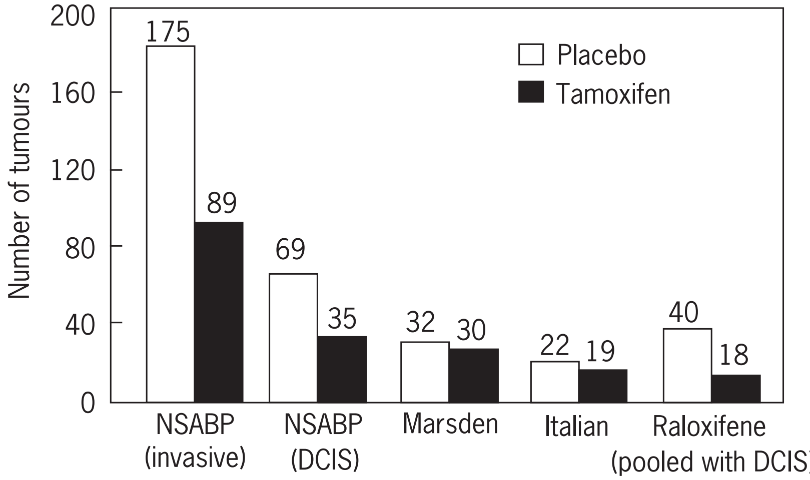


Figure 11 A comparison of the able-to-be evaluated events observed in the studies to reduce the incidence of breast cancer in well women. The National Surgical Adjuvant Breast and Bowel Project (NSABP) P-1 trial is the only completed prospective clinical trial designed to test the worth of an antioestrogen to prevent breast cancer in 13 388 high-risk women, who were selected on a validated mathematical risk model. The figure illustrates the effect of tamoxifen on both invasive and noninvasive ductal carcinoma *in situ* (DCIS) breast cancer. By contrast, the Royal Marsden study is a pilot project, and stated to be a toxicity evaluation in 2471 women at lower risk than the NSABP study. These women were selected based on first-degree relatives with breast cancer. The Italian study is from 3984 normal-risk women, half of whom had already had an oophorectomy. Finally, the raloxifene result is estimated from published abstracts and is a secondary endpoint from 10 553 postmenopausal women around 70 years of age who were participating in osteoporosis trials. The results for raloxifene are for both invasive and noninvasive breast cancer, but an effect of raloxifene on reducing DCIS alone has not been found because the events are too few. (From Jordan and Morrow, 1999, *Endocrine Reviews*, **20**, 253–278.)

of ER-positive breast cancer cells *in vitro*. The finding that the compounds expressed increased oestrogenic properties, i.e. increased uterine weight in the mouse, raised questions about the reasons for the species specificity and created a new dimension for study which ultimately led to the recognition of the target site-specific actions of antioestrogens. However, the finding that tamoxifen would increase immune-deficient mouse uterine weight but prevent the growth of a transplanted ER-positive breast tumour suggested that the tamoxifen–ER complexes were being interpreted differently, i.e. as an oestrogenic or antioestrogenic signal, respectively, at different target sites (Jordan and Morrow, 1999). Additionally, tamoxifen and raloxifene maintained bone density in ovariectomized rats at doses that prevent mammary cancer (Delmas *et al*., 1997). This concept was subsequently referred to as selective oestrogen receptor modulation (SERM) to describe the target site-specific effects of tamoxifen around a women's body (**Figure 12**) and to justify the use of raloxifene as a preventive for osteoporosis. The whole class of drugs is now known as SERMs.

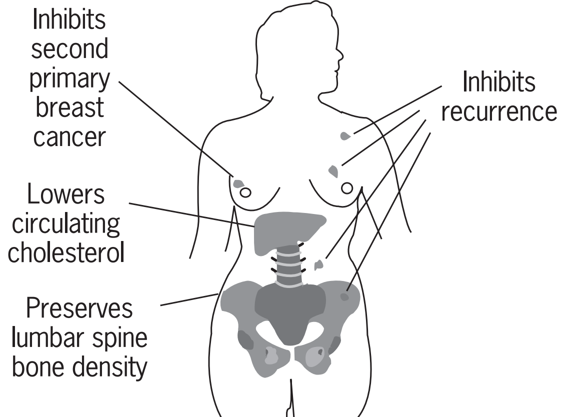


Figure 12 The selective ER modulation observed clinically with tamoxifen in postmenopausal women. Tamoxifen is an antioestrogen in breast cancer but has oestrogen-like actions in bone and causes a decrease in circulating cholesterol.

During the past decade, there has been intense interest in identifying the reason for the target site-specific effects of antioestrogens. One possible mechanism is the presence of two ERs, ERα and ERβ, but, there are no definitive studies that indicate that tamoxifen or raloxifene interact differently with either receptor. Although the precise mechanisms have not, as yet, been discovered, the incentive to achieve success is high. This knowledge will not only permit a rational application of tamoxifen and raloxifene in patients, but also the discovery of new

mechanisms for drug selectivity will open the door for new innovations in drug discovery.

An ideal antioestrogen would exhibit the target site-specific effects that SERMs do but they would be more refined. For example, the use of tamoxifen carries an increased risk of endometrial cancer and has been shown to cause DNA adduct formation in the rat liver (White, 1999). Clearly, a compound that does not cause the negative side effects and that possesses all of the benefits of tamoxifen would be an attractive option. Several different compounds are currently under investigation. In general they are structurally related to the compound raloxifene (**Figure 13**).

The rationale for the use of raloxifene as a breast cancer preventive is based on a hypothesis that was formulated when selective oestrogen receptor modulation was first recognized. The concept was simply stated.

We have obtained valuable clinical information about this group of drugs that can be applied in other disease states. Research does not travel in straight lines and observations in one field of science often become major discoveries in another. Important clues have been garnered about the effects of tamoxifen on bone and lipids (**Figure 12**) so it is possible that derivatives could find targeted applications to retard osteoporosis or atherosclerosis. The ubiquitous application of novel compounds to prevent diseases associated with the progressive changes after menopause may, as a side effect, significantly retard the development of breast cancer. The target population would be postmenopausal women in general, thereby avoiding the requirement to select a high-risk group to prevent breast cancer (Lerner and Jordan, 1990).

This concept, and the supporting laboratory data, subsequently led to the development of raloxifene for osteoporosis, but with the possibility of breast and uterine safety. This would be distinct from the use of oestrogen replacement therapy to prevent osteoporosis because oestrogen causes an increased risk of breast and endometrial cancer.

Figure 13 SERMs currently under investigation as treatments or preventives for osteoporosis, coronary heart disease or breast cancer.

Raloxifene

Raloxifene (Evista®) (**Figure 13**) was originally under development as a breast cancer therapy, but results from an early clinical trial with raloxifene in patients with advanced disease were disappointing. As a result, the development of raloxifene was redirected in the 1990s towards osteoporosis because of encouraging laboratory data on bone density (Jordan *et al*., 1987). Raloxifene is approved for the prevention of osteoporosis in postmenopausal women (Delmas *et al*., 1997; Ettinger *et al*., 1999). The current incentive to test raloxifene further in multiple diseases comes from trials to prevent osteoporosis The Multiple Outcomes of Raloxifene Evaluation (MORE) Trial involved over 7700 women and an additional 2800 women participating in other clinical trials designed to examine the risk of fracture in postmenopausal women who received either placebo or raloxifene were also analysed. The trials were not designed to assess the efficacy of raloxifene as a breast cancer-preventive agent, but based on the hypothesis that a tamoxifen-like drug can reduce breast cancer and maintain bone density, analysis of the data at 40 months of follow-up showed a 65% decrease in the incidence of invasive breast cancer in raloxifene-treated patients compared with patients receiving placebo. Similar to the findings from the NSABP Tamoxifen Prevention Trial (P-1) (Fisher *et al*., 1998), the effect of raloxifene was most striking in reducing the occurrence of ER-positive tumours (74%). Raloxifene appeared to have no effect on the incidence of ER-negative tumours (Cummings *et al*., 1999). An analysis of breast cancer incidence from all randomized osteoporosis trials involving raloxifene is shown in **Figure 11**. This preliminary observation encouraged the NCI to support a second prevention trial that is comparing raloxifene with tamoxifen in postmenopausal women determined to be at high risk of developing breast cancer. The trial, Study of Tamoxifen and Raloxifene (STAR), is recruiting 22 000 volunteers.

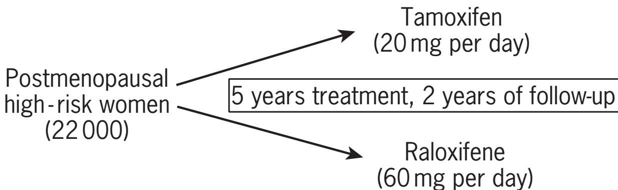


Figure 14 The study of tamoxifen and raloxifene (STAR) is recruiting 22 000 high-risk postmenopausal women to be randomized to tamoxifen (20 mg per day) or raloxifene (60 mg per day). Compounds will be given for 5 years with a 2-year wash out period. Endometrial biopsies are not required, but symptoms of spotting and bleeding will be followed up to determine whether endometrial cancer is present. The primary endpoint is breast cancer incidence, but secondary endpoints of fractures, coronary heart disease and symptoms will be collected.

Results will be available in 2006 (**Figure 14**) (Jordan and Morrow, 1999).

CONCLUSIONS

The development of endocrine therapy for the treatment and prevention of breast cancer has evolved dramatically over the last 100 years. Most patients with an established diagnosis of breast cancer will receive some form of endocrine therapy during the clinical course of their disease. One of the most attractive aspects of the endocrine therapies currently available is the ease with which most patients tolerate the medications.

As laboratory observations have been translated to the clinic, antioestrogens have been shown to reduce the incidence of breast cancer in high-risk populations. One of the intriguing observations from the NSABP Tamoxifen Prevention Trial is that tamoxifen appears to reduce the occurrence of ER-positive breast cancer exclusively (Fisher *et al*., 1998). However, the fact raloxifene is widely used as a preventive for osteoporosis and is being tested in high-risk populations for the prevention of coronary heart disease points to a broader future application for SERMs. As novel SERMs are developed and targeted for different diseases, it is possible to speculate that the incidence of breast cancer will decrease. Nevertheless, with an ageing population, the absolute number of breast cancer cases may remain the same. These tumours may be refractory to traditional endocrine therapy because of antioestrogen drug resistance or the tumour is already ER negative. This will be among the challenges of cancer pharmacology in the twenty-first century.

At present there are no therapeutic interventions that can reduce the risk of prostate cancer. However, measurement of prostate-specific antigen (PSA) can provide guidance for the physician when monitoring elderly men over the age of 70 years.

Prostate cancer is age related and appears to be more hormone responsive than breast cancer, i.e. more men respond initially to endocrine therapy. The treatment of prostate cancer has followed the successful path established for breast cancer. Men no longer have to have their testes removed to reduce circulating androgen levels. LHRH agonists are used routinely to cause a chemical hypophysectomy and reduce gonadotropin levels. Based on advances with nonsteroidal antioestrogens, i.e. drugs for breast cancer treatment, nonsteroidal antiandrogens have been developed to block the androgen receptor in the tumour. Today the treatment of prostate cancer mirrors the treatment of breast cancer. For the future, an effort is being made to study the chemoprevention of prostate cancer by interrupting the prostate-specific metabolic activation of androgens to the steroid dihydrotestosterone that binds to the androgen receptor. Similarly, there are attempts to develop a new class of drugs, the selective androgen receptor modulators (SARMs), that might have multiple applications in men's health.

REFERENCES

Brzozowski, A. M., *et al*. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, 389, 753–758.

Couse, J. F. and Korach, K. S. (1999). Estrogen receptor null mice: what have we learned and where will they lead us? *Endocrine Reviews*, **20**, 358–417; Erratum, **20**, 459.

Cummings, S. R., *et al*. (1999). The effect of raloxifene on risk of breast cancer in postmenopausal women: results from the MORE randomized trial. Multiple Outcomes of Raloxifene Evaluation. *Journal of the American Medical Association*, **281**, 2189–2197.

Delmas, P. D., *et al*. (1997) Effects of raloxifene on bone mineral density, serum cholesterol concentrations, and uterine endometrium in postmenopausal women. *New England Journal of Medicine*, **337**, 1641–7.

Ettinger, B., *et al*. (1999). Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. *Journal of American Medical Association*, **282**, 637–645.

Fisher, B., *et al*. (1996). Five versus more than five years of tamoxifen therapy for breast cancer patients with negative lymph nodes and estrogen receptor-positive tumors. *Journal of the National Cancer Institute*, **88**, 1529–1542.

Fisher, B., *et al*. (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute*, **90**, 1371–1388.

Fisher, B., *et al*. (1999). Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet*, **353**, 1993–2000.

EBCTCG (1998). Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet*, **351**, 1451–1467.

Gail, M. H., *et al*. (1989). Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute*, **81**, 1879–1886.

Gottardis, M. M. and Jordan, V. C. (1988). Development of tamoxifen-stimulated growth of MCF-7 tumors in athymic mice after long-term antiestrogen administration. *Cancer Research*, **48**, 5183–5187.

Howell, A., *et al*. (2000). ICI 182,780 (Faslodex): development of a novel, ‘pure’ antiestrogen. *Cancer*, **89**, 817–825.

Jensen, E. V. and Jacobson, H. I. (1962). Basic guides to the mechanism of estrogen action. *Recent Progress in Hormone Research*, **18**, 387–414.

Jordan, V. C., *et al*. (1987). Effects of anti-estrogens on bone in castrated and intact female rats. *Breast Cancer Research and Treatment*, **10**, 31–35.

Jordan, V. C. and Morrow, M. (1999). Tamoxifen, raloxifene, and the prevention of breast cancer. *Endocrine Reviews*, **20**, 253–278.

Lerner, L. J. and Jordan, V. C. (1990) Development of anti-estrogens and their use in breast cancer: Eighth Cain Memorial Award Lecture. *Cancer Res*earch, **50**, 4177–4189.

MacGregor, J. I. and Jordan, V. C. (1998). Basic guide to the mechanisms of antiestrogen action. *Pharmacological Reviews*, **50**, 151–196.

MacGregor, *et al.* (2000). Allosteric silencing of activating function 1 in the 4-hydroxytamoxifen estrogen receptor complex is induced by substituting glycine for aspartate at amino acid 351. *Cancer Research*, **60**, 5097–5105.

Shiau, A. K., *et al*. (1998). The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell*, **95**, 927–937.

White, I. N. (1999). The tamoxifen dilemma. *Carcinogenesis*, **20**, 1153–1160.

FURTHER READING

Burger, H. G. (2000). Selective oestrogen receptor modulators. *Hormone Research*, **53**, Suppl. 3, 25–29.

Denis, L. J. and Griffiths, K. (2000). Endocrine treatment in prostate cancer. *Seminars in Surgical Oncology*, **18**, 52–74.

Henry, R. Y. and O’Mahony, D. (1999). Treatment of prostate cancer. *Journal of Clinical Pharmacy and Therapeutics*, **24**, 93–102.

Hoffken, K. and Kath, R. (2000). The role of LH–RH analogues in the adjuvant and palliative treatment of breast cancer. *Recent Results in Cancer Research*, **153**, 61–70.

Kuyu, H., *et al*. (1999). Recent advances in the treatment of prostate cancer. *Annals of Oncology*, **10**, 891–898.

Labrie, F. (2000). Screening and early hormonal treatment of prostate cancer are accumulating strong evidence and support. *Prostate*, **43**, 215–222.

Lippman, S. M. and Brown, P. H. (2000). Tamoxifen prevention of breast cancer: an instance of the fingerpost. *Journal of the National Cancer Institute*, **91**, 1809–1819.

Njar, V. C. and Brodie, A. M. (1999). Comprehensive pharmacology and clinical efficacy of aromatase inhibitors. *Drugs*, **58**, 233–255.

Osborne, C. K., *et al*. (2000). Selective estrogen receptor modulators: structure, function, and clinical use. *Journal of Clinical Oncology*, **18**, 3172–3186.

Chapter 93

Antiangiogenic Therapy

Kevin J. Turner and Adrian L. Harris

Imperial Cancer Research Fund, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

CONTENTS

- Why is Angiogenesis an Attractive Therapeutic Target?
- Principles of Antiangiogenic Therapy
- Potential Problems in Antiangiogenic Therapy
- The Future of Antiangiogenic Therapy

WHY IS ANGIOGENESIS AN ATTRACTIVE THERAPEUTIC TARGET?

Tumours depend upon angiogenesis for growth and the development of metastases. This observation has prompted efforts to exploit this dependence in the development of antitumour strategies. This approach has become known as 'antiangiogenesis.' Antiangiogenesis has several theoretical advantages over conventional therapy. Firstly, endothelial cells (ECs) are genetically stable. Unlike tumour cells, the ECs of the neovasculature are normal, unmutated host cells. Spontaneous mutation in tumour cells underlies drug resistance. The relative genetic stability of the endothelium means that the development of drug resistance is less likely. Second, drugs have easy access to their therapeutic target. One of the problems with anticancer therapy is that access of intravenous agents to solid tumours is poor. In antiangiogenic therapy the target cells are those which constitute vessel walls; intravenous antiangiogenic agents contact these cells directly. Third, antiangiogenic therapy has an inherent 'amplification' mechanism. In conventional therapy, cells unaffected by the therapeutic intervention may continue to proliferate. This proliferation of a subpopulation of the tumour leads to tumour regrowth and enhanced drug resistance. In antiangiogenic therapy, successful interference with only a few ECs may lead to disruption of a whole vessel. Loss of a single vessel may deprive a significant number of tumour cells of essential nutrients, leading to cell death. Fourth, ECs active in angiogenesis are distinguishable from host ECs unassociated with the tumour vasculature. Were this not the case, antiangiogenic therapy might affect normal endothelium throughout the body which would markedly restrict its utility. Tumour-associated ECs are distinguished by their rate of proliferation; in human breast cancer, tumour-associated endothelial cells may proliferate up to 50 times more than normal vessels (Fox *et al.*, 1993). Angiogenic ECs also differ from normal endothelium in the antigens that they express. For example, the vascular endothelial growth factor (VEGF) receptor is more abundant on tumour endothelium compared with normal and VEGF also stimulates production and expression of the urokinase receptor on tumour ECs. Similarly, the cell adhesion molecule E-selectin and the integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$ are known to be unregulated in tumour endothelium (Brooks, 1996). Exploitation of these and other differences may lead to antiangiogenic therapy that has no effect on ECs other than those in the tumour neovasculature. Fifth, whilst tumour-associated endothelium differs from normal host endothelium, the ECs in different tumours are relatively similar. It may be that a single antiangiogenic strategy is effective against disparate tumour types. Conversely, there is some heterogeneity between the ECs in different tumours. Exploitation of these differences may allow organ-specific antiangiogenic approaches. (See also the chapter *Angiogenesis*.)

PRINCIPLES OF ANTIANGIOGENIC THERAPY

Antiangiogenic therapy can be targeted at different points in the angiogenesis cascade (**Figure 1**). A study of angiogenic factors in primary tumours has shown that most tumours express at least six different factors (Relf *et al.*, 1997). Despite this multiplicity of factors, inhibitors of individual factors are effective, possibly because of synergy between factors or because some factors are more potent or important than others. Approaches that target ECs directly (vascular targeting) also constitute an important strategy and may bypass problems associated with targeting one of several mediators. The current state of trials involving antiangiogenic agents is summarized in **Tables 1–3**.

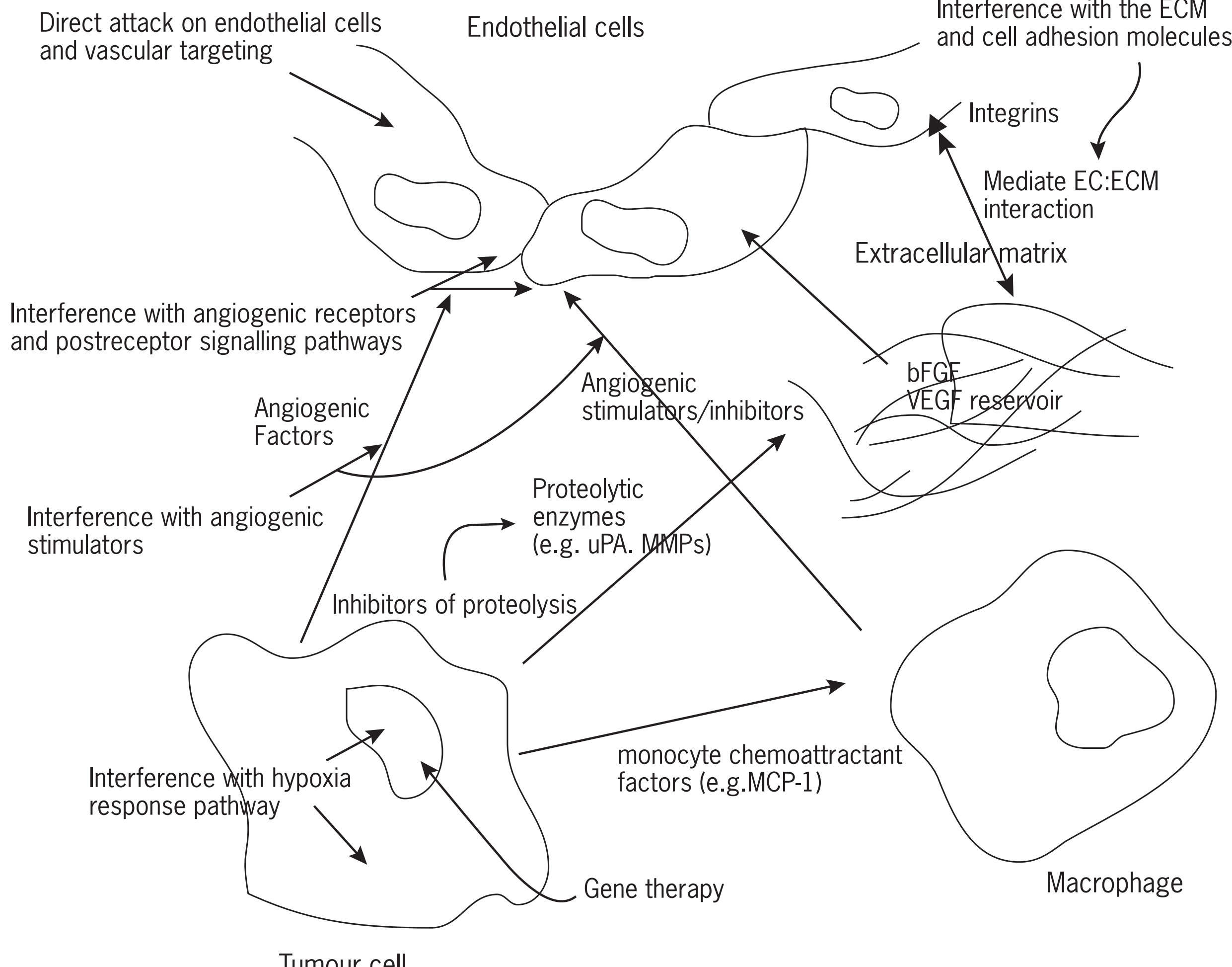


Figure 1 Diagram showing the points at which different classes of antiangiogenic agents may act. (Adapted from Jones and Fujiyama, 1999.)

Interference with Angiogenic Stimulators

VEGF

Monoclonal antibodies to VEGF have reduced both angiogenesis and tumour growth in a number of xenograft animal models. Humanized anti-VEGF monoclonal antibody for the treatment of solid tumours has been used with some success in phase I clinical trials with disease stabilization achieved in 13/25 patients. Phase II trials are currently under way for metastatic renal cancer.

Our understanding of the effects of VEGF antibodies has been enhanced by the use of intravital videomicroscopy techniques that provide 'real-time' information. Treatment with VEGF antibodies results in reductions in the diameter and tortuosity of vessels, and eventually to vessel regression. These data suggest that VEGF is not only needed for angiogenesis, but is also required for the survival of tumour-associated endothelium (Yuan *et al.*, 1996).

An exciting approach is the coupling of toxins to VEGF itself. A fusion molecule consisting of VEGF and the active region of the diphtheria toxin (DT390) demonstrates highly selective toxicity on ECs. This construct slows tumour growth in a murine model of Kaposi sarcoma (KS) (Arora *et al.*, 1999).

FGFs

Interferons are frequently used in the treatment of patients with advanced renal cancer. Exposure of carcinoma cells to interferon-α (IFN-α) *in vitro* and *in vivo* can down-regulate the gene expression and protein production of basic fibroblast growth factor (bFGF) and inhibit growth of human bladder cancer xenografts in nude mice (Dinney *et al.*, 1998). IFN-2α has also been used successfully in the treatment of life-threatening airway haemangiomas in infancy and in the treatment of a giant cell tumour of the mandible. Interferon-α is currently in phase II/III trials.

Table 1 Angiogenesis inhibitors in phase I clinical trials or currently under development (see text for full details)

Name	Class/mechanism of action	Status
VEGF conjugated to diphtheria toxin	Selective targeting of ECs	Under development
Platelet factor 4	Binds to VEGF/bFGF and inhibits binding to receptors	Under development
Dominant mutant VEGFR-2	Binds VEGF and prevents postreceptor triggering	Under development
Soluble mutant VEGFR-1	Binds VEGF and prevents binding to receptor	Under development
SU6668	Inhibits signalling via VEGFR-2 and PDGFR	Phase I
ZD4190	Blocks signalling via VEGFR-2	Under development
Accutin	Disintegrin, blocks binding of ECs to ECM by blocking integrin $\alpha_v\beta_3$	Under development
Endostatin	Mechanism uncertain, but may induce apoptosis	Phase I
Restin	Fragment of collagen XV, mechanism uncertain	Under development
Vasostatin	Fragment of calreticulin, mechanism uncertain	Under development
TSP-1	ECM glycoprotein, mechanism uncertain	Under development
Ang-2	Ang-1 antagonist	Under development
Tie-2 extracellular domain	Reduces binding to endogenous Tie-2	Under development
Gene therapy	Variety of mechanisms	Under development
CM 101	Group B haemolytic streptococcus exotoxin	Phase I

Table 2 Antiangiogenic agents in phase II clinical trials (see text for full details)

Name	Class/mechanism of action	Status
Humanized monoclonal antibody against VEGF	Inhibits VEGF	Phase II, renal cancer
Thalidomide	Inhibits bFGF, precise mechanism unknown	Phase II, against AIDS-KS, breast and prostate cancers
Suramin	Binds VEGF/bFGF	Phase I/II, prostate cancer
Vitaxin	Humanized mouse anti-integrin $\alpha_v\beta_3$ antibody	Phase I/II
Neovastat	MMP inhibitor	Phase I/II
Squalamine	Blocks EC activation and migration	Phase II (in conjunction with other agents)
Combrestatin A4	Induces apoptosis of proliferating endothelial cells	Phase I/II
IL-12	Upregulates IFN-γ and stimulates NK cells	Phase II (AIDS-KS)

Table 3 Antiangiogenic agents in phase III clinical trials (see text for full details)

Name	Class/mechanism of action	Status
Interferon-α	Down-regulation of bFGF	Phase II/III
CAI	Inhibition of calcium-mediated signalling	Phase II/III, ovarian, lung and renal cancers
SU5416	Blocks tyrosine kinase activity of VEGFR-1	Phase I/II/III, in conjunction with conventional chemotherapeutic agents
Marimastat	MMP inhibitor	Phase III
Bay 12-9566	MMP inhibitor	Phase III
AGM 1470	Inhibits methionine aminopeptidase type 2	Phase III
IM-862	Synthetic dipeptide that upregulates IL-12	Phase III (KS)

Thalidomide was originally marketed as a sedative hypnotic in the 1950s. It is now infamous for its potent teratogenic activity causing hypoplastic and aplastic defects in the limbs when ingested during the first 2 months of pregnancy. Thalidomide potently inhibits bFGF-induced angiogenesis *in vivo*. It seems likely that the developmental effects of thalidomide are mediated via this antiangiogenic activity. Thalidomide is currently in stage II trials for the treatment of breast and prostate cancer, gliomas, AIDS related Kaposi sarcoma (AIDS-KS) and diabetic retinopathy. Partial responses have been achieved in hormone-refractory prostate cancer, gliomas and multiple myeloma (Singhal *et al.*, 1999).

Suramin is a polysulfonated naphthylurea that mimics heparin and has been used for 80 years to treat parasitic infections. Virtually all endothelial growth factors

including bFGF and VEGF are heparin binding. Angiogenesis induced by bFGF and VEGF is inhibited in animal models. Most of the clinical experience with suramin is in the treatment of hormone-refractory prostatic carcinoma where it has been evaluated in phase I and II clinical trials with a response rate of approximately 20%. Suramin use is complicated by its long but variable half-life (44–54 days) and its narrow therapeutic index. The most common side effects are haematological and neurological toxicity and adrenal suppression. Analogues with a more favourable therapeutic ratio are under investigation. Some of the problems with use of suramin may be avoided if it is administered locally; some success has been achieved with bladder instillation in the treatment of superficial transitional cell carcinoma.

The anticoagulants fragmin and heparin both have antiangiogenic activity. Analogues to specific regions of the heparin molecule have been shown to inhibit bFGF-induced proliferation in some tumour cell lines. Pentosan polysulfate (PPS) is a synthetic heparin-like compound that is in phase I/II trials for the treatment of AIDS-KS. Disease stabilization has been observed in a small subgroup.

Platelet factor 4 (PF-4) is stored in the α granules of platelets. It is released during platelet aggregation but its physiological role remains unclear. Its use as an experimental control led to the serendipitous observation that it inhibits angiogenesis in animal models. In this context, PF-4 probably works by binding heparin binding proteins such as bFGF and VEGF thus reducing their binding to their receptors. Initial clinical results with PF-4 are disappointing with no beneficial effect observed in the treatment of a variety of tumour types. The short circulating half-life of the currently available forms of PF-4 may underlie this lack of response. Direct injection into KS lesions has resulted in some encouraging results.

Interference with Angiogenic Receptors and Post-receptor Signalling Pathways

The importance of the VEGF–VEGFR-2 receptor–ligand system was investigated in experiments involving the infection of glioblastoma multiforme cells with a retrovirus encoding a dominant mutant of VEGFR-2 which lacks the kinase domain. The mutant receptor dimerized with the wild-type receptor and prevented triggering of the postreceptor pathway after VEGF binding. Glioblastoma multiforme cells infected in this way demonstrated reduced growth *in vitro* (Millauer *et al.*, 1994).

By alternative splicing, the mRNA for VEGFR-1 can also encode a truncated soluble form of the receptor that consists of the extracellular VEGF binding domain but not the intracellular signal transduction moiety. VEGF binds to both species of VEGFR-1 with equal affinity. The role of the soluble receptor variant and the mechanisms that control its expression are unclear. In animal experiments, expression of the soluble form of the receptor was associated with inhibition of tumour growth and metastases and resulted in reduced mortality (Goldman *et al.*, 1998).

Carboxyaminotriazole (CAI) is a low molecular weight synthetic peptide that inhibits proliferation and invasion of a variety of tumour cell types. It works by blocking calcium channels and thus inhibiting the tyrosine kinases associated with a number of angiogenic factor receptors. In *in vitro* angiogenesis assays and in tumour xenograft animal models, CAI has been shown to be antiangiogenic. CAI has been evaluated in phase I and II clinical trials. Encouragingly, CAI resulted in disease stabilization in a number of refractory tumours. Gastrointestinal side effects were the most common but cerebellar ataxia was dose limiting. Phase II/III trials are under way against ovarian, lung and renal cancers.

A number of low molecular weight synthetic peptides have recently been evaluated for their ability to inhibit angiogenesis: Lefluonomide (SU101), SU6668 and SU5416. Lefluonomide is an inhibitor of receptor tyrosine kinases that blocks platelet-derived growth factor (PDGF) signalling. Trials are underway in prostatic, ovarian, lung and brain tumours. SU5416 blocks the tyrosine kinase activity of the VEGFR-1 receptor. Disease stabilization was observed in phase I trials for a number of tumour types and phase I, II and III trials using SU5416 in combination with conventional chemotherapeutic modalities are under way. SU6668 inhibits signalling mediated by both VEGFR-1 and the PDGF receptor. Phase I trials are under way.

ZD4190 is an anilinoquinazoline derivative that inhibits specifically the tyrosine kinase activity of the VEGFR-2 receptor. Significant antitumour activity has been demonstrated following oral administration in animal models.

Signalling by FGFs is mediated by transmembrane receptor tyrosine kinases. Encouraging results in the specific inhibition of this kinase activity have been achieved with pyrimidine-based compounds.

Interference with the Extracellular Matrix (ECM) and Cell Adhesion Molecules

The integrins are a family of transmembrane receptors that mediate cell–cell and cell–ECM interaction. The integrin $\alpha_v\beta_3$ is expressed highly by angiogenic ECs and antibodies to $\alpha_v\beta_3$ can block bFGF-induced angiogenesis. The disintegrins, a family of low molecular weight peptides derived from snake venom, block the binding of human umbilical vein ECs (HUVEC) to the ECM by blockade of $\alpha_v\beta_3$. Accutin is a recently purified disintegrin that has been shown to have *in vivo* antiangiogenic activity in the chick chorioallantoic membrane model of angiogenesis. This effect was mediated by blockade of endothelial $\alpha_v\beta_3$

and induction of apoptosis. A humanized version of a mouse anti-$\alpha_v\beta_3$ antibody named Vitaxin is currently under evaluation in phase I and II clinical trials.

The chemotherapeutic agent doxorubicin has been linked to an α_v integrin-binding motif in an attempt to target drug delivery. The construct enhanced the efficacy and reduced the toxicity of doxorubicin given to nude mice with breast cancer xenografts. (See the chapter *Targeting the Extracellular Matrix*.)

Endogenous Inhibitors of Angiogenesis

The switch to the angiogenic phenotype that is necessary for tumour growth probably results from not only induction of angiogenic stimulators but also the down-regulation of angiogenic inhibitors. Angiostatin and endostatin are angiogenesis inhibitors that show great potential from animal experiments. Angiostatin was initially purified from the urine and blood of mice bearing Lewis lung carcinoma primary xenografts (O'Reilly *et al.*, 1994). Upon removal of the primary tumour (the source of angiostatin), there was marked growth and vascularization of metastases. If purified angiostatin was given intraperitoneally coincident with the removal of the primary tumour, inhibition of metastases was maintained. Furthermore, exogenous angiostatin also reduced the growth of the primary tumour xenografts. Similar results were obtained when the same group subsequently identified endostatin, an endogenous inhibitor of angiogenesis from haemangioendotheliomas. In a dramatic series of experiments, Boehm *et al.* then demonstrated that repeated courses of endostatin caused tumour regression without the emergence of drug resistance (Boehm *et al.*, 1997). Endostatin is currently under evaluation in phase I trials.

The precise mechanisms of action of angiostatin and endostatin are unclear. Angiostatin appears to bind to the α/β-subunits of ATP synthase and thus inhibits matrix-enhanced plasminogen activation with a consequent reduction in invasive activity (Stack *et al.*, 1999). Endostatin is believed to induce endothelial cell apoptosis by the down-regulation of antiapoptotic proteins such as Bcl-2 (Dhanabal *et al.*, 1999).

Angiostatin and endostatin are derived from the proteolytic processing of plasminogen and collagen XVIII, respectively, which are not themselves antiangiogenic. Similarly, Restin is a 22-kDa fragment of human collagen XV that suppresses the growth of renal carcinoma xenografts and Vasostatin is a 180 amino acid fragment of calreticulin that also suppresses human tumour growth in a murine model. Other examples of inhibitory molecules encrypted within larger inactive molecules include a 16-kDa fragment of prolactin which inactivates urokinase by inducing the urokinase inhibitor PAI-1 (plasminogen activator inhibitor), a fragment of prothrombin (prothrombin kringle 2 domain) which inhibited bFGF-induced growth of ECs and the cleaved conformation of the serpin antithrombin. Thrombospondin-1 (TSP-1), another potent inhibitory factor, is an ECM glycoprotein that is also derived from an internal fragment of the whole molecule. Success in determining precisely which residues within TSP-1 are responsible for its inhibitory activity has enhanced the potential of TSP-1 as an antiangiogenic agent. Regulation of the proteolytic processing that produces these endogenous inhibitors is a potential antiangiogenic strategy.

Blood vessel maturation is regulated by angiopoetin-1 (Ang-1), which binds to the tyrosine kinase receptors Tie-1 and Tie-2. Tumours produce a related factor, Ang-2, which binds to Tie-2 but does not activate it. The use of Ang-2 as an antiangiogenic Ang-1 antagonist is under investigation. Similarly, when mice are injected with the extracellular domain of the Tie-2 receptor, tumour growth can be reduced by more than 75% (Lin *et al.*, 1997).

Oncogene Stimulation of Angiogenesis and the Role of Hypoxia

Activation of oncogenes is associated with gain of function and inactivation of tumour-suppressor genes with loss of function. Changes in gene expression of this type may underlie the switch to an angiogenic phenotype and thus provide a further strategy for antiangiogenic intervention. A number of recent observations link oncogenes/tumour-suppressor genes with mediators of angiogenesis.

The tumour-suppressor gene *p53* positively regulates TSP-1 and thus mutations of the *p53* gene result in reduced TSP-1 expression (Dameron *et al.*, 1994). *p53* mutations are associated with many other aspects of tumorigenesis including increased cell proliferation and reduced apoptosis. Similarly, the proangiogenic molecules bFGF, TFG-α and TGF-β are upregulated in cells where there is a mutation in the *ras* oncogene.

The transforming nature of oncogenes might be manifested, at least in part, through angiogenesis. In an elegant series of experiments, Rak *et al.* measured the level of VEGF expression in cultured rat intestinal epithelial cells that had been transfected with a mutant human H-*ras* oncogene (Rak *et al.*, 1995). They demonstrated that transfected cells strongly expressed VEGF mRNA whereas VEGF was undetectable in nontransformed cells. Similar associations have been demonstrated for oncogenes that encode transcription factors such as c-Fos and translation initiation factors such as eukaryotic initiation factor-4e (EIF-4e). EIF-4e may enhance the efficiency of use of VEGF mRNA allowing an increase in protein without increased RNA synthesis (Scott *et al.*, 1998; Crew *et al.*, 2000). Thus, whilst oncogenes might predominantly promote tumour growth via alterations in control of the cell cycle, a secondary and indirect mechanism might be mediated via angiogenesis.

Physiological stresses such as hypoxia and hypoglycaemia are potent mediators of gene expression. Hypoxia induces VEGF gene expression by upregulation of the transcription factors hypoxia inducible factor-1α and -2α (HIF-1α and HIF-2α). These transcription factors are expressed constitutively, but in the presence of oxygen are targeted for degradation by the ubiquitin–proteasome pathway. (The proteasome is an intracellular multiprotein complex that is responsible for the degradation of many cellular proteins. Proteins are targeted for destruction by the proteasome after being tagged with ubiquitin. The ubiquitin–proteasome pathway has a critical role in cell-cycle progression, gene expression and modulating the response to anticancer drugs). In hypoxia, HIF-1α and -2α are not destroyed and they translocate to the nucleus where they up-regulate the expression of a number of genes, including VEGF. Von Hippel–Lindau (VHL) disease is a rare inherited disorder characterized by the development of multiple bilateral tumours (renal cancer, cerebellar haemangioblastomas, pancreatic tumours, phaeochromocytomas and retinal angiomas) (Turner, 2000). The product of the *VHL* tumour-suppressor gene has a pivotal role in modulating the upregulation of VEGF. In RCC cells that are mutant for *VHL*, HIF-1α and -2α are expressed in both normoxia and hypoxia. Transfection with wild-type *VHL* restores hypoxic regulation of HIF-1α/2α expression. Recent work suggests that VHL is necessary for the oxygen-dependent degradation of HIF. VHL appears to bind to HIF and to 'tag' it for destruction by the proteasome (Maxwell *et al.*, 1999).

The hypoxia response pathway is an attractive therapeutic target. Hypoxic cells are resistant to conventional radiotherapy and chemotherapy. This problem might be overcome if drugs could be targeted specifically against hypoxic cells. Conversion of the prodrug 5-fluorocytosine to its active metabolite, 5-fluorouracil, is promoted by the bacterial enzyme cytosine deaminase. Increased expression of this enzyme has been induced *in vitro* using hypoxic induction of modified promoters and this has resulted in enhanced cell death (Dachs *et al.*, 1997).

Gene Therapy

Gene therapy is an attractive means of delivering antiangiogenic factors for a number of reasons (Lau and Bicknell, 1999). First, it is likely that most antiangiogenic compounds would need to be administered for prolonged periods if tumour dormancy is to be maintained. Gene therapy might allow a one-off administration of a vector that results in long-term production of the antiangiogenic peptide. Second, if gene therapy could be targeted preferentially at angiogenic regions of the tumour vasculature then therapeutic levels of an antiangiogenic factor could be achieved at the tumour site with minimal risk to physiological angiogenesis elsewhere. The use of gene therapy as a means to modulate angiogenesis is at an early stage but some success has been achieved in inhibiting the expression of proangiogenic mediators, in interference with signalling processes and in up-regulating inhibitors of angiogenesis. For example, transfection of glioblastoma cells with an antisense construct against VEGF results in reduced VEGF expression and reduced tumorigenicity in nude mice xenografts. Smaller and less aggressive tumours result when DNA encoding the antiangiogenic factors thrombospondin-1 (TSP-1) is transfected into breast carcinoma xenografts in comparison with untransfected cells (Kong and Crystal, 1998). Adenovirus-mediated transfer of angiostatin inhibits angiogenesis and growth of glioblastoma xenografts (Tanaka *et al.*, 1998). It has also been demonstrated that the KDR and E-selectin promoters are selectively expressed in ECs and therefore could be used for targeting the vasculature (Jagger *et al.*, 1997). Development of suitable vectors remains a major obstacle to gene therapy and the problems of efficient gene transduction and prolonged expression of the transgene have been difficult to overcome. Furthermore, the redundancy that exists in the angiogenic cascade means that inhibition of a single or limited number of factors may be of limited use. Future developments in antiangiogenic gene therapy may target the ECs directly.

Inhibitors of Proteolysis

Degradation of the ECM by proteolytic enzymes is essential for EC migration and tumour invasion. Many of these proteases are upregulated by VEGF and bFGF. The matrix metalloproteinases (MMPs) are an important group of degradative proteases that collectively are capable of degrading all the components of the ECM. Both synthetic and naturally occurring inhibitors of MMPs are in clinical trials. Marimastat, a zinc chelator, is a general MMP inhibitor. In initial studies, administration of marimastat achieved favourable reductions in the serum markers of a variety of cancers (prostate, colon, ovary and pancreatic) although this was not associated with tumour regression (Gore *et al.*, 1996). In a phase III study of 400 patients with advanced pancreatic cancer, marimastat was no more effective than gemcitabine, the 'drug of choice.' Further trials involving marimastat are under way. Better tolerated than marimastat is the MMP inhibitor Bay 12-9566, which is orally active and inhibits MMP-2 and MMP-9. Bay 12-9566 is the subject of several international phase III trials. A naturally occurring MMP inhibitor, Neovastat, has been derived from shark cartilage and appears to have antiangiogenic and anti-inflammatory activity. Results of phase I/II trials are awaited. The immunosuppressant cyclosporine inhibits both MMP-2 and MMP-9 in addition to inhibiting EC function, although its clinical role in antiangiogenesis has not yet been assessed.

Urokinase plasminogen activator (uPA) is another important ECM protease. Antibodies to the uPA receptor

suppress bFGF-induced angiogenesis and tumour growth in mice. Transfection of prostate cancer cells with mutant uPA resulted in smaller, less vascular tumours and fewer metastases in a rat xenograft model.

AGM 1470 is an antiangiogenic agent currently being assessed by phase III clinical trials. Most success so far has been achieved in the treatment of AIDS-KS, where partial responses have been achieved and the treatment well tolerated; a phase III trial is under way. Complete remission has also been reported in a patient with metastatic cervical cancer who underwent 22 months of treatment. The mechanism of action of AGM 1470 remains uncertain, although recently it was found to bind irreversibly to, and thus inhibit, the peptidase methionine aminopeptidase type 2 (metAP-2).

It has been demonstrated recently that new vessel formation *in vivo* requires transmembrane MMPs on the EC surface, which function as pericellular fibrinolysins. These EC-bound proteinases are a potential therapeutic target.

Direct Attack on Endothelial Cells and Vascular Targeting

Antiangiogenesis should be distinguished from vascular targeting. Antiangiogenesis involves therapy directed against biochemical targets aiming to inhibit new blood vessel formation and proliferation. In vascular targeting the established tumour vasculature is the therapeutic target; the aim is acute destruction of the tumour vasculature and tumour necrosis.

Transforming growth factor β (TGF-β) induces proliferation in angiogenic but not in nonangiogenic ECs. Part of the TGF-β receptor known as endoglin is recognized by TEC-11, a mouse monoclonal antibody. Staining of tissue sections has demonstrated that TEC-11 binds to ECs in a number of tumour types but not to the ECs in normal tissue. Some success has been achieved in targeting proliferating ECs *in vitro* by linking TEC-11 to the cytotoxic ricin A chain (Derbyshire and Thorpe, 1997).

A bispecific antibody has been developed where one arm is linked to inactive tissue factor and the other arm targets an experimentally induced marker on tumour vasculature ECs (Huang *et al.*, 1997). When the latter arm binds to tumour endothelium the inactive tissue factor is activated by exposure to cell-surface phospholipids. The coagulation cascade is initiated and tumour vessels thrombose. When this antibody was given intravenously to mice with subcutaneous neuroblastomas a third of the mice had complete tumour regression and no thrombi were seen in other organs.

Vascular targeting has been used in the treatment of soft tissue sarcomas of the extremities using TNFα. In this example, an isolated limb perfusion system was used to target TNFα to tumour-associated ECs, causing destruction of the microcirculation of the sarcoma.

CM 101 is a group b haemolytic streptococcus exotoxin that binds preferentially to tumour endothelium. It activates complement C3 and the resulting inflammatory reaction causes vessel disruption and tumour necrosis. Partial responses have been observed in phase I clinical trials (Harris, 1997).

Squalamine is an aminosteroid antibiotic first extracted from dogfish shark liver. It inhibits a sodium–hydrogen exchanger and has antiangiogenic and anti-infective activity. It may work by blocking EC activation and migration and by inhibiting growth factor expression. Phase I trials in brain and breast tumours are under way and a phase II study of squalamine in combination with taxol and carboplatin in the treatment of non-small cell lung cancer began in June 1999.

The cytoskeleton has an essential role in EC proliferation. The tubulin inhibitor Combretastin A4 causes proliferating ECs to apoptose, resulting in vessel shutdown. It is in phase I/II trials.

Other Compounds and Approaches

The interleukins have a variety of anti- and proangiogenic effects. Interleukin-10 (IL-10) inhibits *in vitro* angiogenesis by up-regulating TIMP-1 and by down-regulating MMP-2 and MMP-9 (Stearns *et al.*, 1999). Interleukin-12 (IL-12) induces interferon-γ (IFN-γ) and stimulates proliferation of natural killer (NK) cells. Both IFN-γ and NK cells demonstrate antiangiogenic activity and IL-12 is currently in phase II trials for AIDS-KS and other solid tumours. IM-862 is synthetic dipeptide that up-regulates IL-12 and which has been shown to be efficacious in a phase III trial against KS.

Paclitaxel is an anticancer agent licensed in the treatment of AIDS-KS and recurrent ovarian and metastatic brain tumours. It works by stabilizing microtubules. It may also have antiangiogenic activity and is currently under evaluation in this regard, although toxicity may restrict its usefulness.

Circulating levels of prostate-specific antigen (PSA) are used to screen for prostate cancer and to evaluate response to intervention. Despite its name, PSA is not prostate specific, having been detected in breast, lung and uterine cancers. Intriguingly, elevated PSA levels have been associated with better prognosis in breast cancer. A recent study has suggested that PSA has antiangiogenic activity (Fortier *et al.*, 1999). PSA inhibited EC migration, proliferation and invasion and impaired EC response to stimulation by FGF and VEGF. PSA also inhibited the development of melanoma metastases in a mouse xenograft model.

Recent work on non-small cell lung cancer has assessed the ratio of mature to immature vessels (the vessel maturation index, VMI) using a monoclonal antibody, LH39, which recognizes selectively an epitope unique to

mature vessels. A high VMI was associated with low angiogenesis, an absence of lymph node involvement and low levels of the angiogenic factor thymidine phosphorylase (Kakolyris *et al.*, 1999). This study highlights immature, although not necessarily proliferating, vessels as a future target in antiangiogenesis.

Antiangiogenesis and the Immune Response

Cells of the immune system are able to recognize and destroy tumour cells. There is significant interplay between angiogenesis and the immune response: leucocyte-derived cytokines are angiogenic stimulators, and the adhesion molecules expressed by tumour-associated ECs influences which leucocytes contribute to a localized inflammatory infiltrate. The expression of EC adhesion molecules such as ICAM-1, VCAM and E-selectin is induced by leucocyte derived chemokines such as TNFα, IL-1 and IFN-γ. By contrast, leucocyte adhesion to ECs is inhibited by TGF-β. Tumour microvessels are characterized by diminished adhesion of leucocytes to ECs (Griffioen *et al.*, 1998). It is likely that this reduced adhesion occurs because angiogenic factors down-regulate EC adhesion molecules. This reduced leucocyte adhesion might be advantageous to a tumour since it would impede the antitumour immune response. It has therefore been suggested that the antiangiogenic mediators might promote tumour cell destruction by the immune system by facilitating leucocyte adhesion to tumour-associated ECs (Griffioen *et al.*, 1998). Stimulation of cultured ECs by bFGF resulted in marked down-regulation of ICAM-1; the presence of the antiangiogenic peptide PF4 completely abrogated this effect. It may be that antiangiogenic factors that both inhibit new vessel formation and facilitate leucocyte–EC interaction will be particularly efficacious.

POTENTIAL PROBLEMS IN ANTIANGIOGENIC THERAPY

Whilst tumours cannot grow larger than 2–3 mm^3 without a new blood supply, they remain viable. Antiangiogenic therapy cannot therefore be ultimately cytotoxic. Hence, it is vital that antiangiogenic therapies are well tolerated since they are likely to be given repeatedly or continuously over prolonged periods. For the same reasons, it is important that drug resistance does not emerge. Since antiangiogenic therapies may only be cytostatic, lack of regression must not be equated with lack of effect, and novel ways of assessing response may be required. The utility of noninvasive means of determining tumour blood flow such as colour-flow Doppler and PET (positron emission tomography) scanning is under evaluation in the assessment of antiangiogenic therapies.

Interference with physiological angiogenesis is a concern. Treatment may have to be delayed postoperatively to allow for adequate wound healing, and disturbance of the menstrual cycle may occur.

There is a multitude of factors and pathways involved in angiogenesis and there is likely to be a degree of intrinsic redundancy. Blocked pathways may be circumvented and down-regulation of one factor may be compensated for by up-regulation of another. Some tumours may be able to divorce themselves entirely from angiogenesis-dependent growth. Human lung secondary tumours and some lung primaries may grow by a nonangiogenic-dependent mechanism by the invasion of pre-existing alveolar vessels (Pezzella *et al.*, 1997). There is a need to focus research effort on critical steps and identification of final common pathways. Multiple antiangiogenic treatments may be required synchronously for adequate clinical effect.

Care should be exercised in translating data from mouse xenograft experiments when planning clinical trials. The growth dynamics of experimental tumours may differ substantially from their human counterparts. Few human tumours grow as fast as typical mouse models and thus can be expected to demonstrate less intense angiogenic activity (Augustin, 1998).

THE FUTURE OF ANTIANGIOGENIC THERAPY

It is likely that antiangiogenic treatments will be used in conjunction with more established modalities that target the tumour cells and that are intended to be cytotoxic. Encouraging results have been achieved in preclinical studies (Teicher, 1996) and of particular note is the use of angiostatin in combination with radiotherapy (Mauceri *et al.*, 1998). The toxicity of this combination was selective for the tumour vasculature but did not result in increased toxicity toward normal tissue. Improved identification of markers specific to proliferating ECs will facilitate targeting of antiangiogenic therapy, particularly gene therapy. It may be that certain angiogenic mechanisms predominate in certain patients or tumours. Furthermore, different stages of the same tumour vary in their susceptibility to antiangiogenic agents. Recently, four different antiangiogenic factors were tested at three distinct stages of disease progression in a transgenic mouse model of pancreatic islet cell carcinogenesis (Bergers *et al.*, 1999). The four inhibitors tested showed distinct efficacy profiles depending on the stage of disease being treated. As understanding of angiogenesis grows it may be possible to tailor therapy toward the particular angiogenic profile of a patient or tumour. The results of antiangiogenic therapy to date are encouraging but the most recent preclinical data on angiogenesis inhibitors are particularly exciting. Antiangiogenic treatments, in combination with

other approaches, have great potential in the management of malignant disease and may be used in combination therapy and for maintenance therapy with all tumour types.

REFERENCES

Arora, N., *et al*. (1999). Vascular endothelial growth factor chimeric toxin is highly active against endothelial cells. *Cancer Research*, **59**, 183–188.

Augustin, H. G. (1998). Antiangiogenic tumour therapy: will it work? *Trends in Pharmaceutical Sceince*, **19**, 216–222.

Bergers, G., *et al*. (1999). Effects of angiogenesis inhibitors on multistage carcinogenesis in mice. *Science*, **284**, 808–812.

Boehm, T., *et al*. (1997). Antiangiogenic therapy of experimental cancer does not induce acquired drug resistance. *Nature*, **390**, 404–407.

Brooks, P. C. (1996). Cell-adhesion molecules in angiogenesis. *Cancer Metastasis Reviews*, **15**, 187–194.

Crew, J. P., *et al*. (2000). Eukaryotic initiation factor-4E in superficial and muscle invasive bladder cancer and its correlation with vascular endothelial growth factor expression and tumour progression. *British Journal of Cancer*, **82**, 161–166.

Dachs, G. U., *et al*. (1997). Targeting gene expression to hypoxic tumor cells. *Nature Medicine*, **3**, 515–520.

Dameron, K. M. *et al*. (1994). Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1. *Science*, **265**, 1582–1584.

Derbyshire, E. J. and Thorpe, P. E. (1997). Targeting the tumour endothelium using specific antibodies. In: Bicknell, R., *et al*. (eds), *Tumour Angiogenesis*. 345–346 (Oxford, Oxford University Press).

Dhanabal, M., *et al*. (1999). Endostatin induces endothelial cell apoptosis. *Journal of Biological Chemistry*, **274**, 11721–11726.

Dinney, C. P. N., *et al*. (1998). Inhibition of basic fibroblast growth factor expression, angiogenesis, and growth of human bladder carcinoma in mice by systemic interferon-α administration. *Cancer Research*, 808–814.

Fortier, A. H. *et al*. (1999). Antiangiogenic activity of prostate-specific antigen. *Journal of the National Cancer Institute*, **91**, 1635–1640.

Fox, S. B., *et al*. (1993). Relationship of endothelial cell proliferation to tumor vascularity in human breast cancer. *Cancer Research*, **53**, 4161–4163.

Goldman, C. K., *et al*. (1998). Paracrine expression of a native soluble vascular endothelial growth factor receptor inhibits tumor growth, metastasis, and mortality rate. *Proceedings of the National Academy of Sciences of the USA*, **95**, 8795–8800.

Gore, M., *et al*. (1996). Tumour marker levels during marimastat therapy. *Lancet.*, **348**, 263–264.

Griffioen, A. W., *et al*. (1998). Angiogenesis, a target for tumor therapy. *Journal of Laboratory Clinical Medicine*, **132**, 363–368.

Harris, A. L. (1997). Clinical trials of anti-vascular agent group B *streptococcus* toxin. *Angiogenesis*, **1**, 36–37.

Huang, X., *et al*. (1997). Tumor infarction in mice by antibody-directed targeting of tissue factor to tumor vasculature. *Science*, **275**, 547–550.

Jagger, R. T., *et al*. (1997). Endothelial cell-specific expression of tumor necrosis factor-alpha from the KDR or E-Selectin promoters following retroviral delivery. *Human Gene Therapy*, **8**, 2239–2247.

Kakolyris, S., *et al*. (1999). Assessment of vascular maturation in non-small cell lung cancer using a novel basement membrane component, LH39: correlation with p53 and angiogenic factor expression. *Cancer Research*, **59**, 5602–5607.

Kong, H.-L. and Crystal, R. G. (1998). Gene therapy strategies for tumor antiangiogenesis. *Journal of the National Cancer Institute*, **90**, 273–285.

Lau, K. K. W. and Bicknell, R. (1999). Antiangiogenic gene therapy. *Gene Therapy*, **6**, 1793–1795.

Lin, P., *et al*. (1997). Inhibition of tumor angiogenesis using a soluble receptor establishes a role for Tie-2 in pathologic vascular growth. *Journal of Clinical Investigation*, **100**, 2072–2078.

Mauceri, H. J., *et al*. (1998). The combined effects of angiostatin and ionizing radiation in antitumor therapy. *Nature*, **394**, 287–291.

Maxwell, P. H., *et al*. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*, **399**, 271–275.

Millauer, B., *et al*. (1994). Glioblastoma growth is inhibited *in vivo* by a negative dominant mutant Flk-1 mutant. *Nature*, **367**, 576–579.

O'Reilly, M. S., *et al*. (1994). Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell*, **79**, 315–328.

Pezzella, F., *et al*. (1997). Non-small cell lung carcinoma tumor – growth without morphological evidence of neo-angiogenesis. *American Journal of Pathology*, **151**, 1417–1423.

Rak, J., *et al*. (1995). Mutant ras oncogenes up-regulate vegf/vpf expression – implications for induction and inhibition of tumor angiogenesis. *Cancer Research*, **55**, 4575–4580.

Relf, M., *et al*. (1997). Expression of the angiogenic factors vascular endothelial growth factor, acidic and basic fibroblast growth factor, placenta growth factor, and pleiotrophin in human primary breast cancer and its relation to angiogenesis. *Cancer Research*, **57**, 963–969.

Scott, P. A. E., *et al*. (1998). Differential expression of vascular endothelial growth factor mRNA vs protein isoform expression in human breast cancer and relationship to eIF-4E. *British Journal of Cancer*, **77**, 2120–2128.

Singhal, S., *et al*. (1999). Antitumor activity of thalidomide in refractory multiple myeloma. *New England Journal of Medicine*, **341**, 1565–1571.

Stack, M. S., *et al*. (1999). Angiostatin inhibits endothelial and melanoma cellular invasion by blocking matrix-enhanced plasminogen activation: angiostatin binds ATP synthase on

the surface of human endothelial cells. *Biochemical Journal*, **340**, 77–84.

Stearns, M. E., *et al*. (1999). Interleukin-10 (IL-10) inhibition of primary human prostate cell-induced angiogenesis: IL-10 stimulation of tissue inhibitor of metalloproteinase-1 and inhibition of matrix metalloproteinase (MMP)-2/MMP-9 secretion. *Clincal Cancer Research*, **5**, 189–196.

Tanaka, T., *et al*. (1998). Viral vector-targeted anti-angiogenic gene therapy utilizing an angiostatin complementary DNA. *Cancer Research*, **58**, 3362–3369.

Teicher, B. A. (1996). A systems approach to cancer therapy. *Cancer Metastasis Reviews*, **15**, 247–272.

Turner, K. (2000). Inherited renal cancer. *British Journal of Urology International*, **86**, 155–164.

Yuan, F., *et al*. (1996). Time-dependent vascular regression and permeability changes in established human tumour xenografts induced by an anti-vascular endothelial growth factor/vascular permeability factor antibody. *Proceedings of the National Academy of Sciences of the USA*, **93**, 14765–14770.

FURTHER READING

Teicher, B. A. (ed.) (1999). *Antiangiogenic Agents in Cancer Therapy*. (Humana Press, Totowa, NJ).

Websites

http://cancertrials.nci.nih.gov/NCI_CANCER_TRIALS.
http://www.angio.org.

Chapter 94

Targeting the Extracellular Matrix

Raffaella Giavazzi and Giulia Taraboletti
Mario Negri Institute for Pharmacological Research, Bergamo, Italy

CONTENTS

- Targeting the Extracellular Matrix of Tumours
- Therapeutic Strategies Exploiting the Extracellular Matrix of Tumours
- Use of Matrix-related Molecules for Selective Delivery of Therapeutic Agents
- Matrix-targeted Agents can Act with Multiple Mechanisms of Action
- Conclusion

TARGETING THE EXTRACELLULAR MATRIX OF TUMOURS

The extracellular matrix (ECM), a complex mesh of collagens, proteoglycans and glycoproteins, is a vital supporting scaffold of tissues, as it contributes to their architecture, isolates tissue compartments and acts as a selective macromolecular filter. An important role for the ECM in regulating cell adhesion, motility, differentiation, proliferation and survival is now well recognized (for details, see the chapters *Extracellular Matrix: The Networking Solution* and *Invasion and Metastasis*). The rationale of selectively targeting the tumour ECM is based on the fact that it differs from that of normal tissues in terms of:

- composition: the tumour matrix presents either a different proportion of matrix components or molecules usually not found in the ECM of normal tissues;
- molecular organisation (architecture): the tumour matrix is usually poorly organized;
- turnover: a high matrix-degrading proteolytic activity is a hallmark of tumour malignancy and is necessary for tumour invasion and metastatic dissemination.

Three main molecular players act in the interaction between tumour cells and the ECM:

- matrix molecules: besides acting as a substrate for cell adhesion and tissue organisation, the ECM regulates important cell functions, many of which are related to tumour progression (such as motility, proliferation, invasiveness); therefore, the differences in matrix composition and architecture found in tumours may have profound effects on the behaviour of tumour cells;
- receptors: tumour cells often express an altered pattern of receptors for matrix molecules (mainly integrins), therefore determining a different response/interaction to matrix;
- matrix-degrading enzymes: these enzymes, produced by tumour or stromal cells, greatly affect cell interaction with the extracellular environment.

Although these molecules are ‘normal’ host structures, their association with the transformed tissues makes them quantitatively or qualitatively different from their counterparts in normal tissues, and therefore recognizable by an appropriate treatment. Agents that affect the complex, coordinated interplay between matrix molecules, receptors and proteases can alter the tumour homeostasis, hence affecting tumour progression.

Therapeutic strategies include (**Table 1**):

- agents affecting tumour cell interaction with the ECM, including peptides or recombinant fragments of the active sequences of matrix components, integrin antagonists and ligands of matrix components;
- molecules that modulate the deposition of matrix components;
- inhibitors of matrix degradation.

THERAPEUTIC STRATEGIES EXPLOITING THE EXTRACELLULAR MATRIX OF TUMOURS

Functional Domains of ECM Molecules

The concept of using small matrix-derived molecules for treatment purposes has emerged from studies in the last two decades, showing that synthetic peptides corresponding to minimal active sequences of cell adhesion molecules, such as laminin and fibronectin, where able to prevent metastasis in experimental models (Humphries *et al*., 1986). Since then, a great amount of experimental evidence has accumulated, which confirmed that isolated

Table 1 Examples of therapeutic strategies exploiting tumour-associated extracellular matrix

Analogues of the functional site of matrix components
- Synthetic peptides or recombinant fragments (*endostatin, tumstatin*)
- Molecules that mimic functional regions (*heparin-like agents, laminarin, suramin*)

Integrin antagonists
- Synthetic peptides, recombinant domains containing integrin recognition sequences (*RGD, NGR*)
- Disintegrins
- Nonpeptide integrin antagonists
- Anti-integrin antibodies (*vitaxin*)
- Agents modifying integrin expression (*cytokines*)
- Inhibitors of integrin signalling (*genistein, CAI*)

Molecules that modulate the deposition of matrix components
- Cytokines (*IL-1, TNF-α, IFNs*)
- Differentiating agents (*retinoids*)
- Other (*halofuginone*)

Inhibitors of matrix degradation
- Inhibitors of the serine-, cysteine- and matrix metalloproteases (see **Table 2**)
- Inhibitors of protease synthesis (*cytokines, glucocorticoids, retinoids, halofuginone*)
- Bisphosphonates (*clodronate*)
- Inhibitors of heparanases (*nonanticoagulant heparins, PI-88*)

domains of matrix molecules – usually synthetic peptides or recombinant fragments – could be used as the basis for developing therapeutic agents. Depending on the biological activity of the original molecule, isolated domains could be employed either to mimic an antineoplastic activity of the intact molecule or to act as competitors of its pro-malignant activity.

Recombinant Domains

Given the presence of cryptic sites in matrix components – regions usually not exposed in the native molecule – the activity of isolated domains can be totally unrelated to that of the original, intact molecule (Davis *et al.*, 2000). Examples of this have recently been provided by the identification of fragments of collagen or fibronectin endowed with antineoplastic activity.

Collagen Fragments

Fragments from different collagen molecules have been described to have antiangiogenic and antineoplastic activity. Examples are a noncollagenous fragment of type XVIII collagen (known as endostatin) (O'Reilly *et al.*, 1997), the NC10 domain of type XV collagen (restin, homologous to endostatin) and the NC1 domains of different α-chains of type IV collagen: $\alpha 1$ (also named arresten), $\alpha 2$ (canstatin), $\alpha 3$ (tumstatin) and $\alpha 6$. Although these isolated domains have antineoplastic activity, native collagens type IV, type XVIII and type XV are devoid of such activity. All these compounds have antiangiogenic activity, and endostatin is currently in clinical trials (see the chapter *Antiangiogenic Therapy*). Fragments of type IV collagen could also act by blocking the interaction of tumour cells with the basement membranes, or, since the NC1 domain is thought to be involved in collagen assembly, they could affect matrix organisation. In addition, at least for one of them, tumstatin, a sequence – distinct from the antiangiogenic one – has been shown to have a direct antiproliferative activity for melanoma cells (Maeshima *et al.*, 2000).

Fibronectin Fragments

An example is Anastellin, a III1-C fibronectin fragment, derived from the first type III repeat of fibronectin, has been described to inhibit growth of experimental tumours, mainly acting through an antiangiogenic mechanism. Anastellin is thought to act by altering matrix assembly or interfering with the avβ3 receptor on endothelial cells (Yi and Ruoslahti, 2001).

Synthetic Peptides

Synthetic peptides corresponding to active sequences of matrix molecules, such as fibronectin (RGD, PHSRN, LDV, REDV), laminin (YIGSR, IKVAV) and thrombospondin-1 (GSVTCG, WSPW), have been studied in experimental models (Yamada, 1991). They have shown high potential in inhibiting cell functions associated with the malignant behaviour of tumour cells, including angiogenesis and metastatic dissemination of experimental tumours.

It has been a general experience that small peptides are poor therapeutic agents since, *in vivo*, they diffuse rapidly, are often susceptible to proteolytic degradation and are subjected to a rapid clearance. Chemical modifications of the peptides have been introduced to improve their pharmacological characteristics, including the preparation of multimeric or cyclic peptides (see also Integrin Antagonists, below), use of D-amino acids, synthesis of *retro–inverso* peptidomimetics, conjugation with molecules (polyvinylpyrrolidone, polyethylene glycol or the heparin-like 6-*O*-carboxymethylchitin) or addition of chemical groups and spacers. These modifications improve the therapeutic index of the peptides by increasing their resistance to enzymatic degradation and hence stability in biological fluids, and by altering the peptide clearance. Modified peptides can be radiolabelled for diagnostic use and for radionucleotide therapy.

Molecules Similar to Functional Domains

Natural or synthetic molecules acts as competitive inhibitors of matrix components, thanks to their structural similarity.

Heparin-like Molecules

Low molecular weight heparins, modified heparins lacking the anticoagulant activity, and heparin-like agents, such as laminarin (a sulphated polysaccharide) and suramin (a polysulphonated naphthylurea) which are structural mimics of heparan sulphates, have a potential therapeutic effect. Preclinical and clinical observations suggest that the antineoplastic effect of these molecules is not necessarily related to their anticoagulant activity. Heparin-like molecules have been reported to inhibit cell proliferation by preventing the binding of growth factors to their cell surface low affinity receptors. In addition, many of these compounds have the ability to block heparanase, a heparan sulphate-degrading enzyme associated with the metastatic potential of tumours (see Inhibitors of Heparanases, below). These compounds have been shown to prevent metastasis formation and angiogenesis in preclinical models. Information from early phases clinical trials with low molecular weight heparins indicate an improved cancer outcome and a favourable pharmacological profile (Zacharski *et al.*, 2000).

Integrin Antagonists

Among the great variety of receptors that mediate tumour cell interaction with matrix components, integrins appear to play a major role in tumour progression (Albelda, 1993). The pattern of integrin expression is known to change during cell transformation, and such alterations profoundly affects malignant tumour behaviour. The modulation of integrin expression or the selective block of integrin functions can suppress tumorigenicity and/or metastatic potential in experimental tumour models. This led to the development of specific inhibitors for this class of receptors. An integrin currently drawing great interest in terms of therapeutic applications is $\alpha v\beta 3$, implicated in the progression of human melanoma, in angiogenesis and in the process of cell survival. In addition, $\alpha v\beta 3$ is involved in the proteolytic activity of both tumour and endothelial cells, through binding to MMP-2.

Strategies to block the activity of these receptors include peptide and non-peptide integrin antagonists, disintegrins, anti-integrin antibodies, agents modifying integrin expression on tumour cells and inhibitors of integrin signalling.

RGD-related Peptides

The antineoplastic activity of peptides designed on the integrin-recognition sequence RGD is due to various mechanisms, including a direct antimetastatic activity on tumour cells (Humphries *et al.*, 1986) and inhibition of endothelial cell functions. A cyclic RGD peptide (EMD 121974) is currently undergoing clinical trials as an anti-angiogenic compound. In addition to the integrin antagonist activity, a recent report indicates that RGD-containing peptides induce apoptosis in an integrin-independent way, through direct intracellular activation of caspase 3, a proapoptotic protein.

Non-peptide Integrin Antagonists

They have been identified, taking advantage of chemical libraries generated by combinatorial organic synthesis. These compounds could represent an alternative to the use of peptides that, as mentioned above, are usually poor pharmacological agents. This approach has led to the identification of nonpeptide inhibitors of integrins. One of them (TSRI265), has recently been shown to block collagen degradation by tumour cells as well as angiogenesis, by preventing the RGD-independent interaction of $\alpha v\beta 3$ with MMP-2 (Silletti *et al.*, 2001).

Disintegrins

These molecules, thanks to the presence of an RGD sequence, have the potential to block integrins, hence their name. Several disintegrins (i.e. salmosin, albolabrin, rhodostomin, accutin), have been described to inhibit tumour cell adhesion, invasion, proliferation, extravasation and experimental metastasis. Moreover, as they also interact with integrins on endothelial cells, disintegrins have been reported to inhibit angiogenesis. The therapeutic potential of these compounds is still to be defined.

Anti-integrin Antibodies

Based on the selectivity of $\alpha v\beta 3$ as a marker of angiogenic blood vessels, antibodies against $\alpha v\beta 3$ have been developed as antiangiogenic therapy (Eliceiri and Cheresh, 1999). Vitaxin is a humanized version of the LM609 monoclonal antibody, which blocks the binding of $\alpha v\beta 3$ to adhesion proteins and disrupts ECM-induced signalling, leading to the activation of p53 and to apoptosis of endothelial cells. Vitaxin is currently in early clinical trials as an inhibitor of angiogenesis.

Agents Modifying Integrin Expression

Factors that modulate the expression of integrins on the surface of both tumour and endothelial cells can affect tumour progression. It has been proposed that cytokines, i.e. TNF and IFNγ, suppress $\alpha v\beta 3$ activation, causing detachment and apoptosis of endothelial cells (Rüegg *et al.*, 1998).

Integrin Signalling

Signalling transduction pathways represent another potential target to block tumour cell interaction with the ECM. Different steps of the signalling cascade have been targeted for therapeutic intervention. Inhibitors of tyrosine kinases (such as genistein), of PKC (staurosporine, sphingosine) and of calcium influx (carboxyaminotriazole) have also

been reported to affect matrix-driven tumour cell behaviour. Some of these compounds are currently being subjected to clinical trials.

Molecules that Modulate the Deposition of Matrix Components

Since the ECM composition can profoundly influence the behaviour of tumour cells, agents that affect matrix synthesis–deposition have been studied in tumour progression. Given the critical role of the matrix in angiogenesis, many of these compounds have also been found to have antiangiogenic activity. Several cytokines (such as interleukin-1β and tumour necrosis factor α) and agents that control cell differentiation have been shown to affect matrix synthesis and deposition.

Retinoids

Retinoids, natural and synthetic analogues of vitamin A, are currently being subjected to clinical trials in cancer prevention and therapy. Their antineoplastic activity is the result of several different effects on tumour cell differentiation and interaction with the ECM. Indeed, they have been described as inhibiting the production of proteases and their inhibitors, reducing cell migration and invasion, altering cell adhesion to the matrix and modulating the expression of integrins by tumour cells.

Halofuginone

This low molecular weight quinazoline alkaloid, originally approved for use as an antiparasitic for farm animals, is currently attracting great interest as a pharmacological inhibitor of matrix synthesis and deposition. Halofuginone inhibits gene expression of the $\alpha 1$ chain of type I collagen. In addition, it also acts as an inhibitor of proteases, since it cause a transcriptional suppression of MMP-2 synthesis. The antineoplastic activity of halofuginone, observed in preclinical models, is attributed to a combined antimetastatic activity on tumour cells (reduced ability to invade and extravasate) and antiangiogenic activity (Elkin *et al*., 2000). This compound is under development for treating solid tumours and other diseases associated with fibrosis.

Inhibitors of Matrix Degradation

Tumour-associated lytic enzymes, produced by tumour cells and the adjacent stroma-infiltrating cells, are responsible for the remodelling of the matrix typically found in tumours. Furthermore, matrix-degrading enzymes exert functions that go beyond the physical destruction of connective tissue barriers, including activation of receptors and cell surface molecules, release of growth and angiogenic factors stored in the matrix and generation of active fragments from matrix components. Altogether this makes the proteolytic system highly involved in the process of tumour progression, metastasis and angiogenesis and thus a relevant target for therapeutic interventions. Inhibitors of metalloproteinases (MMPs), serine proteinases (the plasminogen activator system) and cysteine proteinases (cathepsins), as well as inhibitors of heparanases are under extensive investigation in cancer. Owing to their double role in tumour cell invasion and in angiogenesis (see the chapter *Invasion and Metastasis*), they are under development as both antimetastatic and antiangiogenic agents. The blockade of matrix degradation has shown efficacy in inhibiting cell invasion, tumour growth, angiogenesis and metastasis in a variety of *in vitro* systems and animal models. The promising results of the preclinical studies have rapidly led to clinical trials for several of these inhibitors. It is worth noting that the different proteolytic systems interact with each other, through a complex cross-regulation of protease activation among the members of all families. Therefore, the combination of these inhibitors is now foreseen as the optimal treatment in blocking tumour progression.

Matrix Metalloproteinase Inhibitors (MMPIs)

MMPIs are the proteinase inhibitors most advanced in clinical trials (Hidalgo and Eckhardt, 2001). The great interest in these molecules has been generated by earlier studies showing that the addition of exogenous TIMPs or the transfection with TIMP–cDNA reduced tumorigenicity, local invasiveness, metastasis as well as angiogenesis in several experimental models. TIMPs are not pharmacologically suitable for clinical use. Therefore, several companies have focused their efforts on the development of synthetic MMPIs (Giavazzi and Taraboletti, 2001). They differ in origin, MMP inhibitory profile and zinc-binding group (**Table 2**).

Table 2 Matrix metalloproteinase inhibitors entered in clinical trials

Compound	Origin	Status[b]
Batimastat (BB94)	Hydroxamic[a]	Phase I
Marimastat (BB-2516)	Hydroxamic[a]	Phase II–III
Prinomastat (AG-3340)	Hydroxamic[a]	Phase II–III
Metastat (COL-3)	Tetracycline derivative	Phase I
BAY 12–9566	Carboxylic[a]	Phase II–III
Neovastat (Æ-941)	Cartilage derivative	Phase I–IIa
BMS-275291	Sulfhydryl[a]	Phase I
CGS 27023A	Hydroxamic[a]	Phase I

[a]The zinc-binding group is indicated.
[b]As of May 2001.

A key structural feature of MMP inhibitors is the presence of a metal-binding group that is capable of chelating the zinc ion at the active site of the MMP. The first compounds to be tested in cancer patients – batimastat, CGS27023A, prinomastat and marimastat – all have a hydroxamate binding group. Nonhydroxamic acid inhibitors were synthesized later. Recently, approaches based on the screening of compound libraries have successfully identified novel compounds with MMP inhibitory activity. The rational design of inhibitors has received a great input by the recent advent of high-resolution X-ray and NMR analysis of the structures of the MMP–inhibitor complex.

The inhibition profiles of the different MMPIs vary substantially across the MMP family, the main targets being MMP-2 and MMP-9, that are associated with tumour malignancy in many types of cancers. The significance of the inhibitory profile of the different MMPIs in clinical development is the subject of debate. From the early clinical trials with broad-spectrum MMPIs (e.g. marimastat), the observation of adverse effects suggested that the use of more selective inhibitors (e.g. BAY 12-9566) should provide greater specificity, and hence limit side effects. On the other hand, it is now clear that more than one MMP family member is involved in the course of the disease, this being in favour of a broad-spectrum inhibitor. Requisites for suitable MMP inhibitors are also acceptable pharmacokinetics and high oral bioavailability.

Batimastat (BB-94)

The best studied MMPI in preclinical models is batimastat, a synthetic hydroxamate, with a broad spectrum of activity on MMPs including on MMP-1, -2, -3, -7 and -9. Batimastat was tested as a suspension administered by intracavity injection to patients with malignant ascites and pleural effusion. Its clinical development was suspended owing to its poor solubility and poor local tolerability.

Marimastat (BB-2516)

Related to batimastat, this compound is orally available and is currently under investigation in several phase II–III clinical trials. Phase I studies in healthy volunteers showed plasma levels exceeding the concentrations required to inhibit MMP-2. Phase II studies have been conducted on patients with advanced ovarian, colorectal, pancreatic and prostatic cancers, all tumour types for which high expression of MMP-2 and MMP-9 has been described. In these studies the combined analysis of the modulation of serum tumour markers and of drug-related toxicity has been proposed to select the dose range for long-term studies. The main reported drug-related adverse event is a reversible musculoskeletal syndrome. Marimastat is the first MMPI to have completed phase III studies (Brown, 2000). In the first phase III study, on patients with advanced pancreatic cancer, the survival of patients receiving marimastat was not superior to that of patients receiving gemcitabine. In a second phase III trial, on patients with gastric cancer, there was a modest, but nonsignificant, improvement in survival in the marimastat-treated arm compared with placebo. Two marimastat SCLC studies are due to report very soon.

Prinomastat (AG-3340)

Another synthetic hydroxamate, prinomastat, inhibits MMP-2, -3, -9 and -13, but it is poorly active on MMP-1. It was selected among a series of structurally related inhibitors of MMPs, synthesized using a protein structure-based drug design. In patients with advanced cancer, phase I studies with AG-3340 given orally twice daily showed no dose-limiting toxicity. However, also for this hydroxamate, at higher doses reversible joint-related complaints were reported. Phase III trials in combination with chemotherapy in hormone-refractory prostate cancer and advanced non-small-cell lung cancer (NSCLC) did not meet primary efficacy objectives.

CGS 27023A

Designed as a stromelysin (MMP-3) inhibitor, this compound has a broad-spectrum activity. Phase I clinical studies indicate that the drug is generally well tolerated; the major toxicity was cutaneous rash. It is believed that this is one of the factors that prevented the progression of the compound to advanced clinical trials.

BMS-275291

This compound has completed phase I dose escalation in normal volunteers and is expected to enter phase II–III trials on advanced cancer. BMS-275291, structurally different from and more selective than the first generation of MMPIs, is expected to have an improved safety profile.

BAY 12-9566

This nonpeptidic, noncollagen-mimicking, biphenyl MMPI with a carboxylic zinc binding group is an example of an MMPI that claims a selective spectrum of activity mainly against MMP-2 but not MMP-1. Phase I clinical studies indicated that the compound is generally well tolerated. Drug-related adverse effects were transient, with asymptomatic transaminase elevation and mild thrombocytopenia; no musculoskeletal pain has been reported. However, the main phase III trials of BAY 12-9566 in small-cell lung and pancreatic cancers have been stopped because of negative findings and the development of BAY 12-9566 has been halted.

Neovastat (Æ-941)

This complex, derived from shark cartilage, has multiple mechanisms of action that include the inhibition of MMPs and interaction with the VEGF receptor. The lack of a clear identification of the MMPI activity in the extract makes its development difficult. Phase I–II trials are being conducted with Æ-941 in refractory lung, prostate and breast cancer. No serious drug-related adverse events have been reported and a reduction in PSA levels has been observed

in prostate cancer. Phase III trials in patients with renal cell carcinoma and NSCLC are planned.

Tetracycline derivatives

A variety of nonpeptidic natural product MMP inhibitors have been discovered by screening. These include the tetracyclines minocycline and doxocycline and the newer tetracycline analogues that have been chemically modified to eliminate the antibiotic activity. Metastat (COL-3) is a representative compound of this class currently in early clinical trials.

So far, the results from clinical trials with MMPIs have been rather disappointing (Hidalgo and Eckhardt, 2001). As indicated by the results on animal cancer models, MMPIs are intended to prevent tumour growth and spread, but are unable to control heavy tumour burden. Clinical benefit with MMPIs in cancer is therefore likely to be confined to patients with early stage or low burden cancer. By contrast, the majority of the clinical studies with MMPIs are placebo-controlled advanced-stage cancer. This was pointed out by Brown (2000) in the analysis of the results on gastric cancer studies with marimastat. Another observation from the preclinical studies is that tumours regrow after suspension of the administration of the inhibitor. Given the nature of their mechanism of action, MMPIs are not expected to interfere with cell proliferation and, therefore, these compounds are not intended to affect tumour cells as cytotoxic agents do. These observations have led to investigations of the antitumour activity of MMPIs in combination with conventional cytotoxic drugs. Several experimental studies have shown that the combination of MMPI with cytotoxic agents improves tumour response in both drug-sensitive and -resistant tumours and increases animal survival. These findings strongly suggest the use of MMPIs to potentiate the activity of cytotoxic drugs or to maintain the response after reductive therapy with a cytotoxic drug. Finally, another difficulty in the clinical evaluation of the activity of MMPIs has been the absence of valid surrogate endpoints to determine the optimum dose and to identify responsive tumours.

Additional preclinical studies are now needed in order to investigate the possibility that other MMPs (rather than those targeted so far) could be more efficacious targets for therapeutic block of the proteolytic activity of tumours. In this respect, the recently demonstrated important role for MT1-MMP could indicate this protease as one of the new targets. In addition, it is crucial to understand at which stage of the disease a patient would be more likely to benefit from therapy with protease inhibitors (Bergers *et al*., 1999).

Inhibitors of the Plasminogen Activation System

Generation of the serine proteinase plasmin from the plasminogen activation system, which includes the urokinase- and tissue-type plasminogen activators uPA and tPA, the serpins PAI-1 and PAI-2 and the uPA receptor (uPAR), represents a promising target for anti-invasive therapy (Mignatti and Rifkin, 2000). This is proved by the documented role of uPA-catalysed plasmin generation in cancer cell invasion through the extracellular matrix. Moreover, recent findings indicate an additional role for this system in inducing cell migration and invasion through plasmin-independent mechanisms, due to interactions among uPA, uPAR, PAI-1, ECM proteins, receptors and growth factors. The most promising results in cancer therapy have been obtained with the uPAR blockade. Antibodies, antisense oligonucleotides and uPAR antagonists, including small molecular weight peptide inhibitors, affect tumour invasion, metastasis, angiogenesis and tumour growth in animal models.

Inhibitors of Cathepsins

The abnormal extracellular location of these lysosomal aspartyl or cysteine proteases in tumours makes these enzymes potential targets to block tumour-associated matrix degradation. The levels of cathepsins and their endogenous inhibitors have been proposed as prognostic markers in cancer. Pharmacological inhibitors still have to be developed.

Bisphosphonates

These osteoprotective compounds have been shown to be effective in preventing osteolytic metastases. Analogues of pyrophosphate, they bind to hydroxyapatite on bone surface, preventing bone resorption. In addition, they also inhibit tumour cell adhesion to mineralized matrices, reduce the expression of MT1-MMP and are cytotoxic for macrophages and osteoclasts. Clinical trials (with clodronate and pamidronate) have indicated that bisphosphonates have low toxicity and, albeit with some contradictory results, suggest the effectiveness of these compounds in adjuvant therapy in patients at high risk of developing metastases to bone (Diel and Mundy, 2000). Unexpectedly, adjuvant clodronate has been reported to decrease visceral metastasis in women with breast cancer at high risk of distant metastasis. This finding confirms the preclinical evidence that, apart from the known osteoprotective effects, bisphosphonates also have a direct effect on the adhesion and invasiveness of tumour cells.

Inhibitors of Heparanase

Heparanases are heparan sulphate-degrading endoglycosidases, the expression of which correlates with the metastatic potential of tumour cells. Heparanases are also involved in angiogenesis, both directly (by stimulating endothelial cell invasion) and indirectly (by releasing and promoting the activity of matrix-bound angiogenic

factors) (Vlodavski *et al*., 1999). Heparanase inhibitors, non-anticoagulant heparin species or sulphated oligosaccharides (which are structural mimics of heparan sulphate), decrease the primary tumour growth, metastasis and vascularity of experimental tumours. The heparanases inhibitor PI-88 has been shown to be a potent inhibitor of metastasis and angiogenesis. It has received approval for trials in cancer patients.

USE OF MATRIX-RELATED MOLECULES FOR SELECTIVE DELIVERY OF THERAPEUTIC AGENTS

Molecules targeting the matrix or receptors can be employed to improve the delivery of therapeutic agents, such as cytotoxic drugs and vectors for gene therapy. The use of vehicles that have a high affinity for molecules specifically expressed at the tumour site allows the selective delivery of an effective dose of therapeutic agent to the neoplastic tissue with the advantage of reduced general toxicity and dose-limiting side effects.

Anti-integrin Antibodies

Cytotoxic drugs have been conjugated to anti-integrin antibodies, in order to achieve more efficient targeting of the integrin-expressing tumour cells. Coupling a doxorubicin derivative to the LM609 antibody, which recognizes the $\alpha v\beta 3$ on melanoma cells, improved the antineoplastic effect of the drug in mice.

RGD-related Peptides and Matrix Binding Sequences

Peptides recognizing adhesive receptors of tumour cells have been used as vehicles to improve the targeting of cytotoxic drugs to the neoplastic tissues. The RGD-related NGR peptide has been selected for high tumour-homing '*in vivo*' using a phage-displayed peptide library. This peptide, which binds to aminopeptidase N (CD13) on the tumour vasculature, when coupled to doxorubicin has been shown to increase the antitumour properties of the cytotoxic drug in an experimental model (Arap *et al*., 1998).

Receptor-recognizing sequences of matrix components can control tissue tropism of gene therapy. Integrin-binding peptides synthesized with a polylysine chain to form an electrostatic complex with DNA can function as a DNA vector. This has been found useful in the design of nonviral vectors for gene therapy; the presence of specific peptides directs gene delivery to selected cell types.

Incorporation of peptides or matrix-binding sequences by molecular engineering of envelope 'escort' proteins has been used to improve tissue tropism of viral vectors. Incorporation of NGR peptides in envelope proteins has been used to improve the specific targeting of retroviral vectors to the tumour vasculature, in experimental models. Similarly, retroviral vectors bearing the collagen-binding sequences of von Willebrand factor have been shown to accumulate at the tumour site (where collagen matrix is exposed), strongly enhancing the therapeutic effect of the gene therapy.

Antibodies Recognizing the ECM of Tumour Vessels

Antibodies recognizing tumour matrix molecules selectively expressed in tumours have been proposed for diagnostic imaging or as selective vehicles for treatment, coupled to toxic agents. The technology of antibody phage display libraries has led to the development of recombinant antibody fragments to be used as tumour-targeting agents. Potential targets include fibronectin and tenascin isoforms. A human antibody fragment that recognizes the spliced repeat C of tenascin has been shown to localize *in vivo* around vessels and proliferating cells in malignant tumours (high-grade astrocytoma). A high-affinity human antibody fragment (L19) that recognizes the fibronectin isoform containing the ED-B oncofoetal domain has been shown to localize selectively to newly formed blood vessels in tumour tissues. Recently, L19 antibody fused to tissue factor has been shown to induce the selective infarction and eradication of solid tumours in mice (Nilsson *et al*., 2001).

Tumour-associated Proteases

As mentioned above, high production of proteolytic enzymes is a hallmark of aggressive tumours. The possibility has therefore been explored of using these tumour-associated enzymes as local activators of prodrugs. Peptide prodrugs activated by MMPs, plasmin and γ-glutamyl-transpeptidases have been shown to have a protease-dependent cytotoxic effect in preclinical studies.

Fibronectin

A fragment of fibronectin greatly increases the efficiency of retroviral-mediated gene transfer into human haematopoietic cells, *ex vivo*, through a mechanism probably involving colocalization of retroviral particles and target cells. A protocol using a recombinant fibronectin fragment used to assist in *ex vivo* retroviral transduction is currently being subjected to clinical trials.

MATRIX-TARGETED AGENTS CAN ACT WITH MULTIPLE MECHANISMS OF ACTION

An important concept that emerges from the above description of matrix-targeting compounds is that many of them act through multiple mechanisms. Therefore, they can be proposed for different therapeutic approaches.

In some cases, the same compound can act on different molecular targets. Halofuginone inhibits the synthesis of collagen $\alpha 1$(I) and of the protease MMP-2 (see the section Molecules that Modulate the Deposition of Matrix Components, above). Heparin-like molecules can affect coagulation, block the binding of growth factors to cells and act as inhibitors of heparanase (see the sections Molecules that Mimic Functional Domains of ECM Components and Inhibitors of Heparanase, above). The differentiating agents retinoids also modulate the expression of integrins and the production of proteases and matrix components (see the section Molecules that Modulate the Deposition of Matrix Components, above).

The complexity of activities can also occur at the level of cellular targets. Since similar molecular mechanisms govern the invasive processes of tumour and endothelial cells, compounds acting on these common mechanisms can affect both tumour invasion and angiogenesis. The final antineoplastic activity of collagen fragments, integrin antagonists, halofuginone and inhibitors of matrix-degrading enzymes (see the section Inhibitors of Matrix Degradation, above) is ascribable to a combined anti-metastatic and antiangiogenic activity.

Naturally occurring examples of functional complex molecules are provided by the ADAMTS. As indicated by the acronym (*A Disintegrin And Metalloproteases with ThromboSpondin* motifs), these molecules include different active sequences and consequently exert multiple activities. Two of them, METH-1/ADAMTS1 and METH-2/ADAMTS8, have been reported to inhibit angiogenesis (Vazquez *et al*., 1999).

A large family of dietary constituents that, on the basis of epidemiological studies, are suggested to exert a cancer chemopreventive effect, have been described as acting through multiple mechanisms. Among these, epigallocatechin-3-gallate, the major flavonoid of green tea, has been reported to inhibit proteases involved in tumour cell invasion and angiogenesis.

CONCLUSION

Targeting tumour matrix-associated molecules and, more in general, tumour stroma is a relatively new way of thinking in cancer treatment. Interactions between ECM and tumour cells on the one hand and between ECM and host cells on the other play a pivotal role in tumour progression. Metastasis formation, the angiogenic process and the host response to tumour transformation are affected by the environment, of which ECM is a major component. This has stimulated the development of a range of antagonists of matrix receptors and their ligands and inhibitors of ECM-associated functions. The complexity of the underlying molecular mechanisms (multiple classes of molecules interact with each other) and the involvement of distinct pathways at different stages of tumour progression make the development of successful therapy difficult. The identification of selective tumour-associated ECM targets, the development of compounds with suitable pharmacological characteristics and their use in the correct combinations will constitute the main focus of future research efforts.

Although at this time the development of successful drugs is clearly more arduous than suggested by preclinical studies, agents targeting the tumour environment can be considered good candidates to potentiate, rather than replace, conventional antitumour therapies.

REFERENCES

Albelda, S. M. (1993). Role of integrins and other cell adhesion molecules in tumor progression and metastasis. *Biology of Disease*, **68**, 4–16.

Arap, W., *et al*. (1998). Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science*, **279**, 377–380.

Bergers, G., *et al*. (1999). Effects of angiogenesis inhibitors on multistage carcinogenesis in mice. *Science*, **284**, 808–812.

Brown, P. D. (2000). Ongoing trials with matrix metalloproteinase inhibitors. *Expert Opinions in Investigative Drugs*, **9**, 2167–2177.

Davis, G. E., *et al*. (2000). Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules. *American Journal of Pathology*, **156**, 1489–1498.

Diel, I. J. and Mundy, G. R. (2000). Bisphosphonates in the adjuvant treatment of cancer: experimental evidence and first clinical results. *British Journal of Cancer*, **82**, 1381–1386.

Eliceiri, P. B. and Cheresh, D. A. (1999). The role of αv integrins during angiogenesis: insights into potential mechanisms of action and clinical development. *Journal of Clinical Investigation*, **103**, 1227–1230

Elkin, M., *et al*. (2000). Halofuginone: a potent inhibitor of critical steps in angiogenesis progression. *FASEB Journal*, **14**, 2477–2485.

Giavazzi, R. and Taraboletti, G. (2001). Preclinical development of metalloproteasis inhibitors in cancer therapy. *Critical Reviews in Oncology and Hematology*, **37**, 53–60.

Hidalgo, M. and Eckhardt, S. G. (2001). Development of matrix metalloproteinase inhibitors in cancer therapy. *Journal of the National Cancer Institute*, **93**, 178–193.

Humphries, M. J., *et al*. (1986). A synthetic peptide from fibronectin inhibits experimental metastasis of murine melanoma cells. *Science*, **233**, 467–469.

Maeshima, Y., *et al*. (2000). Two RGD-independent $\alpha v\beta$ 3 integrin binding sites on tumstatin regulate distinct anti-tumor properties. *Journal of Biological Chemistry*, **275**, 23745–23750.

Mignatti, P. and Rifkin, D. B. (2000). Non-enzymatic interactions between proteinases and the cell surface: novel roles in normal and malignant cell physiology. *Advances in Cancer Research*, **78**, 103–157.

Nilsson, F., *et al*. (2001). Targeted delivery of tissue factor to the ED-B domain of fibronectin, a marker of angiogenesis, mediates the infarction of solid tumors in mice. *Cancer Research*, **61**, 711–716.

O'Reilly, M. S., *et al*. (1997). Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell*, **88**, 277–85.

Rüegg, C., *et al*. (1998). Evidence for the involvement of endothelial cell integrin v 3 in the disruption of the tumor vasculature induced by TNF and IFN-γ. *Nature Medicine*, **4**, 408–414.

Silletti, S., *et al*. (2001). Disruption of matrix metalloproteinase 2 binding to integrin alpha vbeta 3 by an organic molecule inhibits angiogenesis and tumor growth *in vivo*. *Proceedings of the National Academy of Sciences of the USA*, **98**, 119–124.

Yamada, K. M. (1991). Adhesive recognition sequences. *Journal of Biological Chemistry*, **266**, 12809–12812.

Yi, M. and Ruoslahti, E. (2001). A fibronectin fragment inhibits tumor growth, angiogenesis, and metastasis. *Proceedings of the National Academy of Sciences of the USA*, **98**, 620–624.

Vazquez, F., *et al*. (1999). METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity. *Journal of Biological Chemistry*, **274**, 23349–23357.

Vlodavsky, I., *et al*. (1999). Mammalian heparanase: gene cloning and function in tumor progression and metastasis. *Nature Medicine*, **5**, 793–802.

Zacharski, L. R., *et al*. (2000). Low-molecular-weight heparin and cancer. *Seminars in Thrombosis and Hemostasis*, **26**, 69–77.

FURTHER READING

Dejana, E. and Corada, M. (eds) (1999). *Adhesion Protein Protocols*. (Humana Press, Totowa, NJ).

Gunthert, U., *et al*. (eds) (1996). *Attempts to Understand Metastasis Formation III*. (Springer, Berlin).

Website

Drugs that block matrix breakdown.
http://cancertrials.nci.nih.gov/news/angio.

Chapter 95

Growth Factor Receptor Blockade

Marissa Shrader and Matthew H. Herynk
University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

Robert Radinsky
AMGEN, Inc., Thousand Oaks, CA, USA

CONTENTS

- Introduction
- Growth Factor Receptor Tyrosine Kinases
- Biological Properties of Epidermal Growth Factor Receptor (EGFR)
- The Role of RTKs and EGFR in Human Cancer
- Growth Factor Receptor Blockade Therapy
- Conclusions

INTRODUCTION

Cancer is the second leading cause of death in the United States today. Its development is a complex process in which malignant cells proliferate to form a tumour or a large progeny of transformed cells. This involves multiple genetic alterations resulting in phenotypic changes to the cell characteristic of the malignant process such as increased survival, proliferation and invasion. Although the individual steps in tumorigenesis are not completely understood, the fundamental process involves cells undergoing a series of successive genetic changes, each of which alters the cell genome in a way that confers a selective growth advantage to a specific cellular clone. These genetic changes lead to altered cellular signalling or to an imbalance of proteins that stimulate the cell cycle. The identification and characterization of oncogenes and tumour-suppressor genes have enhanced our understanding of the biochemical pathways associated with the tumour progression.

A major limitation of current cancer therapy is the ability of cancer cells to metastasize. Metastasis, the spread of cells from a primary neoplasm to a distant site, is not a random process. This process involves a number of highly regulated steps, which are influenced by both host factors and the intrinsic properties of the tumour cells (reviewed by Radinsky, 1995). The essential steps in the formation of a metastatic lesion include: (1) growth of neoplastic cells after the initial transforming event(s) that is first supported with nutrients supplied from the local microenvironment; (2) neovascularization must occur next for a tumour mass to exceed approximately $2\,mm^3$ in volume; the regulation, synthesis and secretion of tumour angiogenesis factors play a vital role in this process; (3) local tumour cell invasion of the surrounding host stroma occurs next via blood and lymphatic vessels and is enhanced by the production of lytic enzymes such as matrix metalloproteinases (MMPs) and urokinase from either tumour cells or host inflammatory cells; (4) once the tumour breaches the stroma of the circulatory system, detachment and embolization of small tumour cell aggregates occur, with the majority of tumour cells being rapidly destroyed; for most tumours, fewer than 0.1% of tumour cells that enter the circulation survive to form metastases owing to blood turbulence and the trauma associated with arrest, transcapillary passage and lysis by lymphocytes, monocytes and natural killer (NK) cells; (5) the tumour cells must arrest in the capillary beds of distant organs, either by adhering to capillary endothelial cells or by adhering to the exposed subendothelial basement membrane; (6) extravasation occurs next, by mechanisms similar to those that influence initial invasion; (7) tumour cell survival and proliferation within the organ parenchyma complete the metastatic cascade. To grow in the organ parenchyma, the metastases must develop a vascular network and evade the host immune system. For production of clinically relevant metastases, each of the outlined steps must be completed. Failure to complete one or more steps (e.g. inability to invade host stroma, a high degree of antigenicity, inability to grow in a distant organ's parenchyma) eliminates the cells.

Numerous examples exist in which malignant tumours metastasize to specific organs (Fidler, 1990). In 1889, Paget proposed that the growth of metastases is influenced by the interaction of particular tumour cells (the 'seed') with the unique organ's environment (the 'soil') and that metastases resulted only when the seed and soil were compatible (Paget, 1999). Common regional metastatic involvements have been attributed to anatomical or mechanical

considerations such as efferent venous circulation or lymphatic drainage to regional lymph nodes but distant organ metastases represent a unique pattern of organ specificity. Experimental and clinical confirmation of these observations advocates that the microenvironment of each organ influences the implantation, invasion, survival and growth of distinct tumour cells (Fidler, 1990; Radinsky, 1995). The mechanistic basis of this interaction remains under intense investigation. What influences a tumour cell's response to specific organ environments? What enables one tumour cell to survive and ultimately proliferate as a metastatic lesion and another to lie dormant in some cases for years? The properties of the tumour cell and host factors play a key role in this process. For example, during the metastatic cell's interactions with a number of host cells and systems, biochemical signals from endocrine, paracrine or autocrine pathways, alone and in combination, stimulate or inhibit tumour cell survival and proliferation, with the eventual outcome dependent on the net balance of positive and negative regulators. Hence the successful metastatic cell, referred to three decades ago as the 'decathlon champion' (Fidler, 1990), must also be viewed as a cell receptive to its environment (Radinsky, 1995).

Membrane-bound receptor tyrosine kinases (RTKs) and their corresponding growth factor ligands participate in this process in multiple solid tumour types and their metastases. This chapter focuses on current and novel therapeutic strategies targeting RTKs and/or their corresponding growth factor ligands in the therapy of primary solid tumours and their resulting metastases. Emphasis is given to strategies using monoclonal antibody and small molecule kinase inhibitor approaches targeting specific members of the ErbB family of RTKs including the epidermal growth factor (EGF) and the ErbB2/HER2 receptors. Use of the RTK targeting agents alone and in combination with cytotoxic compounds or radiation is discussed in the context of antitumour effects in terms of mechanisms affecting tumour cell survival, proliferation, invasion and angiogenesis in preclinical animal tumour models and in human clinical trials. (See also chapter on *Signalling by Tyrosine Kinases*.)

GROWTH FACTOR RECEPTOR TYROSINE KINASES

Cell surface molecules carry information from the microenvironment to the cell. Cells communicate by way of cell surface receptors that recognize and respond to signals from the environment. One major family of cell surface molecules is composed of transmembrane proteins with intrinsic tyrosine kinase activity. The superfamily of RTKs is organized into subfamilies depending on structural and sequence similarities (Blume-Jensen and Hunter, 2001). These include the epidermal growth factor receptor (EGFR) family, the platelet-derived growth factor receptor (PDGFR) family, the insulin receptor (IR) family, the fibroblast growth factor receptor (FGFR) family, the hepatocyte growth factor receptor (c-Met) family and the neurotrophin growth factor receptor family (**Table 1**). These RTKs contain several discrete domains, including an extracellular ligand-binding domain, a transmembrane domain, a tyrosine kinase catalytic domain and a C-terminal domain. Interaction of a growth factor with its receptor at the cell surface leads to a tight association enabling growth factors to mediate their activities at nanomolar concentrations.

Receptor tyrosine kinases have activity against tyrosine residues present both within the receptor itself and in downstream adaptor molecules. RTKs bind to specific ligands triggering a conformational change resulting in dimerization of the receptors. Receptor dimerization brings together the tyrosine kinase catalytic domains (intracellular domains), which phosphorylate each other on tyrosine residues, and activate intracellular signal transduction pathways (Blume-Jensen and Hunter, 2001). The mitogenic pathway is mediated through the growth factor receptor binding protein (Grb2) and activation of the Ras and the mitogen-activated protein kinase (MAPK) signalling pathways. Some growth factors, particularly insulin-like growth factor I (IGF-I), can promote cell survival by activating the phosphatidylinositide-3-OH kinase (PI3′K) and its downstream target, the serine–threonine kinase Akt, also known as PKB (protein kinase B), thus promoting cell survival through the activation of anti-apoptotic pathways.

BIOLOGICAL PROPERTIES OF EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR)

The epidermal growth factor receptor (EGFR; also known as HER, human epidermal growth factor receptor, and c-erbB1) is a representative member of the tyrosine kinase family of receptors. EGFR was the first receptor described to have tyrosine kinase activity and the first member of the receptor tyrosine kinase family to be sequenced (Mendelsohn, 1997; Wells, 1999). The EGFR responds to multiple factors from the EGF family of ligands. Each ligand can elicit a different activating or deactivating response that is dependent on ligand and cell type. The overactivation of EGFR can have dramatic effects on a cell's normal functions, leading to uncontrolled migration, survival and proliferation. *Egfr* gene knockout mice usually die in the blastocyst stage or *in utero*, but may live to 20 days post-birth depending on the genetic background of the mice (Wells, 1999). EGFR has a variety of ligands with redundant signalling pathways. EGF is well characterized and can be a potent activator of epithelial cells. Transforming growth factor α (TGF-α) is a potent

Table 1 Growth factor receptor families and their properties in malignancy

Receptor family	Family members	Action
Epidermal growth factor receptor (EGFR)	EGFR HER2/neu/c-erbB2 HER3/c-erbB3 c-erbB4 Xmrk	Epithelial cell growth
Hepatocyte growth factor receptor (c-Met)	c-Met	Mesenchymal cell growth Proangiogenic
	Sea	Plasma cell growth
	Ron	Enhancement of cell invasion and motility
Insulin growth factor receptor	Insulin receptor (I-R) Insulin-like growth factor 1 receptor (IGF-1R) Insulin receptor related protein (IRR) c-Ros	Proliferation Antiapoptosis
Platelet-derived growth factor receptor (PDGFR)	PDGFR-A	Mesenchymal/smooth muscle proliferation
	PDGFR-B Colony stimulating factor 1R (CSF-1R) c-Kit	Mast cell growth
Fibroblast growth factor receptor (FGF-R)	FGFR1/flg FGFR2/bek FGFR3 FGFR4 Keratinocyte growth factor receptor (KGF-R)	Proangiogenic Epithelial cell growth
Neurotrophin growth factor receptor (NGF-R)	NGF-R	Neurite growth and survival
	trkB	

For a review, see Blume-Jensen and Hunter (2001).

activator during liver regeneration and normal epithelial growth and is often expressed in an autocrine manner by different tumour cells (Radinsky, 1995). Mice deficient in TGF-α or EGFR do not die early in development, indicating redundancy in the EGFR/TGF-α signalling pathways that can compensate for the loss of either protein. Other members of the EGF family of ligands include amphiregulin (AR), heparin-binding EGF-related growth factor (HB-EGF) and betacellulin (BTC).

Following stimulation by ligand, the EGFR (ErbB1) is able to heterodimerize with other members of the EGF/c-erbB family of receptors, namely ErbB2 (Her2/neu), ErbB3 and ErbB4 (Blume-Jensen and Hunter, 2001). This heterodimerization allows EGFR to respond to another family of growth factor ligands, the neuregulins (Wells, 1999; Kirschbaum and Yarden, 2000). Activation of intracellular signalling pathways occurs when the RTK binds to one or a few intracellular proteins with Src homology 2 (SH2) domains via typical docking sites containing phosphotyrosine residues (Blume-Jensen and Hunter, 2001). These effector molecules include the adaptor proteins Grb2, Shc, Crk, Vav and Nck and proteins that regulate second messengers such as PI3-kinase, PI-4,5-kinase and PLC-γ (**Table 2**).

Table 2 EGFR intracellular binding partners

Name	Function	Reference
Grb2	Adaptor protein	Batzer *et al.* (1994)
Shc	Adaptor protein	
Crk	Adaptor protein	
Crk II	Adaptor protein	
Vav	Adaptor protein	Dougall *et al.* (1996)
Nck	Adaptor protein	
Gab-1	Adaptor protein	
GAP	ras GTPase	
β-Catenin	Cell–cell adhesion	Runge *et al.* (1999)
Actin	Cytoskeletal interaction	
STAT1α/1β	Transcription factor	
STAT 5	Transcription factor	
PTP 1B	Protein phosphatase	Liu and Chernoff (1997)
SHPTP2	Protein phosphatase	Dougall *et al.* (1996)
Jak2	Tyrosine kinase	
SRC	Tyrosine kinase	
PKC	Serine kinase	Dougall *et al.* (1996)
PKA	Serine–threonine kinase	Chen *et al.* (1996)
PI-3 kinase	Lipid kinase	
PLC-γ	Phospholipase	

THE ROLE OF RTKs AND EGFR IN HUMAN CANCER

The most common cellular lesions found in human cancers involve the overexpression of RTKs in combination with autocrine stimulation of the receptor by ligand. These include TGF-α, PDGF A and B, acidic fibroblast growth factor (aFGF), basic fibroblast growth factor (bFGF) and their specific receptors (**Tables 1** and **2**). Epidermal growth factor receptors are present on many normal and tumour cells (Wells, 1999; Ciardiello *et al*., 2001). Elevated levels and/or amplification of the EGFR has been found in many human tumours and cell lines, including breast cancer, gliomas, lung cancer, bladder cancer (Dinney *et al*., 1997; Perotte *et al*., 1999; Inoue *et al*., 2000), tumours of the female genital tract, the A431 epidermoid carcinoma, prostate carcinoma, pancreatic carcinoma and colon carcinoma (Radinsky, 1995). These results suggest the physiological significance of inappropriate expression of members of the RTK family including EGFR in abnormal cell growth control, making it and other receptor tyrosine kinases prime targets for anticancer therapy.

EGFR Signal Transduction and Bladder Cancer Progression

Transitional cell carcinoma (TCC) of the bladder is the fourth most common malignancy in the United States. Clinical studies evaluating the significance of EGFR expression in human TCC have shown that >50% of human TCCs overexpress EGFR, and that the level of expression directly correlates with tumour grade, stage and survival (Dinney *et al*., 1997; Perrotte *et al*., 1999; Inoue *et al*., 2000; Brons *et al*., 2000). In patients with superficial bladder cancer, EGFR expression correlates with multiplicity, time to disease recurrence and overall recurrence rates. EGFR expression also predicted disease progression to muscle invasive or metastatic TCC and was an independent prognostic factor for death in a multivariate analysis. EGFR expression also has prognostic significance for patients with advanced TCC. Patients with muscle-invasive TCC which overexpress EGFR have a 20% probability of long term cancer-specific survival which is significantly worse than the survival of those whose tumours did not express EGFR. These studies establish the importance of EGFR overexpression in the development and progression of human TCC of the bladder.

EGFR Signal Transduction and Pancreatic Carcinoma Progression

The EGFR and its ligands (including EGF, AR and TGF-α) are also commonly overexpressed in pancreatic cancer and their expression is associated with a worse prognosis (Evans *et al*., 1997). This coexpression of both receptor and ligand and the fact that EGF, AR and TGF-α are mitogens for pancreatic cancer cell lines have led to the hypothesis that an EGFR-dependent loop contributes to the malignant phenotype of this disease. Its importance in the biology of pancreatic cancer is supported by experiments targeting either EGFR or its ligands (see below) (Bruns *et al*., 2000a,b). Immunohistochemical analysis of 87 pancreatic tumours for EGFR, TGF-α and EGF found that coexpression of these proteins correlated with a more rapidly progressive pancreatic cancer in comparison with tumours that failed to express the receptor and its ligands. Similarly, a statistically significant decrease in the survival of patients overexpressing the EGFR and at least one of its ligands was observed. No correlation between survival and the presence of Ki-*ras* or the *p53* tumour-suppressor gene mutations was observed. In addition, a longer median survival in unresectable pancreatic cancer patients treated with the anti-EGFR Mab 425 was reported and one complete response was observed in a patient who remained in an unmaintained remission for 3 years (Evans *et al*., 1997). These studies establish the importance of EGFR overexpression in the development and progression of human pancreatic carcinoma and collectively implicate the EGFR as a potential target in the therapy of carcinomas from multiple origins.

GROWTH FACTOR RECEPTOR BLOCKADE THERAPY

There are several methods of RTK blockade therapy designed to interrupt signal transduction selectively. These therapeutic approaches targeting RTKs include strategies using monoclonal antibodies (MAbs) targeted to the extracellular domain, tyrosine kinase inhibitors that are preferential for individual RTKs, bispecific and single-chain antibodies, overexpression of dominant negative mutant receptors, ligands conjugated to toxins, inhibitors of receptor dimerization or antisense RNA. Each of these approaches will be outlined below in the context of specific therapeutic effects alone and in combination with cytotoxic drug therapy or radiation in preclinical experiments and in some cases clinical trials.

Monoclonal Antibody Therapy of Human Cancers

Directing monoclonal antibodies (MAbs) towards a specific RTK represents a promising anticancer therapeutic strategy. By selectively targeting the extracellular portion of a receptor, ligand binding and subsequent intracellular signalling can be prevented. Also, some antibodies subsequent to receptor binding evoke a host immune response

to the tumour. Several MAbs are currently in clinical trials or in use in the clinic, including Rituximab (anti-CD20), Trastuzumab (anti-ErbB2/Her2) and C225 (anti-EGFR). In 1997, the US Food and Drug Administration (FDA) approved the first MAb for the treatment of human cancer. This was anti-CD20 mouse/human chimaeric antibody, Rituximab/Rituxan^{TM}, used to treat human B cell lymphomas. Clinical trials using this MAb showed a 15–60% objective response rate when administered as a signal agent and a 95–100% response rate when combined with chemotherapy (Coiffer *et al.*, 1998).

In 1998, the FDA approved the second MAb for the treatment of human cancer. Trastuzumab, commonly known as Herceptin^{TM}, is a chimaeric mouse/human anti-ErbB2/HER2/neu receptor monoclonal antibody. This was the first antibody approved to treat solid tumours. ErbB2/HER-2/neu regulates normal epithelial cell growth. Amplification and/or overexpression result in increased receptor activity and increased tumour growth. Clinically, ErbB2/HER-2 overexpression is associated with a more rapid rate of tumour growth, increased rates of metastasis and decreased patient survival. However, overexpression of ErbB2/HER-2 is seen in approximately 20–30% of breast cancer patients and a small number of patients with other solid tumours such as lung, bladder, ovary and pancreas (Slamon *et al.*, 1989).

Binding of the Herceptin MAb antibody to the HER-2 receptor results in competitive inhibition of ligand binding to the receptor. *In vitro*, anti-ErbB2/HER-2 antibodies have been shown to downregulate the kinase activity of the receptor, decrease cell proliferation and increase chemosensitivity in HER-2 overexpressing cells (Benz and Tripathy, 2000). As a single agent in phase II clinical trials, 17% of patients with metastatic breast cancer achieved a 50% or greater reduction in tumour size. In phase III trials when Herceptin was given with paclitaxel or doxorubicin/cyclophosphamide, objective response rates increased to 44 and 53%, respectively (Burris, 2000; Benz and Tripathy, 2000).

Monoclonal antibodies have also been developed to the EGFR. An example is anti-EGFR humanized MAb C225, which is currently in multiple clinical trials (Mendelsohn, 1997). C225 binds to the extracellular domain of EGFR, blocking ligand binding and resulting in blockade of receptor activation (and in some cases increased internalization of the nonactivated receptor) (**Figure 1; see colour plate section**), decreased proliferation and cell-cycle arrest. C225 also exhibits additive to synergistic effects when given in combination with chemotherapy or radiation versus either agent alone. Recent preclinical studies in our laboratory have shown that the treatment of human bladder and pancreatic cancer with C225 results in significantly decreased primary tumour growth and incidence of lung, liver and lymph node metastases (Perotte *et al.*, 1999; Bruns *et al.*, 2000; Inoue *et al.*, 2000).

Anti-EGFR Therapy of Human Pancreatic Cancer in Preclinical Models

Cancer of the exocrine pancreas is characterized by extensive local invasion and early development of metastasis (Evans *et al.*, 1997). At the time of diagnosis, more than 80% of the patients present with either locally advanced or metastatic disease. The inability to detect pancreatic cancer at an early stage, its aggressiveness and the lack of effective systemic therapy are responsible for rapid death from this disease. In fact, only 4% of all patients with adenocarcinoma of the pancreas will survive 5 years after diagnosis. For patients with advanced pancreatic cancer, even the recent introduction of the deoxycytidine analogue gemcitabine does not extend median survival beyond 6–9 months. Therefore, new therapeutic modalities such as RTK blockade therapy are needed for this disease.

As described above, an association exists between pancreatic tumour progression and the expression and function of the EGFR. The pancreas is one of the richest sources for expression of prepro-EGF, thereby raising the question of autocrine growth stimulation of pancreatic cancer by EGF production (Bruns *et al.*, 2000a,b). EGF has been shown to promote pancreatic carcinogenesis in hamsters and the *in vitro* growth of human pancreatic cancer cells (Evans *et al.*, 1997). In addition to EGF, several cultured human pancreatic cancer cell lines produce TGF-α. TGF-α has been shown to be 10–100 times more potent than EGF in stimulating anchorage independent growth of these cell lines; 3-, 15- and 10-fold increases in the mRNA levels of EGFR, EGF and TGF-α, respectively, have been reported in pancreatic carcinomas as compared with normal pancreatic tissue. These data suggest that overexpression of the EGFR and its ligands may contribute to the malignant phenotype in human pancreatic cancer (Yamanaka *et al.*, 1993).

We recently reported that therapy with the anti-EGFR humanized MAb C225 inhibited the growth of established human pancreatic carcinomas growing in the pancreas of athymic nude mice (Bruns *et al.*, 2000a) (**Table 3**). These data show that targeting the EGFR in established disease with anti-EGFR MAbs alone or in combination with gemcitabine results in significant antitumour effects (**Table 3**). The abrogation of tumour growth and distant lymph node and liver metastases was due in part to blockade of EGFR tyrosine kinase activation (**Figure 1**) and reduction of tumour-induced neovascularization (**Figure 2; see colour plate section**) secondary to the downregulation of tumour cell expression of the angiogenic factors VEGF and IL-8 (Parker *et al.*, 1998; Dinney *et al.*, 1997) (**Figure 3; see colour plate section**). These studies confirm that systemic administration of the chimaeric anti-EGFR MAb C225 inhibits growth and metastasis of human pancreatic cancer established in the pancreas of athymic nude mice. Therapy with MAb C225 has a significant antitumour effect

Table 3 Therapy of human pancreatic tumours in nude mice with anti-EGFR MAb C225 and gemcitabine. (Adapted from Bruns *et al.*, 2000a.)

Therapy[a]	Incidence of macroscopic tumours[b]			Total pancreas weight (mg) (range)[e]	Median tumour volume (mm^3) (range)[f]	Median body weight (g) (range)
	Pancreas tumour	Liver metastasis[c]	Regional LN metastasis[d]			
Control	10/10	5/10	10/10	923 (660–1371)	539 (254–860)	22 (16–24)
Gemcitabine	10/10	3/10	6/10	297 (205–485)	152 (59–365)	22 (18–29)
C225	5/10	2/10	8/10	119 (97–157)	0.3 (0–13)	25 (21–28)
C225 + gemcitabine	0/9	0/9	1/9	130 (93–173)	0 (0)	23 (20–27)

[a] 1×10^6 L3.6pl cells were implanted into the pancreas. Therapy began on day 7 when the median tumour size was 18 mm^3. Anti-EGFR MAb C225 (1 mg); gemcitabine (250 mg/kg) alone or in combination was injected i.p. biweekly for 4 weeks. Control mice received saline. Tumour volume, weight and extent of metastases were evaluated at necropsy on day 32

[b] Number of tumour-positive mice per number of injected mice

[c] Number of mice with visible nodules (>1 mm in diameter) per number of injected mice

[d] Number of mice with enlarged regional lymph nodes per number of injected mice

[e] $p < 0.0001$ (unpaired Student's *t*-test) control versus all three therapy groups

[f] Tumour volume was calculated using the equation $v = ab^2/2$, where a is the longest diameter and b is the shortest diameter of the tumour. $p < 0.0001$ control versus all three therapy groups and anti-EGFR MAb C225 versus anti-EGFR MAb C225 plus gemcitabine therapy group.

mediated by direct effects on the tumour cells with respect to inhibition of proliferation, and by inhibition of angiogenesis mediated by downregulation of tumour cell produced angiogenic factors, which are potentiated when used in combination with gemcitabine.

EGFR Signalling and Regulation of Angiogenesis

The downregulation of angiogenic factors produced by the tumour restores the balance between stimulating and inhibitory factors that keeps angiogenesis dormant under normal conditions. The observation that downregulation of the angiogenic stimulus of the tumour cells inhibits the host angiogenic response emphasizes the complexity of tumour–host interactions. The experiments outlined above demonstrate that inhibition of angiogenesis characterizes, in part, the antitumour effect of therapy directed at inhibiting EGFR signalling pathways in human pancreatic tumour cells. This observation was first reported by Petit *et al.* (1997) and later confirmed by several groups, including our own in multiple tumour systems (Perrotte *et al.*, 1999; Bruns *et al.*, 2000a,b; Ciardiello *et al.*, 2001). Both EGF and TGF-α, which are ligands for EGFR, induce angiogenesis. Therefore, downregulation of EGFR signalling pathways could inhibit tumour growth by inhibiting tumour-mediated angiogenesis, independent of any direct cytostatic effect on tumour growth. Preclinical data suggest that EGFR blockade therapies alone and in combination with cytotoxic therapies (e.g. gemcitabine, paclitaxol) inhibit the growth of established human carcinomas growing orthotopically in immunodeficient mice (**Table 3**). The abrogation of tumour growth and distant metastasis was due, in part, to a reduction in tumour-induced neovascularization secondary to the downregulation of tumour cell expression of the angiogenic factors vascular endothelial growth factor (VEGF), bFGF and interleukin-8 (IL-8), leading to endothelial and tumour cell apoptosis and ultimately to regression of established tumours (**Figures 2 and 3**). Thus, the presence of the proangiogenic factors VEGF, bFGF and IL-8 were necessary to maintain a viable microcirculation and maintain tumour growth. The withdrawal of the angiogenic stimulus by EGFR blockade therapy resulted in regression of the tumour's neovasculature and ultimately tumour cell death.

The mechanisms by which EGFR signalling pathways regulate VEGF and IL-8 are unclear, but it is established that upregulation of these factors follows activation of the EGFR signalling pathways by EGF or TGF-α. Transcription of VEGF is potentiated by activation of the four AP-1 transcription factor binding sites within its promoter; the IL-8 promoter has one AP-1 site (Petit *et al.*, 1997). After activation of EGFR signalling pathways, *ras* and *raf* are activated, resulting in phosphorylation of c-Fos and c-Jun, leading to increased AP-1 activity (Petit *et al.*, 1997). This increase in AP-1 transcription factor activity leads to transcription of genes with AP-1 sites in their promoter. Since VEGF and IL-8 all share AP-1 binding sites, they are potential targets for therapies that downregulate EGFR signalling pathways and reduce AP-1 activity. EGFR activation has also been shown to stimulate VEGF expression in multiple tumour cells (Ciardiello *et al.*, 2001). Petit *et al.* (1997) reported that *in vitro* treatment of the human epidermoid carcinoma cell line A431 with MAb C225 downregulated VEGF and that after *in vivo* therapy tumours showed a reduction in microvessel density counts (Petit *et al.*, 1997). Similar observations have now been reported in multiple tumour models (Ciardiello *et al.*, 2001).

Alternative Mechanisms for the Tumoricidal Effect of EGFR Blockade

A cooperative growth inhibitory and cytotoxic effect of C225 in combination with several chemotherapeutic agents and radiation has recently been shown to augment antitumour activity in several mouse xenograft models and in human clinical trails (Wells, 1999). Anti-EGFR MAb C225 substantially enhanced the cytotoxic effects of doxorubicin, *cis*-diammine-dichloroplatinum, gemcitabine and paclitaxol on well-established xenografts (Baselga *et al.*, 1993; Mendelsohn, 1997). Furthermore, clinical trials with squamous cell carcinoma of the lung have demonstrated the capacity of the anti-EGFR MAb C225 to localize in such tumours and to achieve saturating concentrations in the blood for >3 days without toxicity. The molecular pathways for this effect are unclear, but those affecting DNA repair, multidrug resistance, cell cycle checkpoint control or as discussed above angiogenesis, may be involved. For example, treatment with C225, but not EGF, triggered a specific physical interaction between the internalized EGFR and DNA-dependent protein kinase (DNA-PK) implicated in the repair of DNA double-strand breaks (Mendelsohn, 1997). This significantly reduced the level and activity of DNA-PK in the nucleus with a concomitant increase in DNA-PK levels in the cytosol, suggesting that EGFR blockade and downregulation by C225 may impair DNA repair by reducing the nuclear level of DNA-PK. The capacity of C225 to modulate tumour cell cycle distribution may also play a central role regulating the increased sensitivity to chemotherapeutic agents (e.g. gemcitabine, paclitaxol) and radiation. This may involve cell cycle checkpoint control as an activator of cell death (Wells, 1999). EGFR blockade results in cellular arrest at the G_1 restriction point, and cells damaged by chemotherapy or radiation typically arrest in G_2-M to repair DNA alterations. Mendelsohn (1997) hypothesized that when tumour cells simultaneously ignore two checkpoint signals (e.g. activated by EGFR blockade and cytotoxic drug or radiation treatment), cell death occurs. Nonmalignant epithelial cells, which obey checkpoint control signals, may be less susceptible to the cytotoxic effects of these combination treatments. These data indicate that inhibition of EGFR signalling probably acts through several potential mechanisms to sensitize tumour cells to cytotoxic agents or radiation and provides possible mechanisms for enhanced effects of therapy of human carcinomas with EGFR blockade agents in combination with cytotoxic agents or radiation.

Receptor Tyrosine Kinase Inhibitors

RTK-specific inhibitors are small molecules that directly inhibit RTK activity and hence phosphorylation by physically interacting with the kinase domain and blocking ATP binding or inhibiting tyrosine phosphorylation. Newer generation compounds compete at the ATP-binding site and prevent phosphorylation by blocking the phosphate source, ATP. These compounds are also very receptor specific. Most RTK inhibitors are targeted against the EGFR family (Noonberg and Benz, 2000).

In an attempt to block selectively EGFR signalling, Buchdunger *et al.* (1995) synthesized a new class of dianilinophthalimide (DAPH) tyrosine kinase inhibitors. DAPH2,4,5-bis(4-fluoroanilino)phthalimide, is a potent and selective inhibitor of the EGFR tyrosine kinase and the PKC $\beta 2$ enzyme. *In vivo*, DAPH2 significantly reduced the growth, in a dose-dependent manner, of human A431 and 253J B-V bladder carcinoma xenografts growing in immunodeficient mice (Dinney *et al.*, 1997; Parker *et al.*, 1998).

Bos *et al.* (1997) found that PD153035, an RTK inhibitor, suppresses the growth of a number of EGFR-overexpressing human cancer cell lines. Administration of the RTK inhibitor in combination with MAb C225 further increased the antitumour activity of PD153035. The effects of ethyl 2,5-dihydroxycinnamate (EtDHC), a novel EGFR TK inhibitor, were tested on human glioblastoma cell lines expressing different EGFR mutants. They found that EtDHC was more potent in inhibiting EGFR tyrosine kinase activity in cell lines overexpressing the wild-type *EGFR* gene, resulting in decreased cell growth and DNA synthesis. In cell lines expressing truncated EGFR, a higher dose of EtDHC was required to achieve similar effects. This finding was in contrast to work with another EGFR TK inhibitor, tyrophostin AG 1478, that preferentially inhibited glioblastoma cells expressing truncated EGFR (Han *et al.*, 1996).

A new class of irreversible RTK inhibitors were recently engineered against EGFR and ErbB2/HER-2 (Fry, 2000). The inhibitor PD168393 inactivates the EGFR tyrosine kinase by covalently modifying a cysteine residue in the ATP binding pocket. It was also found that the irreversible inhibitor was nine times more effective than a similar reversible compound (PD174265) in a human A431 epidermoid carcinoma xenograft model. An even greater difference (30-fold) was seen against HER-2 phosphorylation in MDA-MB-453 human breast carcinoma cells.

The efficacy of RTK inhibitors has recently been described. Faust *et al.* (1999) found that PD153035, an RTK inhibitor against the EGFR TK, caused apoptosis and growth inhibition *in vitro* of human squamous cell carcinoma of the head and neck (SCCHN). Murakami *et al.* (1999) described the activity of Sporostatin, a novel and specific inhibitor of the EGFR tyrosine kinase. Sporostatin was isolated from a fungus of *Sporomiella* sp. M5032 and is noncompetitive with either substrate or ATP. They also reported that sporostatin inhibits the autophosphorylation of the EGFR in A431 cells. More recent reports describe similar results with RTK inhibitors of the EGFR and ErbB2/HER-2 receptor. To date, multiple EGFR-preferential RTK inhibitors from a number of companies including AstraZeneca, OSI, Novartis and

Pfizer are in human clinical trials (Cassinelli *et al*., 2000); Fry, 2000; Kirschbaum *et al*., 2000 and Vincent *et al*., 2000).

Bruns *et al*. (2000b) recently published a study of a novel EGFR RTK inhibitor, PKI166, from Novartis Pharma. They found that administration of PKI166 significantly inhibited the growth and liver and lymph node metastases of human pancreatic cells growing in the pancreas of nude mice. As a single agent, PKI166 reduced the growth of pancreatic tumours by 45%. Primary tumour growth was further reduced when PKI166 was combined with gemcitabine (85% versus control). Combination therapy also resulted in a decreased incidence of liver and lymph node metastases and prolonged survival. These results are similar to those seen with anti-EGFR MAb C225 administered in combination with gemcitabine (see above). The observed decrease in tumour growth was due, in part, to a decrease in tumour microvessel density that was attributed to increased endothelial cell death following therapy with PKI166 alone and in combination with gemcitabine.

Bispecific and Single-chain Antibodies

Bispecific antibodies (BsAb) are recombinant proteins that contain one bivalent variable region (Fv) arm targeting an epitope on the surface of a tumour cell and another targeting an immune effector cell, such as a T lymphocyte. BsAb recruit immune effector cells to the tumour target and generate an increased host–tumour interaction, leading to tumour-cell death.

Negri *et al*. (1995) tested the antitumour effect of the anti-EGFR/anti-CD3 BsAb M26.1 with the constant region removed. They found that the BsAb M26.1 targeted human T lymphocytes to the EGFR in an ovarian cancer xenograft immunotherapy model (IGROV1) and increased the survival of the treated animals. Valone *et al*. (1995) used BsAb MDX-210 in a phase Ia/Ib clinical trial in patients with advanced breast or ovarian cancer that overexpresses ErbB2/HER-2/neu receptor. MDX-210 binds to type I Fc receptors for immunoglobulin G (IgG) (Fc gammaRI) and to the HER-2/neu receptor. MDX-210 targets Fc gammaRI-positive immune effector cells such as monocytes and macrophages to tumour cells that overexpress HER-2. In the phase I clinical trial, one partial and one mixed tumour response were observed among 10 assessable patients receiving MDX-210. Goldstein *et al*. (1997) developed a similar antibody that targets the EGFR and Fc gammaRI-positive immune cells. This BsAb, H22-EGF, inhibited the growth of EGFR-overexpressing cells and mediated the cytotoxicity of these cells in the presence of Fc gammaRI-positive cells. Similar results have been seen with single-chain antibody therapy.

Single-chain antibodies (scFv) are recombinant proteins composed only of immunoglobulin heavy- and light-chain variable regions linked by a short, flexible polypeptide. Single-chain antibodies have rapid tumour penetration *in vivo* owing to their small size. These antibodies can be exogenously administered or generated within a tumour cell. ScFv are able to inhibit ligand binding and kinase activity or prevent receptor maturation and cell surface expression. The development of scFv is still in an immature stage as problems with delivery, stability and gene transfer are considered.

Beerli *et al*. (1994) developed a secreted form of an scFv that competes with EGF binding to the EGFR. This antibody was expressed in EGFR-transformed cells and was shown to inhibit receptor activation in an autocrine manner. This group also developed scFv against the extracellular domain of the ErbB2/HER-2 receptor using a different approach from above. The single-chain antibodies were expressed intracellularly and targeted to the lumen of the endoplasmic reticulum (ER). By binding the extracellular domain of the newly synthesized ErbB2/HER-2 receptor, the scFvs prevented the transit of the receptors through the ER to the cell surface, thus inactivating the receptors. Similar approaches have been utilized to inactivate HER-2 in T47D human mammary carcinoma cells. EGF and neu differentiation factor (NDF) induced the activation of mitogen-activated protein kinase (MAPK) and growth stimulation by NDF was inhibited in cells devoid of cell surface ErbB2/HER-2 receptors.

Dominant-negative Mutant Receptors

Previous work has provided evidence that structurally defective receptors can suppress the action of wild-type receptors, thus behaving as dominant-negative mutations. This causes various decreased receptor responses resulting in decreased cell proliferation, cell migration, etc. Wells (1999) found that the kinase-negative mutant EGFR K721A was internalized following ligand binding but was then recycled to the cell surface. Wild-type receptors, however, were degraded following internalization. Based on this information they determined that intracellular trafficking of the EGFR was determined by a sorting mechanism that recognizes EGFR molecules based on their kinase activity. Therefore, the mutant receptors were causing the suppression of the wild-type receptors. In culture, EGF stimulated the dimerization of both types of receptors in cultured NIH/3T3 cells that coexpressed human wild-type EGFR and a mutant EGFR that was lacking most of its cytoplasmic domain. However, only homodimers of wild-type receptors were activated by tyrosine autophosphorylation following ligand binding. Heterodimerization with mutant receptors resulted in the inactivation of the wild-type receptors, resulting in decreased kinase activity (Kashles *et al*., 1991).

More recent studies have shown that NIH/3T3 cells transformed with the activated *HER-2* oncogene could

be phenotypically reverted by a dominant-negative mutant HER-2 receptor (Messerle *et al*., 1994). These variants were mutated in the transmembrane region and the ATP-binding site of the kinase domain. The authors also found a significant decrease of mutant growth in soft agar (80–90%) compared with the transformed cells. Qian *et al*. (1994) showed that a kinase-deficient HER-2 could suppress the function of wild-type EGFR in transformed mouse fibroblast cells. EGFR normally heterodimerizes with HER-2 following ligand activation and becomes cross-phosphorylated. However, heterodimers formed by HER-2 mutants and wild-type EGFR resulted in a lack of kinase activity. The kinase-deficient ErbB2/HER-2 receptor protein behaves as an anti-oncogene factor and causes reversion of the transformed phenotype.

CONCLUSIONS

A primary goal of cancer research is an increased understanding of the molecular mechanisms mediating the process of tumour progression and metastasis. Experimental analysis of the interaction between cancer cells and the microenvironment has increased our understanding of the biological mechanisms mediating tumour formation and organ-specific metastasis. Insight into the molecular mechanisms regulating these processes has produced a foundation for new therapeutic approaches. As reviewed here, RTKs and their corresponding ligands act to influence tumour cell growth, differentiation, invasion, metastasis and angiogenesis. Taken together, the abundance of growth-promoting factors, the disturbance of growth-inhibitory pathways, tumour-induced angiogenesis and the presence of gene mutations combine to give cancer cells a distinct growth advantage which clinically contributes to rapid tumour progression, metastasis and poor survival. The analyses presented here of new and novel RTK blockade therapies add new evidence to support the concept that tumour progression and metastasis are not a random process; they constitute a regulated process that can be analysed at the molecular level in the context of the relevant organ environment. This new knowledge will lead to the further design and implementation of more effective RTK blockade therapies alone and in combination with cytotoxic compounds or radiation for this dreaded disease.

REFERENCES

Baselga, J., *et al*. (1993). Antitumor effects of doxorubicin in combination with anti-epidermal growth factor receptor monoclonal antibodies. *Journal of the National Cancer Institute*, **85**, 1327–1333.

Batzer, A. G., *et al*. (1994). Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor. *Molecular Cell Biology*, **14**, 5192–5201.

Beerli, R. R., *et al*. (1994). Autocrine inhibition of the epidermal growth factor receptor by intracellular expression of a single-chain antibody. *Biochemical and Biophysical Research Communications*, **204**, 666–672.

Benz, C. and Tripathy, D. (2000). ErbB2 overexpression in breast cancer: biology and clinical translation. *Journal of Women's Cancer*, **2**, 33–40.

Blume-Jensen, P. and Hunter, T. (2001). Oncogenic kinase signaling. *Nature*, **411**, 355–365.

Bos, M., *et al*. (1997). PD153035, a tyrosine kinase inhibitor, prevents epidermal growth factor receptor activation and inhibits growth of cancer cells in a receptor-dependent number. *Clinical Cancer Research*, **3**, 2099–2106.

Bruns, C. J., *et al*. (2000a). Epidermal growth factor receptor blockade with C225 plus gemcitabine results in regression of human pancreatic carcinoma growing orthotopically in nude mice by antiangiogenic mechanisms. *Clinical Cancer Research*, **6**, 1936–1948.

Bruns, C. J., *et al*. (2000). Blockade of the epidermal growth factor receptor signaling by a novel tyrosine kinase inhibitor leads to apoptosis of endothelial cells and therapy of human pancreatic carcinoma. *Cancer Research*, **60**, 2926–2935.

Buchdunger, E. *et al*. (1995). 4,5-Bis(4-fluoroanilino)phthalimide: a selective inhibitor of the epidermal growth factor receptor signal transduction pathway with potent *in vivo* antitumor activity. *Clinical Cancer Research*, **1**, 813–821.

Burris, H. A., III (2000). Docetaxel (taxotere) in Her-2-positive patients and in combination with trastuzmab (Herceptin). *Seminars in Oncology*, **27**, 19–23.

Cassinelli, G., *et al*. (2000). Clavilactones, a novel class of tyrosine kinase inhibitors of fungal origin. *Biochemical Pharmacology*, **59**, 1539–1547.

Chen, P., *et al*. (1996). Mitogenic signaling from the EGF receptor is attenuated by a phospholipase C-gamma/protein kinase C feedback mechanism. *Molecular Biology of the Cell*, **7**, 871–881.

Ciardiello, F., *et al*. (2001). Inhibition of growth factor production and angiogenesis in human cancer cells by ZD1839 (Iressa), a selective epidermal growth factor receptor tyrosine kinase inhibitor. *Clinical Cancer Research*, **7**, 1459–1465.

Coiffer, B., *et al*. (1998). Rituximab (anti-CD20 monoclonal antibody) for the treatment of patients with relapsing or refractory aggressive lymphoma: a multicenter phase II study. *Blood*, **92**, 1927–1932.

Dinney, C. P. N., *et al*. (1997). Therapy of human transitional cell carcinoma of the bladder by oral administration of the epidermal growth factor-receptor protein tyrosine kinase inhibitor, 4,5-dianilinophthalimide. *Clinical Cancer Research*, **3**, 161–168.

Dougall, W. C., *et al*. (1996). Association of signaling proteins with a nonmitogenic heterodimeric complex composed of epidermal growth factor receptor and kinase-inactive p185c-neu. *DNA Cell Biology*, **15**, 31–40.

Evans, D. B., *et al*. (1997). Cancer of the pancreas. In: de Vita, V. T. *et al*. (eds), *Cancer: Principles and Practice of Oncology*, 5th edn. 1054–1087 (J. B. Lippincott, Philadelphia).

Faust, R. A., *et al*. (1999). Apoptosis and growth inhibition of head and neck tumor cell line induced by epidermal growth factor receptor tyrosine kinase inhibitor. *Oral Oncology*, **35**, 290–295.

Fidler, I. J. (1990). Special Lecture: critical factors in the biology of human cancer metastasis: Twenty-eighth G. H. A. Clowes Memorial Award Lecture. *Cancer Research*, **50**, 6130–6138.

Fry, D. W. (2000). Site-directed irreversible inhibitors of the erbB family of receptor tyrosine kinases as novel chemotherapeutic agents for cancer. *Anti-Cancer Drug Design*, **15**, 3–16.

Goldstein, J., *et al*. (1997). Cytolytic and cytostatic properties of an anti-human Fc gammaRI (CD64) × epidermal growth factor bispecific fusion protein. *Journal of Immunology*, **158**, 872–879.

Han, Y., *et al*. (1997). Preferential inhibition of glioblastoma cells with wild-type epidermal growth factor receptors by a novel tyrosine kinase inhibitor ethyl-2,5-dihydroxycinnamate. *Oncology Research*, **9**, 581–587.

Inoue, K., *et al*. (2000). Paclitaxel enhances the effects of the anti-epidermal growth factor receptor monoclonal antibody ImClone C225 in mice with metastatic human bladder transitional cell carcinoma. *Clinical Cancer Research*, **6**, 4874–4884.

Kashles, O., *et al*. (1991). A dominant negative mutation suppresses the function of normal epidermal growth factor receptors by heterodimerization. *Molecular Cell Biology*, **11**, 1454–1463.

Kirschbaum, M. H. and Yarden, Y. (2000). The ErbB/HER family to tyrosine kinases: a potential target for chemoprevention of epithelial neoplasms. *Journal of Cell Biochemistry*, **34**, 52–60.

Liu, F. and Chernoff, J. (1997). Protein tyrosine phosphatase 1B interacts with and is tyrosine phosphorylated by the epidermal growth factor receptor. *Biochemical Journal*, **327**, 139–145.

Mendelsohn, J. (1997). Epidermal growth factor receptor inhibition by a monoclonal antibody as anticancer therapy. *Clinical Cancer Research*, **3**, 2703–2707.

Messerle, K., *et al*. (1994). NIH/3T3 cells transformed with the activated erbB-2 oncogene can be phenotypically reverted by a kinase deficient, dominant negative erbB-2 variant. *Molecular and Cell Endocrinology*, **105**, 1–10.

Murakami, Y., *et al*. (1999). Sporostatin, a novel and specific inhibitor of EGF receptor kinase. *Anticancer Research*, **19**, 4145–4150.

Negri, D. R., *et al*. (1995). *In vitro* and *in vivo* stability and anti-tumour efficacy of an anti-EGFR/anti-CD3 F(ab′)2 bispecific monoclonal antibody. *British Journal of Cancer*, **72**, 928–933.

Noonberg, S. B. and Benz, C. C. (2000). Tyrosine kinase inhibitors targeted to the epidermal growth factor receptor subfamily. *Drugs*, **59**, 753–767.

Paget, S. (1889). The distribution of secondary growths in cancer of the breast. *Lancet*, **1**, 571–573.

Parker, C., *et al*. (1998). Preferential activation of epidermal growth factor receptor in human colon carcinoma liver metastases growing in nude mice. *Journal of Histochemistry and Cytochemistry*, **46**, 595–602.

Perrotte, P., *et al*. (1999). Anti-epidermal growth factor receptor antibody C225 inhibits angiogenesis in human transitional cell carcinoma growing orthotopically in nude mice. *Clinical Cancer Research*, **5**, 257–265.

Petit, A. M. V., *et al*. (1997). *American Journal of Pathology*, **151**, 1523–1530.

Qian, X., *et al*. (1994). Kinase-deficient neu proteins suppress epidermal growth factor receptor function and abolish cell transformation. *Oncogene*, **9**, 1507–1514.

Radinsky, R. (1995). Modulation of tumor cell gene expression and phenotype by the organ-specific metastatic environment. *Cancer and Metastasis Reviews*, **14**, 323–338.

Runge, D. M., *et al*. (1999). STAT 1alpha/1beta, STAT 3 and STAT 5: expression and association with c-MET and EGF-receptor in long-term cultures of human hepatocytes. *Biochemical and Biophysical Research Communications*, **265**, 376–381.

Slamon, D. J., *et al*. (1989). Studies of HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science*, **244**, 707–712.

Valone, F. H., *et al*. (1995). Phase Ia/Ib trial of bispecific antibody MDX-210 in patients with advanced breast or ovarian cancer that overexpress the proto-oncogene HER-2/neu. *Journal of Clinical Oncology*, **13**, 2281–2292.

Vincent, P. W., *et al*. (2000). Anticancer efficacy of the irreversible EGFR tyrosine kinase inhibitor PD 0169414 against human tumor xenografts. *Cancer Chemotherapy and Pharmacology*, **45**, 231–238.

Wells, A. (1999). Molecules in focus: EGF receptor. *International Journal of Biochemistry and Cell Biology*, **31**, 637–643.

Yamanaka, Y., *et al*. (1993). Coexpression of epidermal growth factor receptor and ligands in human pancreatic cancer is associated with enhanced tumor aggressiveness. *Anticancer Research*, **13**, 565–570.

FURTHER READING

Baselga, J. (2000). New technologies in epidermal growth factor receptor-targeted cancer therapy. *SIGNAL*, **1**, 12–21.

Blume-Jensen, P. and Hunter, T. (2001). Oncogenic kinase signaling. *Nature*, **411**, 355–365.

Ciardiello, F. (2001). EGFR-targeted agents potentiate the antitumor activity of chemotherapy and radiotherapy. *SIGNAL*, **2**, 4–11.

Evan, G. I. and Vousden, K. H. (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature*, **411**, 342–348.

Fidler, I. J. and Radinsky, R. (1996). Editorial: search for genes that suppress cancer metastasis. *Journal of the National Cancer Institute*, **88**, 1700–1703.

Herynk, M. and Radinsky, R. (2000). The coordinated expression of growth factor receptors in human colon carcinoma metastasis. *In Vivo*, **14**, 587–596.

Radinsky, R. (1995). Modulation of tumor cell gene expression and phenotype by the organ-specific metastatic environment. *Cancer and Metastasis Reviews*, **14**, 323–338.

Wells, A. (2000). The epidermal growth factor receptor (EGFR) – a new target in cancer therapy. *SIGNAL*, **1**, 4–11.

Chapter 96

Signal Transduction Pathway Targeting

Stanley B. Kaye and Paul Workman

Institute of Cancer Research and Royal Marsden Hospital, Sutton, Surrey, UK

CONTENTS

- Introduction
- Farnesyl Transferase Inhibitors (FTIs)
- Cyclin-dependent Kinase Inhibitors (CDKIs)
- Receptor Tyrosine Kinase Inhibitors (RTKIs)
- Protein Kinase C Inhibitors (PKCIs)
- Other Targets
- Future Perspectives
- Acknowledgement

INTRODUCTION

The term signal transduction is used to describe the molecular processes involved in the communication between the cell and its environment, and in the regulation of cell fate (Downward, 2001). The rationale for developing signal transduction inhibitors as anticancer agents is clear. New anticancer agents are targeted increasingly to the specific abnormalities in the sequence and level of expression of a series of key genes that combine together to drive the progression of human cancer (Hanahan and Weinberg, 2000; Evan and Vousden, 2001). Many of the genetic abnormalities in cancer cells lead to the activation of proliferative signal transduction pathways, deregulation of cell cycle control and the activation of antiapoptotic and cell survival pathways (Hanahan and Weinberg, 2000). In addition, pathways leading to angiogenesis, invasion and metastasis are upregulated (Hanahan and Weinberg, 2000).

The new generation of molecular therapeutics targeted specifically to these deregulated pathways should be more effective and less toxic than the broadly antiproliferative cytotoxic drugs which dominate current therapy (Garrett and Workman, 1999; Workman, 2000). This promise has still to be fully evaluated in hypothesis-testing clinical trials, but progress with the anti-erbB2 monclonal antibody Herceptin (trastuzumab), the Bcr-Abl inhibitor Glivec (imitanib, STI571; CGP57148B) and the epidermal growth factor receptor tyrosine kinase inhibitor Iressa (ZD1839) has very clearly exemplified the clinical potential of molecular therapeutics targeted to deregulated signal transduction pathways (Garrett and Workman, 1999; Workman, 2000, 2001a).

Although most cancers are likely to be driven by the combinatorial effects of several different oncogenes and tumour-suppressor genes, it is possible that therapeutic modulation of any one deregulated target or pathway may be sufficient to give a therapeutic effect (Workman, 2000, 2001a). Such effects are likely to be cytostatic rather than cytotoxic. Hence clinical trials will have to be planned very carefully in order to identify therapeutic activity (Gelmon *et al.*, 1999). In addition, pharmacodynamic endpoints will be required to confirm the intended molecular mode of action and to guide dose selection (Garrett and Workman, 1999; Gelmon *et al.*, 1999; Workman, 2000, 2001a). It is essential to bear in mind that signal transduction inhibitors may have multiple actions, not only in terms of molecular locus, but also with respect to the biological effects achieved, e.g. cell cycle arrest, induction of apoptosis or inhibition of angiogenesis. Although single agent activity may be seen, it seems more likely that optimal use of the new molecular therapeutics will be in combination. Such combinations may involve multiple agents targeted to various signal transduction pathways.

However, in considering the introduction of signal transduction inhibitors into clinical practice, it should be emphasized that it will certainly be the case that conventional cytotoxics will continue to provide a large proportion of the available therapy for the next 5–10 years. Hence an important challenge will be to develop more effective ways of using cytotoxic agents, in particular to overcome the key clinical problem of drug resistance. Here signal transduction inhibitors may play a key role. It is well understood that the response of cancer cells to chemotherapy or radiotherapy represents a balance between cell death (usually by apoptosis) and cell survival (generally following damage repair during cell cycle arrest) and that this balance can be altered by increasing or decreasing the activity of those cell signals leading either to apoptosis or damage repair. Thus it is evident that signal transduction inhibitors can fundamentally alter the chemo- and radiosensitivity of cancer cells (Dent and Grant, 2001). Their clinical development must therefore recognize this

considerable opportunity as part of a combined therapeutic approach. This is not to dismiss their potential as single agents, which already is becoming evident in clinical trials, but it serves to remind those involved in their development of the importance of pursuing both avenues.

Previous reviews have described progress with the development of signal transduction inhibitors (e.g. Gibbs, 2000; Workman, 2000, 2001b). The emphasis of the present review will be on the progress made and challenges ahead with the introduction of these agents into clinical trials and routine practice. The current clinical status of the various candidate signal transduction inhibitors will be summarized briefly according to the target, with reference made in each case to both single agent and combination studies. Particular attention will be focused on:

- farnesyl transferase inhibitors;
- cyclin-dependent kinase inhibitors;
- receptor tyrosine kinase inhibitors;
- protein kinase C inhibitors.

This chapter will concentrate mainly on small molecule enzyme inhibitors, although other approaches to signal inhibition are also in clinical development. Selected antibodies acting on signal transduction will be discussed. Antisense inhibitors based around phosphorothioate oligonucleotides have been shown to retain biological activity following continuous intravenous administration. Targets for these agents include the H-*ras* oncogene, c-*Raf*-1 kinase (and hence the MAP kinase signalling cascade), *protein kinase C*α and also the antiapoptotic oncogene *bcl-2*. In each case, only modest toxicity has been noted; pharmacodynamic studies have indicated the potential for reduced target expression in peripheral blood mononuclear cells and hints of activity in phase I trials in various tumours have been sufficient to justify further development of antisense drugs, both as single agents and in combination with conventional chemotherapy (Cunningham *et al*., 2000).

FARNESYL TRANSFERASE INHIBITORS (FTIs)

The receptor tyrosine kinase→Ras→MAP kinase pathway is frequently activated in tumours by mutation or overexpression of *ras* oncogenes, growth factors or growth factor receptors (Workman, 2001b). This stimulated the discovery of FTIs, on the basis that inhibition of the post-translational modification and membrane localisation of Ras would block proliferative signal transduction. At least five FTIs have entered clinical trials in humans and data are available on four, R115777 (Janssen), SCH6636 (Schering-Plough), L778123 (Merck) and BMS 214662. Clinical experience has recently been summarized (Johnston, 2000). In all cases, prolonged exposure schedules have been investigated in phase I trials. R115777 and SCH6636 have used oral formulations, L778123 and BMS 214662 are administered intravenously.

R115777 has been administered over 21 days, and the dose-limiting toxicity (at 400 mg b.d.) is a combination of myelosuppression, fatigue and peripheral neuropathy. Higher doses, up to 1300 mg b.d. for 5 days every 2 weeks, are feasible, but not on a continuous basis (Zujewski *et al*., 2000). Hints of clinical activity were seen in the initial studies in patients with pancreatic, colon and non-small-cell lung cancer and plasma levels of R115777 at the MTD exceeded those which were clearly active *in vitro*. More recently, two disease-specific studies with R115777 have indicated significant antitumour efficacy, in six of 20 patients with refractory acute leukaemia (Lancet *et al*., 2000), and three of 27 patients with advanced breast cancer (Johnston *et al*., 2000). In the leukaemia study, *ex vivo* inhibition of FT enzyme activity and MAP kinase phosphorylation inhibition was confirmed in sequential bone marrow samples. Combination schedules involving R115777 together with gemcitabine and 5-fluorouracil (5FU)–leucovorin have also been completed, and toxicity has been predictable and manageable.

SCH6636 was administered in various oral regimes, and in addition to myelosuppression, other aspects of dose-limiting toxicity included diarrhoea, vomiting and nephrotoxicity. Regimes have included oral dosing for 1 week followed by a 3-week rest period (Adjei *et al*., 2000), a 2-week on/2-week off schedule and continuous oral dosing (Hurwitz *et al*., 1999). Using the first of these, at the MTD of 400 mg b.d., a biological surrogate marker of activity, i.e. inhibition of prelamin A farnesylation in buccal mucosa, was assessed, and effective inhibition was noted. With this schedule a patient with non-small-cell lung cancer obtained a partial response, while stable disease was noted in a number of other patients, with treatment duration ranging up to 9 months. Combination regimens have included SCH6636 together with paclitaxel and gemcitabine.

For L778123, an intravenous formulation was used with a 7-day administration in a 3-week cycle. At the maximum tolerated dose (MTD) of 560 mg/m^2 day^{-1}, dose-limiting toxicity comprised myelosuppression and nausea (Britten *et al*., 1999). A second study, with a 14- or 28-day continuous infusion, reached an MTD of 840 mg/m^2, with dose-limiting toxicity comprising myelosuppression and cardiac conduction changes (Rubin *et al*., 2000a). At a half of this dose, a surrogate tissue assay – inhibition of farnesylation of hDJ2 in peripheral blood mononuclear cells – showed effective inhibition. L778123 has been given together with paclitaxel, and also radiotherapy (based on preclinical evidence of enhanced radiation-induced apoptosis).

For BMS 214662, three phase I trials (using single 1 h, daily ×5, and weekly intravenous infusions) are ongoing. Nausea, vomiting and diarrhoea have been noted as the major toxicities, and clinical activity has been seen in patients with breast, colorectal and non-small-cell lung

cancer. Dose-related inhibition of farnesylation was noted in peripheral blood cells (Ryan *et al*., 2000). Further trials, combining the drug with conventional chemotherapy, are also planned.

It should be emphasized that the precise molecular mechanism responsible for the antitumour activity of FTIs is not at all clear and remains controversial (Workman, 2001b). A number of cellular proteins are farnesylated and inhibition of any of these, in addition to Ras, may contribute to the therapeutic activity. One potentially important consequence of treatment with FTIs is the switch from farnesylation to geranylgeranylation of RhoB, leading to apoptosis.

In summary, whatever their precise mode of action, FTIs showed a range of manageable antiproliferative and other side effects, and inhibition of the target enzyme in surrogate tissues has been confirmed. However, no molecular studies have yet been reported using solid tumour biopsies from treated patients. Interesting antitumour activity for single agent therapy has been seen, particularly with R115777. Major challenges for the future include the choice of dose and schedule, with the balance appearing to lie between continuous exposure and normal tissue toxicity. An additional question will be the choice of tumours to test, since a large number either have mutated or deregulated Ras; however, the relevance of this is now unclear. Finally, these agents have also demonstrated synergy with a range of cytotoxic agents, perhaps because RhoB, which is probably an important intracellular target, is involved in apoptosis and DNA repair. Combination trials, using FTIs with chemotherapy will thus be viewed with particular interest.

CYCLIN-DEPENDENT KINASE INHIBITORS (CDKIs)

Cyclin-dependent kinases are frequently deregulated in cancer by a variety of means, including overexpression of the cyclin and kinase components, together with mutation or loss of negative regulators such as p16 (Garrett and Workman, 1999). The first CDKI to be tested clinically was flavopiridol. This inhibits CDK1, -2 and -4 (and at higher doses it also inhibits other protein kinases including EGF receptor tyrosine kinase). Flavopiridol has antitumour activity in its own right and has also shown schedule-dependent synergy *in vitro* with paclitaxel. In the first phase I trial using a 72-h intravenous infusion every 2 weeks, dose-limiting toxicity initially comprised secretory diarrhoea (at $62.5\,\mathrm{mg/m^2\,day^{-1}}$). With antidiarrhoeal prophylaxis a higher dose ($98\,\mathrm{mg/m^2\,day^{-1}}$) was achievable, before dose-limiting hypotension was also noted (Senderowicz *et al*., 1998). In addition a characteristic ‘proinflammatory syndrome’ comprising ’flu-like symptoms and local tumour pain occurred at the higher dose levels. Pharmacokinetic analyses indicated that concentrations which inhibited all three CDKs were achievable, and one patient with renal cancer achieved a partial response, with minor responses in patients with colon cancer and lymphoma. However, a subsequent phase II trial in renal cancer, using the same schedule, has shown minimal activity (Stadler *et al*., 2000).

Of greater interest is a second phase I trial which involved a 24-h infusion of flavopiridol, either preceded or followed by a 24-h infusion of paclitaxel (Schwartz *et al*., 1998). Apoptosis was measured in peripheral blood lymphocytes, and a clear difference was noted with respect to drug sequencing. The sequence of paclitaxel followed by flavopiridol led to a marked increase in apoptosis, whereas an actual reduction was noted with the reverse sequence. These observations, although preliminary and not involving tumour cells, may have important implications for future combination regimes, since on theoretical grounds the interaction between a cell cycle inhibitor and conventional chemo- or radiotherapy may be either synergistic or antagonistic, and scheduling could be a key determinant.

Another cell cycle inhibitor which directly inhibits CDK1 and -2 is staurosporine. The benzoyl derivative UCN-01 has other actions, and at lower concentrations it inhibits the G2/M checkpoint kinase Chk1. This would be expected to potentiate the effects of chemotherapy in tumour cells. A phase I trial of UCN-01 also involved a 72-h i.v. infusion, and dose-limiting side effects included vomiting, hyperglycaemia and pulmonary toxicity (Senderowicz, 2000). Antitumour activity was seen in patients with melanoma and anaplastic lymphoma. Measurements on saliva of free drug levels indicated that concentrations compatible with modulation of the cellular response to a DNA-damaging agent were clinically achievable. An extremely long half-life was also noted, thought to be due to extensive binding of UCN-01 to α-acidic glycoprotein. Shorter infusion phase I protocols are planned. However, at this stage, the drug’s toxicity will probably preclude extensive further investigation in combination schedules. Nevertheless, the preclinical data would argue persuasively for further studies of compounds of a similar type in combination with drugs such as 5FU or platinum, or alternatively with radiation.

A third CDKI in clinical development (within the EORTC) is E7070, a novel chloroindolylsulfonamide which inhibits CDK2 and cyclin E in cancer cells at concentrations ranging from 1.4 to $131.4\,\mu\mathrm{g\,mL^{-1}}$ *in vitro*. Phase I clinical trials have employed a daily $\times 5$ (1 h), 5-day continuous infusion, and weekly infusion (1 h) schedules (Raymond *et al*., 2000a). Myelosuppression is dose-limiting, clinical responses have been seen in patients with breast and uterine cancer and pharmacodynamic studies are ongoing.

A range of new CDKIs are now emerging with the potential for selectivity against particular CDKs. It might be envisaged that a CDK4 inhibitor could mimic the effects of restoring p16 function. A CDK2 inhibitor could

induce apoptosis in tumour cells with deregulated E2F expression. However, the biological and clinical effects of selective CDK inhibition are likely to be more complex.

RECEPTOR TYROSINE KINASE INHIBITORS (RTKIs)

An alternative approach to signal transduction inhibition is to target at the level of cell membrane receptors, rather than further downstream. There are numerous examples of overexpression of growth factors and/or RTKs in cancer, leading to the potential for autocrine growth stimulation, and in some cases mutations are also seen (Garrett and Workman, 1999). A range of targets have been identified, and inhibitors of varying specificity have been tested clinically. These include inhibitors of EGF (or ErbB-1), PDGF and VEGF RTKs.

There are at least five molecules in clinical development as EGF RTKs: ZD1839 (Iressa; AstraZeneca), CP358774 (Pfizer), CI-1033 (Parke-Davis), EKB 1869 (Wyeth) and PKI 166 (Novartis). In addition a human-to-murine chimaeric monoclonal antibody, C225 (Imclone) has been developed for the purpose of direct inhibition of EGFR, while a second humanized monoclonal antibody (Herceptin) directed against another member of the ErbB family of receptors, the erbB2 or Her-2/neu receptor, has attracted the most attention because of clear evidence of clinical benefit in randomized trials in breast cancer.

ZD1839 is completing phase I clinical trials, using an oral preparation, delivered over 14 or 28 days. At the highest doses given, dose-limiting toxicity includes diarrhoea and skin rash; indeed, a skin rash first visible on the face has been seen at all doses above 250 mg and appears to be characteristic of agents targeting EGF RTK, presumably because of the high level of receptor expression in normal skin (Baselga *et al*., 2000a). Objective tumour responses to ZD1839 given as a single agent have been noted in patients with non-small-cell lung and prostate cancer, with hints of activity in ovarian and head and neck cancer. Phase II single-agent studies are therefore planned, and phase III trials in non-small-cell lung cancer, combining ZD1839 with platinum-based chemotherapy, are also under way. The potential for significant augmentation of cytotoxic chemotherapy, and perhaps circumvention of clinical drug resistance, is particularly intriguing because of the persuasive laboratory data pointing to the role of EGF as a survival signal in cells exposed to DNA-damaging agents. In due course, the clinical schedules used in combination trials may need to take account of these interesting possibilities.

There are fewer clinical data available on the other compounds in development, but experience to date indicates a similar pattern of toxicity. Skin rash has been seen with CP358774 (now renamed OSI-774) and the agent is proceeding to phase II trials, hints of activity having already been detected in non-small-cell lung, head and neck, ovarian and renal cancer (Karp *et al*., 1999). CI-1033 is interesting in that it is in fact an inhibitor of all four members of the ErbB family of receptors and is currently in phase I trials.

As regards CI-225, clinical data on both single-agent and combination regimes are available. Antibody was given either as a single intravenous dose or weekly. As with ZD1839, a characteristic skin rash was noted, otherwise toxicity was minimal. A dose of 200 mg/m^2 was considered optimal, based on pharmacokinetic data, indicating saturation of systemic clearance at that level (Baselga *et al*., 2000b). Moreover, an earlier clinical trial with a murine anti-EGF receptor monoclonal antibody indicated more than 50% receptor saturation of tumour biopsies following treatment of patients at that dose.

Combination trials with chemotherapy and radiation therapy have been particularly interesting. Head and neck and colorectal cancer patients with disease progressive on prior cisplatin-containing or CPT-11 containing treatment were retreated with the same dose of cisplatin (or CPT-11) and weekly injections of CI-225. Seven of eight patients demonstrated a partial response (Rubin *et al*., 2000b). In a separate trial, 15 patients were treated with weekly CI-225 during a course of radiation therapy for locally advanced disease; the overall response rate was 93% (Ezekiel *et al*., 1999).

For Herceptin, initial clinical trials using a weekly intravenous schedule showed that it was well tolerated, and a modest 15% response rate was noted in patients with Her-2/neu positive breast cancer (Baselga *et al*., 1996). Following evidence of synergy in xenograft studies, a randomized trial of chemotherapy (paclitaxel or doxorubicin–cyclophosphamide) with or without Herceptin was conducted in patients with advanced breast cancer, and both response and overall survival were significantly improved (Slamon *et al*., 2001). Herceptin is rapidly finding its way into adjuvant clinical trials in HER-2/neu-positive breast cancer and is also being evaluated in combination with other drugs, particularly platinum, in other cancers. The regulatory approval of Herceptin for use in a genomically defined subgroup of patients represents an important landmark in the evolution of the new molecular therapeutics acting on signal transduction pathways and other cancer genome targets.

There are two PDGF RTKIs which have undergone clinical development, SU101 (Sugen) and CGP57148, now renamed STI571 or Glivec/Gleevec (Novartis). Both are given in continuous oral regimes, and the data with Glivec are particularly intriguing. In addition to its property as an inhibitor of PDGF RTK with an IC_{50} of 300 $nmol\,L^{-1}$, Glivec also has been found to have remarkable potency (IC_{50} 38 $nmol\,L^{-1}$) as an inhibitor of the tyrosine kinase associated with the Bcr/Abl fusion protein present in patients with chronic myeloid leukaemia (CML). This led to a modification of the phase I development plan, with the focus

switching to patients with refractory CML. A remarkable level of clinical activity has been noted, with 96% of patients demonstrating a clinical response at doses of Glivec above 300 mg daily for up to 5 months (Druker and Lydon, 2000). It is too soon to evaluate fully the contribution that the drug will make in the management of this disease, particularly in those with lymphoid blast crisis, but already it is evident that it will have an important clinical role, especially as toxicity so far has been minimal. Indeed the drug has recently obtained rapid regulation for the treatment of CML. For the field in general, this experience illustrates the potential for a new signal transduction inhibitor to make an impact, when a susceptible target can be identified for a specific inhibitor. The potential for drug resistance has been noted and combinations of Glivec with cytotoxic chemotherapy will now be examined. Clinical evaluation of STI571 is now continuing in patients with solid tumours known to express PDGF. These include glioma, prostate and small-cell lung cancer. In this last tumour the inhibition of c-kit-associated tyrosine kinase may also be of importance. Indeed, recent reports of remarkable responses in the rare subtype of c-kit positive sarcoma – the gastrointestinal stromal sarcoma (GIST) – are of great interest.

A second PDGF RTKI SU101 has also had fairly extensive clinical assessment. This is in fact a new formulation of an old agent, leflunomide, originally developed as an immunomodulatory agent in 1982. It does have immunosuppressive properties; these may possibly to relate to its potency as a PDGF RTKI and consequent regulation of B and T cell activity, in addition to its antimetabolite properties. The precise role of inhibition of PDGF receptor signalling versus the more conventional antimetabolite effects is not clear. Preclinical studies confirmed specificity of receptor inhibition, and indicated activity in a range of PDGF receptor expressing tumours *in vitro* and *in vivo*. Clinical trials mostly used weekly intravenous 24-h infusions of SU101. Fatigue, nausea and myelosuppression were noted at the highest doses. Tumour regressions were noted in the first trials in patients with recurrent glioma, and a level of significant clinical benefit as high as 43% was claimed in nonrandomized phase II trials (Norman, 2000). However, CNS toxicity proved to be problematic, and after two treatment-induced deaths occurred, the clinical development of the agent was discontinued.

The clinical development of specific VEGF RTKIs has taken a slightly different route, since here the main target is the endothelial cell–tumour cell interaction, the focus being inhibition of angiogenesis. An example of an agent studied in some detail clinically is SU5416 (Sugen). This primarily blocks VEGF-mediated Flk-1 receptor signalling, but does not inhibit tumour cell growth directly. Phase I trials have involved twice weekly intravenous infusions, and toxicity at the highest dose has included vomiting and headache, but at lower doses the drug is well tolerated. Dose-related changes in vascular permeability have been noted, using gadolinium-enhanced magnetic resonance imaging (MRI) scans, and phase II trials of the agent are under way (O'Donell *et al.*, 2000). In addition, further clinical trials of SU5416 in combination with chemotherapy are proceeding because of additive/synergistic effects noted in human tumour xenografts (Cropp and Hannah, 2000). In particular, the combination of SU5416 and 5FU–leucovorin in patients with colorectal cancer is being actively pursued; preliminary clinical data suggest that significant tumour control can be achieved. In addition, a more potent VEGF RTK inhibitor, SU6668, is also being studied in phase I trials. This has the potential advantage of inhibiting signalling via VEGF, PDGF and also FGF, all of which appear to have a role in tumour angiogenesis.

Alternative approaches to VEGF signalling are being taken with a number of other agents; perhaps the agent with the longest history is thalidomide, which is a potent down-regulator of several peptide signalling molecules, including TNFα and VEGF. Interestingly, thalidomide has demonstrated single-agent antitumour activity in a number of diseases, including renal cancer (Eisen *et al.*, 2000), Kaposi sarcoma and glioma, and insight into the contribution which VEGF antagonism makes would clearly be of value in developing this approach.

One further example of a signalling agent in the clinic is the natural flavonoid quercetin. This is a rather nonspecific protein kinase inhibitor, with antitumour activity in human tumour xenografts. A phase I trial using an intermittent i.v. schedule showed that reversible nephrotoxicity was dose-limiting, and tyrosine kinase inhibition was noted in peripheral blood cells (Ferry *et al.*, 1996). A subsequent phase I trial examined the combination of quercetin with carboplatin, because of hints of activity in ovarian cancer, and this approach is certainly of interest, particularly with agents of better defined specificity.

A key issue for the future will be the extent to which specificity towards a single receptor is or is not desirable. It could be argued that well-defined agents, with the ability to inhibit two or three receptors, may find broader application. Examples which will be studied in the clinic in the near future include SU5418, which inhibits VEGF, PDG and FGF receptors. (See also chapter *Growth Factor Receptor Blockade*.)

PROTEIN KINASE C INHIBITORS (PKCIs)

Protein kinase C is one of the most ubiquitous of intracellular signal transduction targets for inhibition (Caponigro *et al.*, 1997) At least 12 isoforms exist, and these are intimately involved in cell signals leading to proliferation and differentiation. PKC isoenzymes consist of a single-polypeptide chain divided into a regulatory and a catalytic domain, and a number of modulators acting mainly at one or other domain have been tested clinically. Single-agent antitumour activity is clearly possible, the best example being tamoxifen, which is both a PKC inhibitor (IC_{50} 40–100 μmol L^{-1}) and

an oestrogen receptor blocker (Couldwell *et al*., 1996). However, there is probably greater potential for these agents as modulators of currently available cytotoxic drugs. This may occur through a number of mechanisms. Staurosporine is one of the most potent PKC inhibitors, but is poorly selective. Among several analogues with greater specificity for certain PKC isoenzymes, CGP 41251 (Midostaurin, Novartis) has been most extensively tested. It is an *N*-benzoyl derivative, most effectively inhibiting cytosolic PKC (IC_{50} 0.5 $\mu mol\,L^{-1}$). It has modest activity in its own right in a range of xenografts, and can reverse P-glycoprotein-mediated drug resistance, restoring doxorubicin sensitivity in a number of models (Crosios, 2000).

A phase I trial of CGP 41251 involving daily oral dosing of 28 days, used a novel surrogate endpoint to test for target (PKC) inhibition. This comprised a measurement of TNF and IL-6 production in response to mitogen stimulation of white blood cells taken from treated patients, the response being significantly lowered as a consequence of (reversible) PKC inhibition. Toxicity comprised nausea and lethargy, but the trial was concluded when the number of daily tablets reached 12, rather than through confirmation of a maximum tolerated dose (Propper *et al*., 2001). Preliminary reports of clinical improvement have been described in a phase II study of CGP41251 in patients with low-grade lymphoma, and clinical trials of combination treatment with chemotherapy and radiotherapy are ongoing.

A second PKC antagonist, bryostatin, has also entered clinical trials in recent years. This is a macrocyclic lactone, derived from a marine bryozoan, and in different conditions it is in fact capable of both PKC antagonist and agonist activity. Clinical trials have involved intravenous infusions of various durations, and the dose-limiting toxicity of myalgia may relate to the drug's ability to stimulate cytokine production (Jayson *et al*., 1995). Responses were seen in patients with ovarian cancer and lymphoma, and phase II trials are ongoing, while combination regimes are also being planned.

OTHER TARGETS

Clearly the range of cell signals which could represent targets for new drug development in cancer is considerable, and likely to expand further as the sequencing of the human genome is finalized (International Genome Consortium, 2001; Venter *et al*., 2001). Complete sequencing of the genomes of cancer cells is likely to identify new cancer gene targets (Futreal *et al*., 2001; Workman, 2001a). **Table 1** lists some of the main targets for which drugs are in preclinical and clinical development for cancer treatment. These include p70S6 kinase, phosphorylation and activation of which is inhibited by a novel rapamycin analogue CCI-779 (Wyeth). This compound is actually an inhibitor of mTOR and is proving to be of considerable interest, not least because of modest toxicity in phase I, associated with clinical activity in renal and breast cancer (Raymond *et al*., 2000b). mTOR inhibitors may have particular utility in tumours which have lost the function of the *PTEN* tumour-suppressor gene, leading to activation of the PI3 kinase pathway, although this is by no means clear. Also of potential interest in view of the role of PI3 kinase in proliferation and apoptosis would be the development of PI3 kinase inhibitors.

Table 1 Current genomic targets for drugs in preclinical and clinical development

EGF receptor tyrosine kinase	Farnesyl transferase
ErbB2 receptor tyrosine kinase	Protein kinase C
VEGF receptor tyrosine kinase	Hsp90
PDGF receptor tyrosine kinase	Histone deacetylases
FGF receptor tyrosine kinase	Histone acetylases
Src tyrosine kinase	p53-mdm2 interaction
Raf-1 kinase	Integrins
MEK kinase	Proteases
Bcr/Abl kinase	Telomerase
Cyclin-dependent kinase	HIF-1α
mTOR	Poly-ADP ribose polymerase
p70S6 kinase	DNA protein kinase
PI3 kinase	Chk 1 kinase
Aurora kinase	Wnt signalling pathway

The degradation of intracellular proteins, including those responsible for signalling apoptosis (including p27, p21, p53 and NF-kappa B), is now known to be a primary function of the ubiquitin–proteasome pathway (Garrett and Workman, 1999). This could be a crucially important cell signalling regulation pathway to target for inhibition, and a potent molecule, PS-341, has now been developed for this purpose. As with other signalling inhibitors, a rationale could be developed both for an antitumour affect on its own, and for an anticipated enhancement of the affect of conventional chemotherapy. Both predictions have been borne out in laboratory studies, and the first phase I trials with an oral dose schedule of twice-weekly PS-341 indicate the feasibility of this approach. Effective inhibition of the 20S-proteasome has been demonstrated in lymphocytes with minimal toxicity seen so far, and hints of antitumour activity have been noted in a patient with melanoma (Hamilton *et al*., 2000). Further clinical trials (single agent and chemotherapy combinations) are awaited with interest.

The molecular chaperone protein HSP90 is also proving to be a target of considerable interest, because of its central role in the intracellular processing of key signalling proteins (Kelland *et al*., 1999). Inhibition of HSP90 and its close homologue GRP94 leads to proteasomal degradation of a range of important oncogenic client proteins, including c-Raf-1, CDK4, Akt/PKB, erbB2 and also mutant p53, leading to cell cycle arrest and apoptosis (Hostein *et al*., 2001). A phase I trial of the geldanamycin analogue,

17AAG, which targets the HS90 chaperone is now under way, with proof of principle studies focusing on measurement of elevated HSP70 levels in peripheral blood lymphocytes and tumour biopsies (Clarke *et al*., 2000), as well as changes in client proteins such as c-Raf-1 (Banerji *et al*., 2000).

Modulators of histone acetylation, and hence of gene transcription, also have exciting potential (Kouzarides, 2000; Workman, 2001c).

FUTURE PERSPECTIVES

The preclinical development and clinical testing of signal transduction inhibitors is a rapidly evolving field. However, two issues of particular note are:

1. phase I trials require to be designed carefully as 'proof-of-principle' hypothesis-testing studies, with extensive use of pharmacodynamic endpoints (Gelmon *et al.*, 1999);
2. the combination of these novel cytostatic agents with conventional chemotherapy must be an essential part of their clinical evaluation, because of the complex potential nature of these interactions at the level of the tumour cell (Dent and Grant, 2001; Waldeman *et al.*, 1997).

The next 5 years will be very stimulating as a range of mechanism-based signal transduction inhibitors undergo clinical evaluation as part of an overall strategy of individualized, genome-based cancer therapy.

ACKNOWLEDGEMENT

The authors' work is funded by the Cancer Research Campaign (CRC) and PW is a CRC Life Fellow.

REFERENCES

Adjei, A. A., *et al*. (2000). Phase I trial of the farnesyl transferase inhibitor SCH66336; evidence for biological and clinical activity. *Cancer Research*, **60**, 1871–1877.

Banerji, U., *et al*. (2000). PK-PD relationships for the HSP90 molecular chaperone inhibitor 17AAG in human ovarian cancer xenografts. *Clinical Cancer Research*, **6**, Suppl., 4545S.

Baselga, J., *et al*. (1996). Phase II study of weekly i.v. recombinant humanized anti HER2 monoclonal antibody in patients overexpressing HER2/neu. *Journal of Clinical Oncology*, **14**, 737–44.

Baselga, J., *et al*. (2000a). Continuous administration of ZD1839, a novel EGFrtk inhibitor in patients with 5 selected tumour types: evidence of activity and good tolerability. *Proceedings of the American Society for Clinical Oncology*, **19**, 177a.

Baselga, J., *et al*. (2000b). Phase I studies of anti-epidermal growth factor receptor chimeric antibody C225 alone and in combination with cisplatin. *Journal of Clinical Oncology*, **18**, 904–914.

Britten, C. D., *et al*. (1999). The FTase inhibitor L788,123 in patients with solid cancers. *Proceedings of the American Society for Clinical Oncology*, **18**, 156a.

Caponigro, F., *et al*. (1997). Protein kinase C: a worthwhile target for anticancer drugs? *Anti-Cancer Drugs*, **8**, 26–33.

Clarke, P. A., *et al*. (2000). Gene expression profiling of human colon cancer cells following inhibition of signal transduction by 17-allylamino-17-demethoxy geldanamycin, an inhibitor of the Hsp90 molecular chaperone. *Oncogene*, **19**, 4125–4133.

Couldwell, W. T., *et al*. (1996). Treatment of recurrent malignant glioma with chronic oral high dose tamoxifen. *Clinical Cancer Research*, **2**, 619–622.

Cropp, G. F. and Hannah, A. L. (2000). SU5416, a molecularly targeted novel antiangiogenic drug: clinical PK and safety review. *Clinical Cancer Research*, **6**, Suppl., 4518S.

Crosios, K. (2000). Midostaurin. *Current Opinion in Oncologic, Endocrine and Metabolic Investigational Drugs*, **2**, 74–82.

Cunningham, C. C., *et al*. (2000). Phase I trial of c-Raf kinase antisense oligonucleotide ISIS 5132 in patients with advanced cancer. *Clinical Cancer Research*, **6**, 1626–1631.

Dent, P. and Grant, S. (2001). Pharmacologic interruption of the mitogen-activated extracellular-regulated kinase/mitogen-activated protein kinase signal transduction pathway. Potential role in promoting cytotoxic drug action. *Clinical Cancer Research*, **7**, 775–783.

Downward, J. (2001). The ins and outs of signalling. *Nature*, **411**, 759–762.

Druker, B. J. and Lydon, N. B. (2000). Lessons learnt from the development of an Abl tyrosine kinase inhibitor for CML. *Journal of Clinical Investigation*, **105**, 3–7.

Eisen, T., *et al*. (2000). Continuous low dose thalidomide; phase II study in advanced melanoma, renal cell, ovarian and breast cancer. *British Journal Cancer*, **82**, 812–817.

Evan, G. I. and Vousden, K. H. (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature*, **411**, 342–348.

Ezekiel, M., *et al*. (1999). Phase I trial of anti-EGFr antibody in combination with irradiation for locally advanced head and neck malignancies. *Proceedings of the American Society for Clinical Oncology*, **18**, 388a.

Ferry, D. F., *et al*. (1996). Phase I clinical trial of the flavonoid quercetin: pharmacokinetics and evidence for *in vivo* tyrosine kinase inhibition. *Clinical Cancer Research*, **2**, 659–668.

Futreal, P. A., *et al*. (2001). Cancer and genomics. *Nature*, **409**, 850–855.

Garrett, M. D. and Workman, P. (1999). Millenium review. Discovering novel chemotherapeutic drugs for the third millennium. *European Journal of Cancer*, **35**, 2010–2030.

Gelmon, K. A., *et al*. (1999). Anticancer agents targeting signalling molecules and cancer cell environment: challenges for drug development. *Journal of the National Cancer Institute*, **19**, 1281–1287.

Gibbs, J. B. (2000). Mechanism-based target identification and drug discovery in cancer research. *Science*, **287**, 1969–1973.

Hamilton, A., *et al*. (2000). Proteasome inhibition by PS341. A phase I study. *Clinical Cancer Research*, **6**, Suppl., 4549S.

Hanahan, D. and Weinberg, A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Hostein, I., *et al*. (2001). Inhibition of signal transduction by the Hsp90 inhibitor 17-allylamino-17-demethoxygeldanamycin results in cytostasis and apoptosis. *Cancer Research*, **61**, 4003–4009.

Hurwitz, H. I., *et al*. (1999). Phase I and pharmacokinetic study of SCH66336, a novel FTI using a 2 week on, 2 week off schedule. *Proceedings of the American Society for Clinical Oncology*, **18**, 156a.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.

Jayson, G. C., *et al*. (1995). Phase I study of bryostatin I in patients with advanced malignancy using a 24 hr. intravenous infusion. *British Journal of Cancer*, **72**, 461–468.

Johnston, S. R. D. (2000). Farnesyl transferase inhibitors – a novel targeted therapy for cancer. *Lancet Oncology*, **2**, 18–26.

Johnston, S. R. D., *et al*. (2000). Phase II study of the farnesyl transferase inhibitor R115777 in patients with advanced breast cancer. *Proceedings of the American Society for Clinical Oncology*, **19**, 83a.

Karp, D., *et al*. (1999). Phase I dose escalation study of the EGFrtk inhibitor CP358774 in patients with advanced cancer. *Proceedings of the American Society for Clinical Oncology*, **18**, 388a.

Kelland, L. R., *et al*. (1999). DT-Diaphorase expression and tumour cell sensitivity to 17-allylamino, 17-demethoxygeldanamycin, an inhibitor of heat shock protein 90. *Journal of the National Cancer Institute*, **91**, 1940–1949.

Kouzarides, T. (2000). Acetylation: a regulatory mechanism to rival phosphorylation? *EMBO Journal*, **19**, 1176–1179.

Lancet, J., *et al*. (2000). Use of farnesyl transferase inhibitor R115777 in relapsed and refractory acute leukaemia; results of phase I trial. *Proceedings of the American Society for Clinical Oncology*, **19**, 3a.

Norman, P. (2000). SU 101. *Current Opinion in Oncologic, Endocrine and Metabolic Investigational Drugs*, **2**, 104–114.

O'Donnell, A., *et al*. (2000). Phase I trial of the VEGF inhibitor SU5416 incorporating dynamic MRI assessment of vascular permeability. *Proceedings of the American Society for Clinical Oncology*, **19**, 177a.

Propper, D. J., *et al*. (2001). Phase I and PK study of PKC412, an inhibitor of protein kinase C. *Journal of Clinical Oncology*, **19**, 1485–92.

Raymond, E., *et al*. (2000a). Combined results of 4 Phase I and PK studies of E 7070. *Clinical Cancer Research*, **6**, Suppl., 4529S.

Raymond, E., *et al*. (2000b). CCI-779, an ester analogue of rapamycin that interacts with PTEN/P13 kinase pathways: Phase I study utilizing a weekly intravenous schedule. *Clinical Cancer Research*, **6**, Suppl., 4549S.

Rubin, E., *et al*. (2000a). Phase I trials of the FTase inhibitor L778 123 on 14 or 28 day dosing schedule. *Proceedings of the American Society for Clinical Oncology*, **19**, 178a.

Rubin, M. S., *et al*. (2000b). IMC 225 – an EGFr monoclonal antibody for patients with EGFr-positive tumours refracting to previous therapy. *Proceedings of the American Society for Clinical Oncology*, **19**, 474a.

Ryan, D. P., *et al*. (2000). Phase I clinical trial of the FTase inhibitor BMS 214662 in patients with advanced solid tumours. *Proceedings of the American Society for Clinical Oncology*, **19**, 185a.

Schwartz, G. K., *et al*. (1998). Flavopiridol enhances the biological effects of paclitaxel: a Phase I trial in patients with advanced solid tumours. *Proceedings of the American Society for Clinical Oncology*, **17**, 188a.

Senderowicz, A. M., *et al*. (1998). Phase I trial of continuous infusion flavopiridol, a novel cyclin-dependent kinase inhibitor in patients with refractory neoplasms. *Journal of Clinical Oncology*, **16**, 2986–2999.

Senderowicz, A. M. (2000). Preclinical and clinical development of CDK inhibitors. *Clinical Cancer Research*, **6**, Suppl., 4482S.

Slamon, D. J., *et al*. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer. *New England Journal of Medicine*, **344**, 783–792.

Stadler, W., *et al*. (2000). Flavopiridol, a novel CDK inhibitor, in metastatic renal cancer; a University of Chicago phase II consortium study. *Journal of Clinical Oncology*, **18**, 371–375.

Venter, J. C., *et al*. (2001). The sequence of the human genome. *Science*, **291**, 1304–1351.

Waldeman, T., *et al*. (1997). Cell-cycle arrest versus cell death in cancer therapy. *Nature Medicine*, **3**, 1034–1037.

Workman, P. (2000). Towards genomic cancer pharmacology: innovative drugs for the new millennium. *Current Opinion in Oncologic, Endocrine and Metabolic Investigational Drugs*, **2**, 21–25.

Workman, P. (2001a). New drug targets for genomic cancer therapy; successes, limitations, opportunities and future challenges. *Current Cancer Drug Targets*, **1**, 33–47.

Workman, P. (2001b). Signal transduction pathways: a goldmine for therapeutic targets. In: Sebti, S. M. and Hamilton, A. D. (eds), *Farnesyltransferase Inhibitors in Cancer Therapy*. 1–20 (Humana Press, Totowa, NJ).

Workman, P. (2001c). Scoring a bull's-eye against cancer genome targets. *Current Opinion in Pharmacology*, **1**, 342–352.

Zujewski, J., *et al*. (2000). Phase I and pharmacokinetic study of farnesyl protein transferase inhibitor R115777 in advanced cancer. *Journal of Clinical Oncology*, **18**, 927–941.

FURTHER READING

Downward, J. (2001). The ins and outs of signalling. *Nature*, **411**, 759–762.

Evan, G. I. and Vousden, K. H. (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature*, **411**, 342–349.

Futreal, P. A., *et al*. (2001). Cancer and genomics. *Nature*, **409**, 850–855.

Garrett, M. D. and Workman, P. (1999). Millenium review. Discovering novel chemotherapeutic drugs for the third millennium. *European Journal of Cancer*, **35**, 2010–2030.

Gelmon, K. A., *et al*. (1999). Anticancer agents targeting signalling molecules and cancer cell environment: challenges for drug development. *Journal of the National Cancer Institute*, **19**, 1281–1287.

Gibbs, J. B. (2000). Mechanism-based target identification and drug discovery in cancer research. *Science*, **287**, 1969–1973.

Hanahan, D. and Weinberg, A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.

Venter, J. C., *et al*. (2001). The sequence of the human genome. *Science*, **291**, 1304–1351.

Workman, P. (2000). Towards genomic cancer pharmacology: innovative drugs for the new millennium. *Current Opinion in Oncologic, Endocrine and Metabolic Investigation Drugs*, **2**, 21–25.

Workman, P. (2001a). New drug targets for genomic cancer therapy; successes, limitations, opportunities and future challenges. *Current Cancer Drug Targets*, **1**, 33–47.

Workman, P. (2001b). Signal transduction pathways: a goldmine for therapeutic targets. In: Sebti, S. M. and Hamilton, A. D. (eds), *Farnesyltransferase Inhibitors in Cancer Therapy*. 1–20 (Humana Press, Totowa, NJ).

Workman, P. (2001c). Scoring a bull's eye against cancer genome targets. *Current Opinion in Pharmacology*, **1**, 342–352.

Chapter 97

DNA Repair and the Cell Cycle as Targets in Cancer Therapy

Manfred F. Rajewsky
University of Essen Medical School and West German Cancer Center Essen, Essen, Germany
Rolf Müller
Phillips University of Marburg Medical School, Marburg, Germany

CONTENTS

- DNA Damage and Repair
- The Mammalian Cell Cycle
- Acknowledgements

DNA DAMAGE AND REPAIR

General Aspects

The genetic information specifying the phenotypes, functional activities and the adaptive and reparative capacity of our cells is stored in the linear sequence of approximately 3×10^9 copies of the four bases guanine, cytosine, adenine and thymine, aligned in the DNA macromolecule. This central molecular blueprint needs to be preserved intact and must be reproduced by DNA replication with utmost fidelity whenever cells prepare for division.

During the lifetime of cells and organisms, the structural and functional integrity of the genome is constantly compromised by various forms of DNA damage that either occurs 'spontaneously' or is caused by exogenous and endogenous DNA-reactive agents. On the one hand, DNA damage that accumulates in cells can be cytotoxic, i.e. cause programmed cell death (apoptosis) or other forms of cell breakdown. However, DNA lesions can also result in the mutation of genes whose inactivation or functional alteration subverts regular cell type-specific properties and may enhance cancer risk.

In recent years it has become evident that the recognition and processing of cellular DNA damage is tightly interconnected with the transcriptional status of target genes, the regulation of DNA repair gene transcription and the control of cell cycle progression and apoptosis via so-called DNA damage checkpoints. The efficacy of distinct modes of DNA repair differs interindividually, with age and among different cell types (most likely also depending on their state of differentiation).

Human cancer cells very often exhibit defects in 'checkpoint' and cell cycle regulation – notably in the p16^{INK4a}–cyclin D1–pRb pathway (see The Mammalian Cell Cycle below) – and even among histologically similar tumours there is wide variability with respect to DNA repair capacity. The latter is a reflection of the genetic and phenotypic heterogeneity and instability of cancer cells, which also explains the frequent generation of variant cell subpopulations within tumours. Together, these properties strongly influence the sensitivity of malignant cells toward DNA-reactive therapeutic agents and represent potential targets for modulation to achieve more effective therapy.

Spontaneous and Induced DNA Damage

In spite of the extraordinary precision of the molecular machinery for DNA replication (notably of the DNA polymerases), replication errors resulting in mismatched bases in the DNA double helix do occur, albeit at relatively low frequency. This is not surprising given the fact that about 10^{10} bases need to be paired correctly with their complementary bases in a diploid human cell per round of cell division, corresponding to about 10^{25} bases during the average human lifetime (assuming a total of 10^{15} cell divisions). Other forms of spontaneous (i.e. unpreventable) structural DNA alterations include noninstructional abasic sites in the DNA molecule, resulting from the loss (via hydrolytic cleavage of *N*-glycosylic bonds) of purine or, less frequently, pyrimidine bases; the conversion of cytosine or 5-methylcytosine to uracil or thymine, respectively, via deamination; a broad spectrum of DNA lesions caused by reactive oxygen species (ROS) that are produced, e.g., in the course of cellular metabolic processes; and alkylated bases formed in DNA by reaction with *S*-adenosylmethionine or with alkylating compounds generated by bacterially catalysed nitrosation of endogenous amides or amines.

To a large extent, however, potentially mutagenic and cytotoxic DNA lesions result from exposure to exogenous DNA-reactive agents, including chemicals of environmental or nutritional origin, solar ultraviolet (UV) light, ionizing radiation and cancer chemotherapeutic drugs. The respective DNA lesions are 'agent-specific,' i.e. they reflect the chemical nature and DNA reactivity of a given compound. The molecular structures of such DNA lesions are therefore highly diverse, and in many cases remain to be clarified because an unknown, but undoubtedly large, number of exogenous DNA reactants still await identification and new chemicals, including cancer therapeutic drugs, continue to be synthesized.

Nevertheless, a considerable number of specific DNA lesions have already been structurally characterized, notably the DNA reaction products of various mono- and bifunctional alkylating *N*-nitroso compounds, some polycyclic hydrocarbons and heterocyclic amines, aflatoxin B_1, UV light, ionizing radiation and ROS. Exogenous DNA alterations have been operationally classified into 'bulky,' helix-distorting lesions that include large-sized base monoadducts or the intra- and interstrand DNA cross-links typically induced by bifunctional anticancer agents (e.g. chloroethylnitrosoureas, platinum derivatives; see later sections) versus small, essentially nondistorting base adducts. (See also the chapters *Mechanisms of Chemical Carcinogenesis* and *The Formation of DNA Adducts*.)

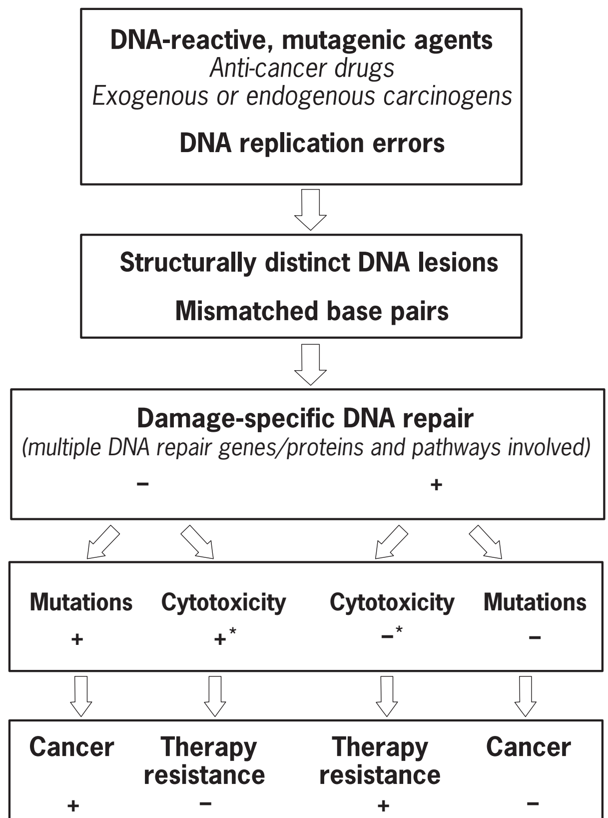


Figure 1 DNA repair: counteragent in mutagenesis and carcinogenesis, accomplice in cancer therapy resistance.

DNA Repair Mechanisms and Molecular Pathways Involved

Timely and error-free repair is the key protective mechanism counteracting the generation of mutations from premutational DNA alterations, thereby holding cancer risk in check. However, because the DNA machinery of the cell can process both potentially mutagenic and cytotoxic DNA lesions, DNA repair also enables cells to cope with an otherwise deleterious load of DNA damage (**Figure 1**).

A large number of genes and proteins have been, and continue to be, identified as being directly or indirectly involved in DNA repair. Not unlike the immune system with its surveillance functions at the cellular level, the DNA repair machinery of our cells represents a modular system with multiple recognition and effector mechanisms, built-in redundancies and 'backup pathways,' allowing the subtle distinction between specific types and sites of DNA damage (Friedberg *et al*., 1995; Wood, 1996; Rajewsky *et al*., 1998).

Human examples of defects in DNA repair genes associated with a dramatically increased risk of skin cancer or colorectal cancer, are the familial genetic disorders Xeroderma pigmentosum (XP) and hereditary nonpolyposis colorectal cancer (HNPCC) respectively.

To enhance the power of the cells' response to DNA damage, the DNA repair machinery joins forces with the molecular controls of cell cycle progression and DNA replication under the surveillance of checkpoint proteins and pathways (Weinert, 1998). In addition to being involved in the control of the transcriptional activity of DNA repair genes, checkpoint proteins signal the arrest of proliferating cells in the G_1 phase of the cell cycle, thus providing more time for DNA repair to avoid the replication of damaged DNA in S phase. Cells may become arrested in mid-G_1 or just prior to the transition of cells from G_1 to S phase. Arrest may also occur in G_2, presumably to avoid the segregation of damaged chromosomes in mitosis (M). In kinetic terms, G_1 arrest in response to DNA damage is at least two-componential (Agami and Bernards, 2000). An initial rapid induction phase is mediated by the proteolytic degradation of cyclin D1, a very important regulatory protein responsible for cell cycle progression through G_1 and often overexpressed in malignant cells. Thereafter, G_1 arrest is maintained through increased stability of the transcription factor p53, the tumour-suppressor protein most frequently inactivated in human cancers.

The p53 protein also promotes DNA repair through the transcriptional induction of target genes such as *p53R2*, which encodes a ribonucleotide reductase catalysing the

conversion of ribonucleoside diphosphates to the deoxyribonucleotides required in major DNA repair reactions (Tanaka *et al*., 2000). Unprotecting cells through inactivation of p53 therefore increases their sensitivity toward DNA-reactive drugs (see Cancer Therapy Resistance). If the extent of DNA damage surpasses a critical, not yet well-specified level, checkpoint proteins such as p53 direct cells into a suicidal apoptotic pathway rather than attempting to accomplish recovery by DNA repair in vain.

A simple scheme of major DNA repair pathways is shown in **Figure 2**. Two principal modes of DNA repair are distinguished: (1) one-step repair (OSR) (i.e. the direct reversal of DNA damage) and (2) excision repair, including nucleotide excision repair (NER), base excision repair (BER) and mismatch repair (MMR) (see Jiricny, 1998, and the next section). Not included in this scheme are (1) the repair of DNA double-strand breaks (induced, e.g., by ionizing radiation, ROS or anticancer agents) by homologous recombination or nonhomologous DNA end joining, and (2) the bypass of DNA lesions by low-fidelity DNA polymerases, a potentially very important error-prone (i.e. mutagenic) repair process, the molecular mechanisms of which have not yet been fully clarified. In a recent model of mutagenic lesion bypass, polymerase ι first misincorporates a deoxynucleotide opposite a helix-distorting lesion or abasic site in DNA, and polymerase ζ then extends from the mispair to complete the bypass (Johnson *et al*., 2000).

OSR is performed, for example, by the repair protein O^6-alkylguanine-DNA alkyltransferase (MGMT) through direct removal of an alkyl group from the O^6-atom of guanine in the DNA of cells exposed to alkylating agents (Pegg, 1990). With increasing size of the alkyl group ($\geq$ethyl), the relative contribution of MGMT to the repair of O^6-alkylguanines in DNA decreases and excision repair steps in as a backup modality.

NER, BER and MMR are multistep pathways involving different proteins and protein complexes. Acting upon a

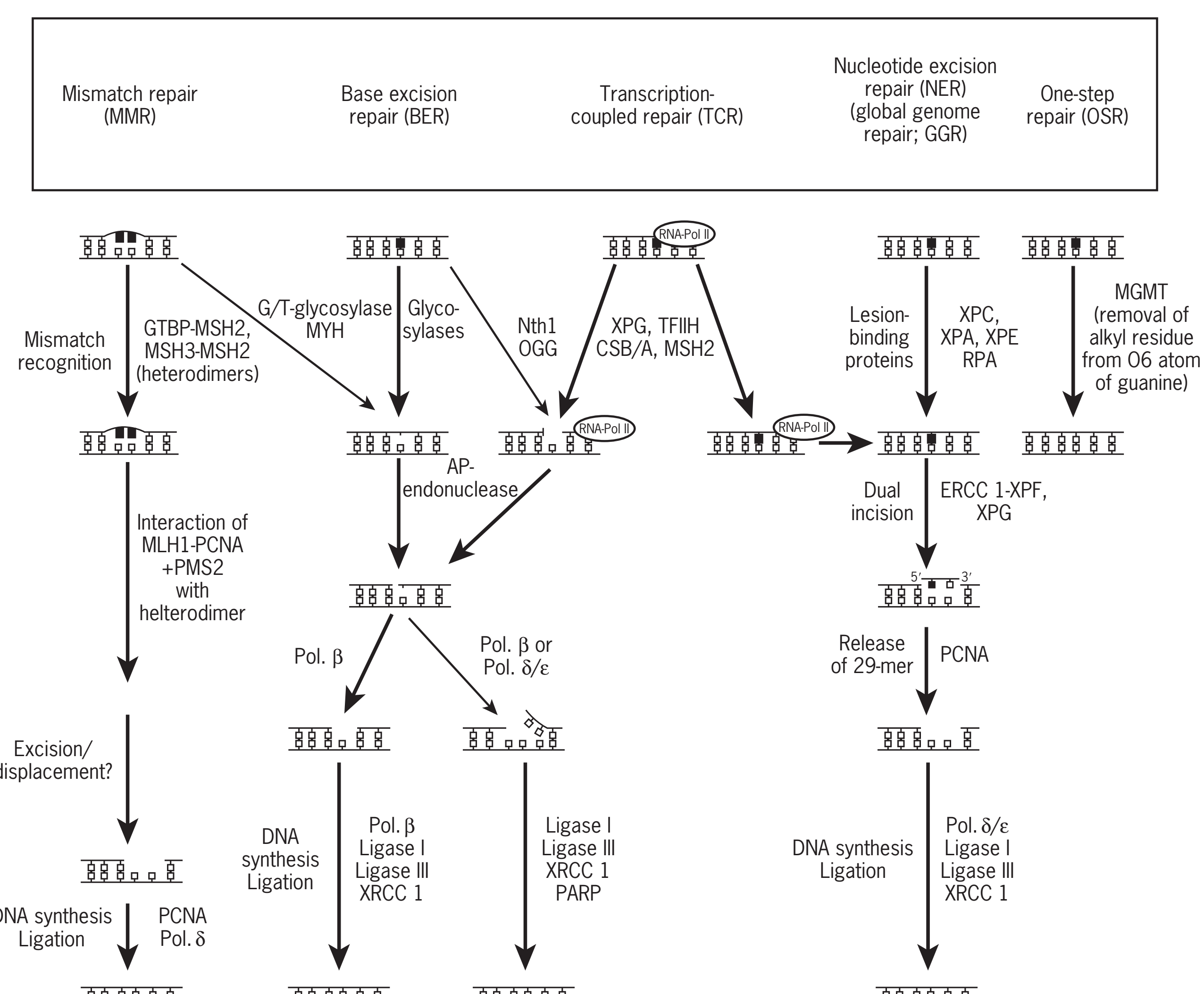


Figure 2 DNA repair pathways in mammalian cells.

broad spectrum of structurally diverse, mostly 'bulky' DNA lesions, NER is a highly versatile pathway involving the concerted action of about 30 different proteins. The human excision nuclease (comprising six repair factors) removes a 24–32 nucleotides long oligomer from the damaged DNA strand after incision on either side of the lesion, followed by DNA repair synthesis (by DNA polymerases δ and ε) using the intact strand as a template, and closing of the gap by ligation. Compared with BER, there appears to be much less redundancy in the NER pathway. So-called 'global' or 'overall' NER (GGR) operates on genomic DNA independent of gene transcription; certain NER proteins may also participate in genetic recombination (as probably required in the repair of DNA interstrand cross-links). Depending on DNA sequence context and chromatin structure, extensive positional heterogeneity regarding the efficiency of NER within defined genes has been found by analyses at very high or even single-nucleotide resolution.

A discrete pathway, transcription-coupled repair (TCR), is responsible for the rapid removal of DNA lesions that block transcription by RNA polymerase. TCR has been found to operate much faster than GGR, usually, but not in all cases, with a bias in favour of the transcribed strand (Le Page *et al*., 2000).

In BER, the removal of a single modified base from one DNA strand is performed by DNA glycosylases via hydrolytic cleavage of the *N*-glycosylic bond. Some of these glycosylases exhibit pronounced lesion specificity, others recognize multiple, structurally different damaged bases and certain glycosylases display considerable promiscuity in terms of the range of reactions that they catalyse. The apurinic/apyrimidinic (AP) site left behind after cleavage of the *N*-glycosylic bond is hydrolysed 5′ by an AP endonuclease, and the 5′-deoxyribose phosphate is excised by a phosphodiesterase. The resulting single-nucleotide gap is then filled by polymerase β and ligated. Alternatives to this common BER pathway include the excision of a short oligonucleotide patch containing the AP site and filling of the gap by polymerase δ or ε. Base alterations caused by a large variety of agents and processes (e.g. spontaneous deamination, ionizing radiation, ROS, alkylating agents, DNA replication errors) are processed by BER. A common feature of the DNA lesions recognized by BER glycosylases is that they do not significantly distort the DNA helix.

Sensitive analytical methodology has been developed for detecting DNA damage and for determining DNA repair kinetics in human cells (Pfeifer, 1996). Notably, these techniques include (1) the radiolabelling and subsequent radiochromatographic analysis of DNA adducts as a function of time after their formation (^{32}P-postlabelling) and (2) immunoanalytical methods, i.e. the application of poly- or monoclonal antibodies to quantify distinct DNA lesions in DNA isolates from cell or tissue samples, in individual cells and in individual genes.

DNA Repair and Cancer Therapy

Cancer Therapy Resistance

Most present-day anticancer drugs and ionizing radiation interact with target cell DNA and exert their cytotoxic effects preferentially in replicating cells. In addition to the primary DNA lesions, secondary DNA alterations induced in the course of repair processes also contribute to the cytotoxic effects of DNA-reactive agents. In principle, DNA repair (except for MMR; see below) synergizes with other protective mechanisms used by cancer cells to outwit cytotoxic therapy (Chaney and Sancar, 1996). Cancer cells which efficiently carry out a mode of repair specifically required to counteract a particular anticancer agent may thus survive, and actually be selected for, in the course of therapy.

Cancer therapy resistance is multifactorial and based on the exceptional ability of malignant cells to adapt to altered microenvironmental conditions, including their exposure to cytotoxic agents. That DNA repair ranks high among the defence mechanisms of cancer cells becomes particularly impressive under laboratory conditions, when only in repair-defective target cells, and not in their repair-proficient counterparts, a given DNA-damaging drug is able to trigger apoptosis or other forms of cell death.

Notably, resistance of cancer cells to the widely used chloroethylnitrosoureas (e.g. *N*,*N*′-bis(2-chloroethyl)-*N*-nitrosourea (BCNU)) has proven to be positively correlated with their MGMT activity (more precisely, the size of the cellular MGMT pool and the rate of MGMT biosynthesis). In the formation of cytotoxic interstrand cross-links in target cell DNA, chloroethylnitrosoureas initially alkylate the O^6 atom of guanine. The resulting O^6-alkylguanine is a substrate of one-step repair by MGMT. Cellular MGMT levels have been found to be highly variable in different types of human tumours.

A notable exception to the rule that DNA repair enhances therapy resistance is MMR, the primary function of which is to correct mismatches generated by nucleotide misinsertion during semiconservative DNA replication or the bypassing of DNA lesions by low-fidelity DNA polymerases (see earlier). Moreover, MMR is involved in the initiation of p53-mediated and -independent apoptosis in response to excessive DNA damage.

In general, therefore, cells defective in MMR are hypermutable, their genome becomes destabilized and the probability of mutations in genes critically associated with carcinogenesis is enhanced. MMR-defective cancer cells, however, such as *p53* mutants, exhibit increased resistance to various anticancer drugs, i.e. these cells tolerate even high loads of DNA damage caused by alkylating agents such as temozolomide and procarbazine, or by cisplatin. This complicates the therapeutic situation, because MMR-deficient tumour cells will be at a selective advantage relative to MMR-proficient normal cells of the host. In the

case of germ-line defects of MMR genes, as exemplified by HNPCC, patients would be at higher risk of developing secondary cancers owing to the hypermutability of their cells.

Similar observations are made in an animal model of toxicity and carcinogenesis induced by a DNA methylating agent: while mice with an inactivating germ-line mutation of the *MGMT* gene ($MGMT^{-/-}$) are hypersensitive toward the lethal cytotoxicity of the agent, their sensitivity reverts to the wild-type ($MGMT^{+/+}$) level when the *MGMT* 'knockout' is combined with the inactivation of the MMR gene *MLH1*. At the same time, the highly elevated cancer susceptibility of $MGMT^{-/-}$ mice remains unchanged in the $MGMT^{-/-}MLH1^{-/-}$ 'double knockout' animals (Kawate *et al.*, 1998).

Individualized Cellular DNA Repair Profiles

As pointed out above, diverse genetic alterations and changes in signal transduction pathways and metabolism accumulate in the course of carcinogenesis and cancer progression. In principle, therefore, and if properly diagnosed, distinct molecular traits of malignant cells, including defects in specific modes of DNA repair, may be exploited to achieve more selective targeting by anticancer drugs.

Different DNA-reactive anticancer agents produce, or may be designed to produce, structurally distinct DNA lesions. These, in turn, require processing by different modes of DNA repair to protect cells from cytotoxic damage. Pretherapeutic determination of the DNA repair profiles of individual cancers (complemented by appropriate controls of the patient's normal cells) may thus permit us to specify the repair characteristics of the target cells as a basis for choosing the most effective drug or drug combination. Criteria to be used for the choice of drug would be (1) that the repair pathway responsible for processing the major cytotoxic DNA lesion produced by the drug shows the greatest possible differential between the cancer cells (low or missing activity) and critical normal cells (regular activity), and (2) that the chemical nature of the major cytotoxic DNA lesion is such that there is the least possible redundancy between different pathways for its repair. Novel or structurally modified drugs that meet the latter criterion better than existing agents may still need to be designed based on molecular recognition and mechanics in human DNA repair.

Cellular DNA repair profiling of individual cancers may be performed in two ways.

1. Via DNA damage dosimetry in primary cancer cells (leukaemic cells; tumour cells microdissected from surgical/biopsy specimens) as a function of time after exposure to model drugs *in vitro*, i.e. by direct measurement of the kinetics of repair of a panel of specific DNA lesions selected for being processed by distinct DNA repair pathways (Buschfort *et al.*, 1997). As an example, **Figure 3** shows the repair kinetics of a model DNA alkylation product (O^6-alkylguanine), as determined in normal and leukaemic lymphocytes isolated from three patients with chronic lymphatic leukaemia (CLL) by monoclonal antibody-based immunoanalysis of individual CLL cells.
2. By 'gene expression profiling' using gene chip (cDNA arrays) and differential display technology for comprehensive analysis of the RNA transcripts of a representative panel of human DNA repair genes, including repair-associated 'checkpoint,' cell cycle and apoptosis regulatory genes. This approach differs from the direct measurement of repair rates in the DNA of individual cells mainly in that repair gene expression is analysed at the level of RNA transcripts (not yet, unfortunately, at the level of functional repair proteins). Hence the gene expression data obtained do not readily translate into repair rates for specific DNA lesions. Moreover, the analyses require large numbers of cells and can therefore only provide overall averages for potentially heterogeneous cell populations. Nevertheless, this

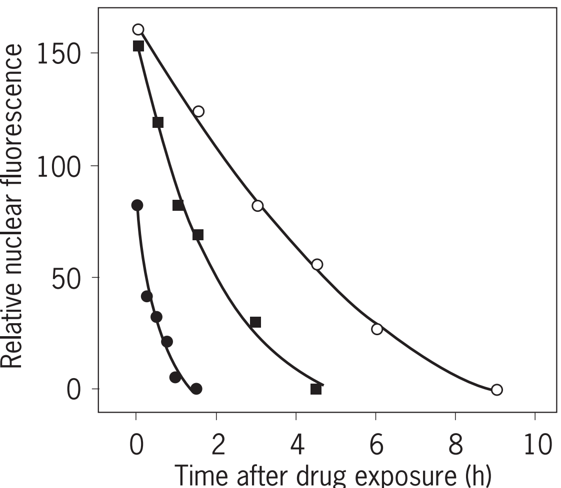


Figure 3 Pretherapeutic cellular DNA repair profiling in human chronic lymphatic leukaemia (CLL): repair kinetics of a model cytotoxic DNA alkylation product (O^6-alkylguanine; O^6-AG) in the nuclear DNA of primary CLL lymphocytes determined by monoclonal antibody-based immunocytological analysis. [O^6-AG]-specific nuclear immunofluorescence quantified via digital imaging of electronically intensified fluorescence signals (data points, mean values for 100 individual cells). Standard input of O^6-AG into DNA generated by pulse exposure of cells to *N*-ethyl-*N*-nitrosourea *in vitro*. Open circles, CLL that subsequently proved to be sensitive to clinical treatment with alkylating drugs (slow repair of O^6-AG in DNA); filled squares, CLL with intermediate drug sensitivity; filled circles, CLL clinically resistant to alkylating drugs (very fast repair). (From Buschfort *et al.*, unpublished work.)

technology permits the simultaneous, semiquantitative evaluation of the expression levels of multiple genes and may thus complement the direct analysis of DNA repair kinetics in a powerful way.

Cotherapeutic Inhibition of Drug-specific DNA Repair Mechanisms

An important objective is the development of specific inhibitors for distinct modes of DNA repair, to be used as 'cotherapeutic drugs' accompanying individualized cancer therapy (see the previous section). The application of repair inhibitors would appear particularly advantageous when the DNA repair pathway to be temporarily inactivated represents the 'backup' pathway for a primary pathway, or vice versa, that is defective or down-regulated in the tumour cells, contrary to normal host cells requiring protection.

The potential toxicity of repair inhibitors would necessitate careful evaluation regarding undesired side effects before they could be clinically applied. The only selective inhibitor of a distinct mode of DNA repair thus far available, and introduced to clinical testing, is the MGMT inhibitor O^6-benzylguanine. However, current advances in the use of high-throughput methodology for drug discovery should facilitate the search for further inhibitors specifically interfering with other DNA repair pathways.

Selective Protection of Haematopoietic Cells from DNA Damage-mediated Cytotoxicity Through Transgenic Enhancement of Repair

As an alternative to the above strategies attempting circumvention or temporary inhibition of DNA repair, it is also possible to exploit enhanced cellular DNA repair capacity to achieve more efficacious cancer therapy. Instead of focusing on the cancer cells, however, such an approach must be directed towards the protection of those normal cells that are of critical importance for the survival of the host.

Owing to their high drug sensitivity, the haematopoietic cells of the bone marrow constitute a major, if not the main, limiting factor in cancer chemotherapy. Normal bone marrow cells express the *MGMT* gene at an exceedingly low level, i.e. these cells are almost defenceless when exposed to drugs that exert their cytotoxic effect via the formation of O^6-alkylguanine in DNA (e.g. the chloroethylnitrosoureas). With the use of viral vectors, however, the *MGMT* gene can be transduced and expressed in haematopoietic progenitor cells *ex vivo* prior to their re-transfer into the host. When a mutant *MGMT* gene insensitive to the inhibitory effect of O^6-benzylguanine is applied, the transduced cells will remain protected against the cytotoxic action of chloroethylnitrosoureas even in the presence of the MGMT inhibitor which, at the same time, will sensitize cancer cells expressing the *MGMT* gene to these drugs (Reese *et al*., 1996).

THE MAMMALIAN CELL CYCLE

Basic Regulatory Mechanisms

Our understanding of the molecular mechanisms, pathways and molecules regulating cell proliferation has grown considerably in recent years and provides a new basis for rational approaches to cancer therapy. In this section we will give an introduction to the key regulatory processes involved in cell cycle control and show how this knowledge can be applied to the design of new therapeutic strategies.

Cell cycle progression in mammalian cells is controlled through fundamentally different regulatory pathways that are controlled either by the cell's microenvironment or by intrinsic checkpoints. Progression through G_1 across the restriction point (R-point; see **Figure 4**) is controlled by external signals which are transmitted, for example, by mitogens or through cell adhesion processes. Beyond this point, cell cycle progression is governed by a genetic programme that is largely independent of extracellular signals but subject to internally controlled checkpoints. These checkpoints ensure proper DNA replication, DNA integrity and mitotic cell division. A central role in cell cycle progression is exerted by the cyclin-dependent kinases (**Figure 4**) (Pavletich, 1999), which are composed of a regulatory cyclin subunit (e.g. cyclin A, B1, B2, D1, D2, D3 or E) and a catalytic kinase subunit (e.g. Cdk1/Cdc2, Cdk2, Cdk4 or Cdk6).

Cyclin-dependent Kinases

The activity of Cdk–cyclin complexes is controlled by multiple mechanisms, including Cdk phosphorylation, the regulation of cyclin mRNA expression and protein stability and the association of Cdk–cyclin complexes with specific protein inhibitors. These regulatory mechanisms are briefly reviewed below.

Regulation of Cyclins by Phosphorylation and Dephosphorylation

Two domains of the Cdk subunits are critical with respect to their regulation by phosphorylation. Whereas phosphorylation of Thr161 is required for kinase activity, the phosphorylation of Thr14 and Tyr15 is inhibitory. Thr161 is dephosphorylated during mitosis and becomes phosphorylated during the subsequent cell cycle by cyclin-activating kinase (CAK), composed of cyclin H and Cdk7. Thr14 and Tyr15 are phosphorylated by the dual-specificity kinase WEE1 and probably other kinases and are

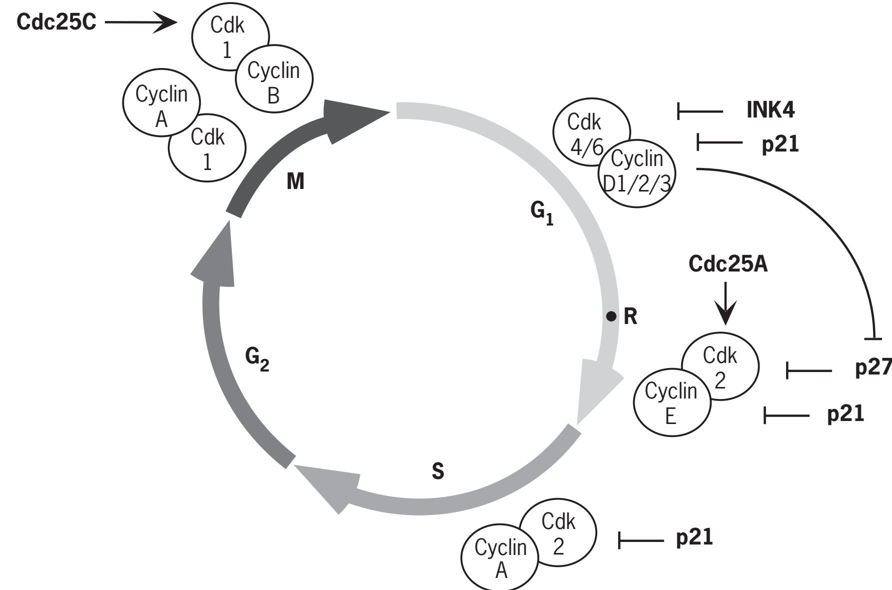


Figure 4 Phase-specific function of cyclin-dependent kinases (Cdk-cyclin complexes), their activating phosphatases (Cdc25A and C), and the Cdk inhibitors (INK4 family members, p21 and p27) during the mammalian cell cycle.

dephosphorylated by different Cdc25 phosphatases acting at different stages of the cell cycle (**Figure 4**).

Phase-specific Expression of Cyclins

Most of the regulatory Cdk subunits, the cyclins, are expressed in a phase-specific manner. Thus, the cyclin E gene is switched on in late G_1, the cyclin A gene at the G_1/S transition and the cyclin B genes in late S. In contrast, transcription of cyclin D genes fluctuates only marginally during a regular cell cycle (in contrast to mitogen-stimulated G_0 cells). Likewise, expression of most Cdk genes is not cell cycle regulated in normally cycling cells.

Phase-specific Proteolysis of Cyclins

The probably most important mechanism regulating the steady-state levels of cyclin A, B and E during the cell cycle is their periodic proteolysis. Following phase-specific ubiquitination by a complex and intricately regulated enzyme system, the modified proteins are degraded in the proteasome. The process is initiated by the attachment of ubiquitin to an ubiquitin-activating enzyme (E1) which is followed by its transfer to an ubiquitin-conjugating enzyme (E2). Subsequently, the ubiquitin moiety is transferred either directly or in conjunction with an ubiquitin ligase (E3) to the target protein. After linkage of additional ubiquitin molecules the polyubiquitinated protein is targeted for degradation. During the cell cycle, E3 complexes are instrumental with respect to the ubiquitination of cyclins and also other proteins (see below), and their subsequent degradation, thus allowing the transition from one phase to the next. One of the best studied examples in this context is the anaphase-promoting complex or cyclosome (APC/C). APC/C is an E3 complex that is responsible for the ubiquitination and subsequent degradation of cyclin B and anaphase inhibitors prior to completion of M phase, thereby controlling chromosome segregation and mitotic exit.

Cdk Inhibitors

Based on their structure, the Cdk inhibitors (CKIs) identified to date can be grouped into two different families: the INK4 family comprising $p15^{INK4b}$, $p16^{INK4a}$, $p18^{INK4c}$ and $p19^{INK4d}$, and the KIP/CIP family consisting of $p27^{KIP-1}$, $p57^{KIP-2}$ and $p21^{CIP-1}$. Of these, the INK4 proteins p15 and p16 and the KIP/CIP proteins p27 and p21 appear to play a major role in human oncogenesis owing to frequent mutations (INK4), deregulated expression (p27) or induction in response to chemo- or radiotherapy (p21). Whereas p21 is an inhibitor of multiple Cdks, the INK family members specifically inhibit cyclin D kinases, and the target of p27 is primarily Cdk2–cyclin E (**Figure 4**). p27 also associates with cyclin D kinases, but does not inhibit their enzymatic activity. Cyclin D complexes can thereby sequester and inactivate p27 (**Figure 4**), resulting in the upregulation of cyclin E kinase activity.

The CKIs are involved in different pathways controlling cell cycle progression, and are themselves regulated by different mechanisms. p21 plays an essential role in the cell's response to certain types of stress, such as DNA damage or metabolic perturbations, which induce a dramatic increase in *p21* gene transcription. This up-regulation of

the *p21* gene is brought about to a large extent by the tumour-suppressor and transcriptional activator p53, which is activated by DNA damage (discussed in detail below) and directly binds to cognate sites in the *p21* gene promoter. p27 levels, on the other hand, are up-regulated under unfavourable growth conditions, such as high cell density or lack of mitogens, to keep Cdk activity low in G_0 cells. However, as soon as cells progress though G_1, p27 is phosphorylated by Cdk2–cyclin E, which has been proposed to serve as a recognition signal for polyubiquitination and subsequent degradation by the proteasome. INK4 proteins seem to play a major role in cellular senescence where their expression is dramatically upregulated. In the case of p15, gene expression is blocked by the Myc oncoprotein, which appears to be an important step in Myc-induced immortalization.

Restriction Point Control

The G_1 Cdk–cyclin complexes regulate cell cycle progression across the R-point through phosphorylation of the retinoblastoma protein Rb and its kins p107 and p130 (Harbour and Dean, 2000). With respect to tumorigenesis, Rb seems to be the most relevant family member, since only Rb has been found to be mutated or deleted in human tumours and only the targeted disruption of the *Rb* gene predisposes mice to the development of tumours. In early–mid-G_1 the transcription factor E2F is found in complexes with Rb and chromatin remodelling factors, including histone deacetylase (HDAC) (**Figure 5**). These complexes actively repress transcription via E2F binding sites in the respective target genes. The phosphorylation of the E2F–Rb–HDAC complexes in mid-G_1 by cyclin D–Cdk4/6 disrupts the binding of Rb and HDAC. This permits the subsequent hyperphosphorylation of Rb by cyclin E–Cdk in late G_1 and disruption of the E2F–Rb complexes (**Figure 5**). This leads to the generation of transcriptionally active free E2F heterodimers, and consequently the induction of numerous E2F target genes.

The relevance of R-point control for tumorigenesis is emphasized by the fact that the Rb pathway is defective in basically all human tumours owing to direct genetic alterations or the indirect deregulation of its components as a consequence of mutations affecting other genes (asterisks in **Figure 5**). Genetic alterations directly affecting the Rb pathway include (1) loss or inactivation of Rb itself, (2) amplification or translocation-mediated deregulation of the cyclin D1 gene, (3) mutations of Cdk4 rendering it resistant to inhibition by INK4 proteins and (4) deletion or mutation of INK4 proteins, such as p15, p16 or p18. Most

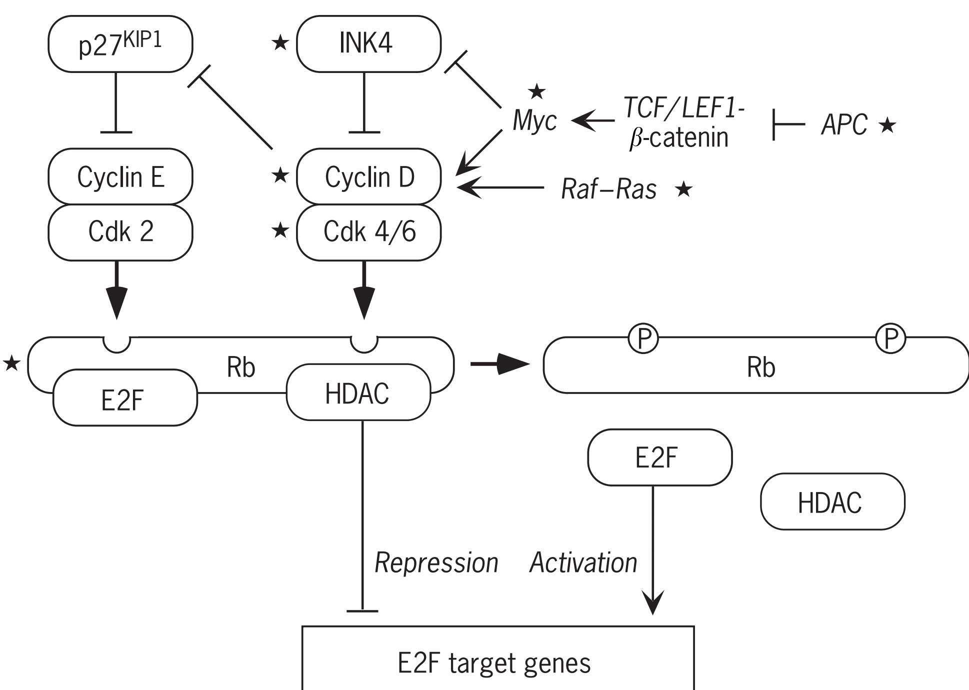


Figure 5 The Rb–E2F pathway and its regulation by cyclin-dependent kinases (cyclin D–Cdk4/6 and cyclin E–Cdk2) and their inhibitors (p27 and the INK4 family members p15, 16 and p18). The repressor function of the E2F–Rb complex is dependent on the association of Rb with a histone deacetylase (HDAC). The latter interaction is disrupted by cyclin D–Cdk4/6-mediated phosphorylation (P). This is followed by cyclin E–Cdk2-mediated phosphorylation of Rb, allowing the release of free, transcriptionally active E2F heterodimers. Also shown are other pathways (Myc and Ras–Raf) that are able to modulate the Rb–E2F pathway. APC: adenomatous polyposis coli tumour-suppressor protein. Asterisks denote frequent genetic alterations in human cancers.

human tumours show one of these genetic alterations. However, in those tumours that do not fall into this category, such as colon carcinoma, the Rb pathway is deregulated as a consequence of other genetic alterations. For example, the Myc oncoprotein (Eilers, 1999) can activate the cyclin D2 promoter through direct transcriptional activation, and can also repress the *p15* gene, and activated Ras and Raf can induce the cyclin D1 promoter via a MAP kinase pathway. Myc expression is frequently upregulated not only because of gene amplification or translocations, but also as a consequence of genetic alterations in other loci. In certain forms of colon carcinoma, for instance, the adenomatous polyposis coli tumour suppressor APC is lost and can therefore no longer constrain the transcriptional activator complex β-catenin–TCF/LEF-1, a potent inducer of the *myc* gene .

In view of its major role in human tumorigenesis, the Rb pathway and its regulators are obviously of major interest with respect to therapeutic intervention, as will be discussed in further detail below.

Checkpoint Control

DNA repair and cell cycle progression are linked via checkpoints that are activated in response to DNA damage and arrest the cell cycle in specific phases. Of particular importance in this context are the p53- and Cdc25A-governed checkpoints in G_1 (Mailand *et al*., 2000), the G_2 checkpoint controlled by Cdc25C and the mitotic APC/C-dependent checkpoint (**Figure 6**). Other checkpoints have been described, in particular in M phase (Muhua *et al*., 1998; Scolnick and Halazonetis, 2000), but their function and regulation are less well understood.

Linking DNA Damage and Checkpoint Control

Sites of damaged DNA are recognized by specific proteins that initiate the cell's response, such as DNA repair, cell cycle arrest or apoptosis (Rotman and Shiloh, 1999). Three protein kinases have been implicated in this step: the DNA-dependent protein kinase (DNA-PK), the ataxia telangiectasia mutated (ATM) kinase and the ATM-related (ATR) kinase. DNA-PK plays a critical role in mammalian DNA double-strand break repair and recombination events, such as immunoglobulin gene rearrangements. Its regulatory subunit Ku, a dimer of the Ku70 and Ku80 proteins, binds to DNA and recruits the catalytic subunit, DNA-PKcs. In cells lacking DNA-PK activity owing to a targeted genetic disruption, the p53-controlled G_1 checkpoint is intact, suggesting that the ATM and ATR (and perhaps other) kinases play a more important role in initiating checkpoint control in response to DNA damage. Subsequent to their activation, ATR and ATM phosphorylate the checkpoint kinase Chk1 and Chk2, respectively (**Figure 7**). The reason for the existence of two parallel pathways is not entirely clear, but may be related to a mediator function in response to different kinds of DNA damage. In the final step of the signalling cascade, Chk1

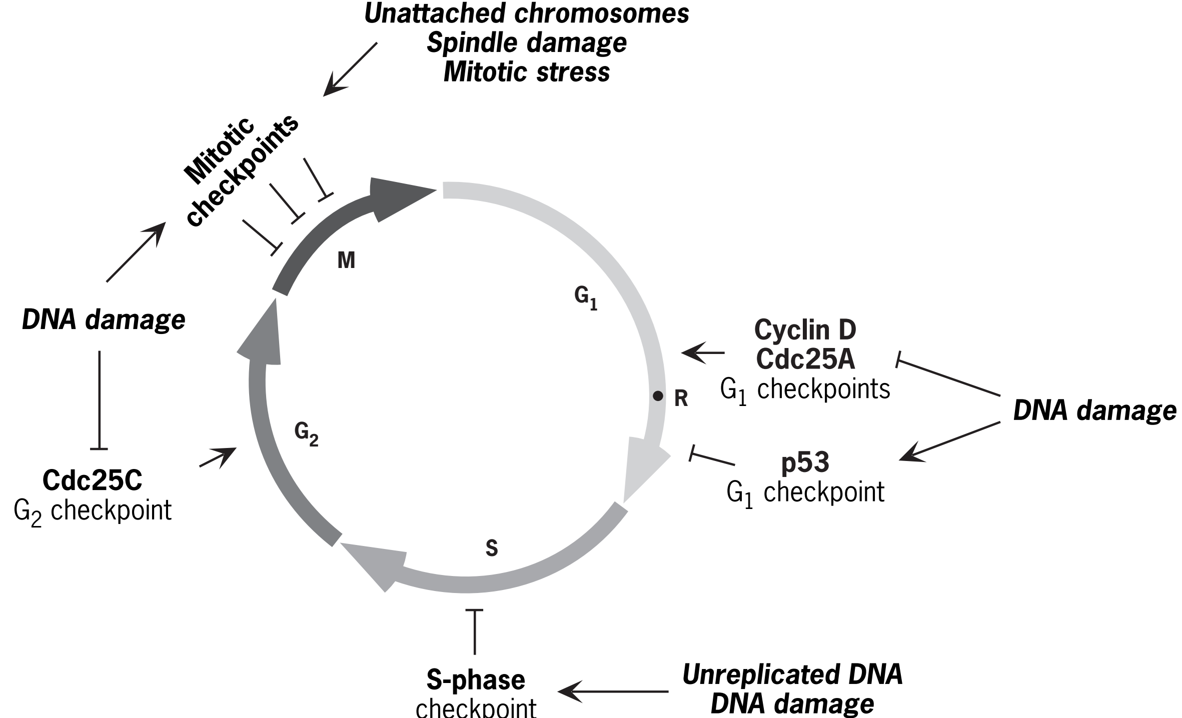


Figure 6 Checkpoints monitoring the integrity of the genome during cell cycle progression.

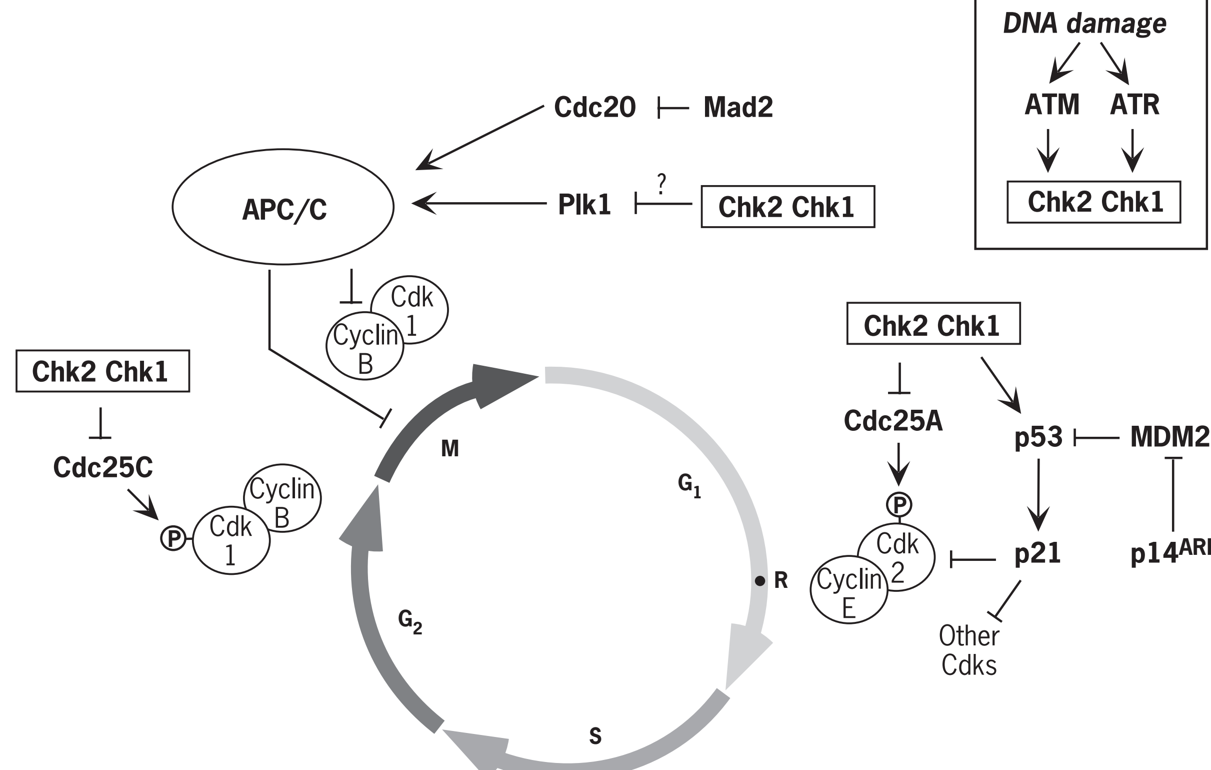


Figure 7 Regulation of the G_1 (p53 and Cdc25A), G_2 (Cdc25C) and spindle (APC/C) checkpoints in mammalian cells. APC/C: anaphase-promoting complex/cyclosome.

and Chk2 phosphorylate their target proteins, such as p53, Cdc25A, Cdc25C and other protein kinases (see below), resulting in the modulation of their activity (**Figure 7**).

p53-Dependent Checkpoint Control in G_1

A major role in checkpoint control in the G_1 phase of the cell cycle is exerted by the p53 tumour-suppressor pathway (**Figures 6** and **7**). In response to DNA damage, steady-state levels of p53 rise owing to its phosphorylation-mediated stabilization affected by the checkpoint kinases Chk1 and Chk2 and presumably others. The steady-state level of p53 is regulated by MDM2, an oncoprotein that associates with unphosphorylated p53 and targets it for ubiquitin-mediated degradation. MDM2 itself is targeted for proteolysis by the tumour suppressor $p14^{ARF}$ (or mouse $p19^{ARF}$). As a consequence of phosphorylation, p53 can therefore accumulate to high levels and activate the G_1 checkpoint via the transcriptional induction of the Cdk inhibitor p21. These observations indicate that p53 is regulated by a complex network of regulatory pathways that affect its function via different mechanisms.

p53 is not only capable of inducing a cell cycle arrest, but can also trigger apoptosis. The latter occurs, for example, in case of severe DNA damage, but is also induced by the untimely or deregulated expression of oncoproteins or cell cycle regulators, such as Myc and E2F1. Therefore, to suppress the proapoptotic potential of a deregulated Rb–E2F pathway, tumour cells acquire potent antiapoptotic mechanisms, among them the elimination of p53 function due to the loss of p53 itself or alteration of its regulators, such as MDM2 or $p14^{ARF}$. It should be noted, however, that Myc and E2F1 can induce apoptosis also via p53-independent pathways, so that other antiapoptotic mechanisms also play an essential role in suppressing their proapoptotic function. Nevertheless, the importance of deleting p53 function during tumour progression is clearly illustrated by the fact that the p53 pathway is defective in more than 50% of all human malignancies. This also emphasizes its relevance with respect to the development of new anticancer therapies.

Cdc25A-dependent Checkpoint Control in G_1

Another G_1 checkpoint activated in response to DNA damage is governed by Cdc25A, a protein phosphatase that is essential for the activation of cyclin E kinase activity prior to progression past the R point (**Figures 6** and **7**). Cdc25A itself is another target of the ATM–Chk2 and ATR–Chk1 pathways. The phosphorylation of Cdc25A leads to its ubiquitin-mediated proteolytic degradation with the consequence that Cdk2 is not dephosphorylated at Thr14 and Tyr15, which leaves the Cdk2–cyclin E complex in an inactive state (Mailand *et al*., 2000). In addition,

DNA damage can induce the proteolytic degradation of cyclin D (Agami and Bernards, 2000), thus preventing hyperphosphorylation of Rb. Each of these events is sufficient to arrest cell cycle progression at the R-point, so that DNA repair can occur prior to the onset of DNA replication. Since many anticancer drugs exert their function through DNA damage, this checkpoint may have a negative impact on their efficacy and therefore represent a potential target for the development of new drugs.

The G_2 Checkpoint

The same DNA damage-induced signalling pathway leading to Cdc25A phosphorylation in G_1 also mediates the phosphorylation of Cdc25C in G_2 (**Figures 6** and **7**). Cdc25C is instrumental in dephosphorylating and thereby activating Cdk1 (Cdc2) in a nuclear complex with cyclin B, a prerequisite for M-phase entry. The Chk1/Chk2-mediated phosphorylation of Cdc25C results in its association with a p53-induced specific isoform of the 14-3-3 protein and thus in its retention in the cytoplasm. This renders Cdc25C inactive, so that progression into mitosis is prevented until DNA repair is completed. By analogy with Cdc25-governed G_1 checkpoint, the G_2 checkpoint also seems to be a suitable target for therapeutic intervention with respect to improving the efficacy of DNA-damaging agents.

Mitotic Checkpoints

Multiple checkpoints operate during mitosis (**Figures 6** and **7**) to ensure that chromosomes segregate with maximum fidelity, to warrant genomic stability and to prevent premature cytokinesis. Mitotic checkpoints have been described to monitor prophase-to-metaphase, metaphase-to-anaphase and anaphase-to-telophase transitions. The checkpoint governing the prophase-to-metaphase transition is, for instance, activated by mitotic stress, such as drugs interfering with microtubule polymerization or function. Activation of this checkpoint is dependent on a gene termed *chfr*, which is frequently mutated in human cancer cells (Scolnick and Halazonetis, 2000). The checkpoint monitoring the anaphase-to-telophase transition is activated by misaligned spindles and delays cytokinesis. In yeast, this checkpoint requires a microtubule-associated protein that is a homologue of human EB1, a protein interacting with the adenomatous polyposis coli tumour-suppressor gene product (Muhua *et al*., 1998).

The best understood mitotic checkpoint operates at the metaphase-to-anaphase boundary, and is also referred to as the spindle assembly checkpoint (**Figure 7**). The target of this checkpoint is the E3 ubiquitin ligase APC/C. The checkpoint is activated early in mitosis unless all kinetochores are attached to microtubules of the mitotic spindle, or in response to microtubule-damaging agents such as taxol. Once activated, the checkpoint induces a signalling cascade involving several checkpoint kinases, eventually leading to the oligomerization of Mad2 at kinetochores. These Mad2 tetramers dissociate from kinetochores and inhibit APC/C activity through binding to its positive regulatory subunit cdc20. The inhibition of APC/C by Mad2 results in the accumulation of proteins whose destruction is required for sister chromatid separation (such as the human oncogene securin/PTTG), thus invoking an arrest in metaphase.

More recently it has been recognized that the checkpoint monitored by the APC/C is also activated in response to DNA damage (Smits *et al*., 2000). A major mediator in this process appears to be the Polo-like kinase Plk1 known as an activator of the APC/C. Plk1 itself is a substrate for the checkpoint kinases Chk1/2, and thus another downstream target of the pathways activating the p53 and Cdc25 checkpoints (**Figure 7**). Phosphorylation of Plk1 results in its inhibition as a consequence of DNA damage in G_2 or M phase, and consequently in a mitotic block due to low APC/C activity. These observations suggest that Plk1 is an important activator of DNA damage-mediated checkpoint control, and thus a potentially interesting therapeutic target.

Application to New Therapeutic Approaches

Cdk Inhibitors

As cancer is mainly a proliferative disease, the inhibition of specific proteins driving the cell cycle is an obvious strategy for the rational discovery of new anticancer drugs. In this context, it is of particular interest that the interference with coordinated cell cycle progression can result in apoptosis of tumour cells, which led to the definition of a new class of antitumour agents that function through a direct inhibition of proteins driving the cell cycle. One of the prototypes of this class of compounds is the synthetic flavone flavopiridol (Senderowicz, 1999). Flavopiridol is a general inhibitor of Cdks, induces cell cycle arrest and apoptosis and is not influenced by many of the genetic alterations conferring resistance on human tumour cells. Accordingly, flavopiridol has shown promising tumour responses in preclinical models and is currently undergoing clinical trials. Numerous other chemical Cdk inhibitors have recently been identified and are currently being evaluated for their antitumour properties (Sielecki *et al*., 2000). It should be noted, however, that these compounds may have other unknown activities that might add to their antitumour properties. Thus, the results obtained with flavopiridol cannot be exclusively attributed to the inhibition of Cdks. In this context, the induction of Bcl-2-independent mitochondrial depolarization is noteworthy, as it may infer a direct impact on mitochondria (Achenbach *et al*., 2000). Likewise, the flavopiridol-induced sensitization of tumour cells to taxol (Bible and Kaufmann, 1997) is tantalizing since it may suggest that flavopiridol interferes with a mitotic checkpoint.

Manipulation of Checkpoint Control

Other interesting targets for therapeutic intervention are the proteins governing checkpoint control, e.g. in response to DNA damage. As discussed above, checkpoint control can invoke a transient cell cycle block, but can also trigger apoptosis. Both types of checkpoints are relevant to tumour therapy. While the functionality of an apoptosis-inducing mechanism in response to drug- or radiation-induced cellular damage is desirable, checkpoint control leading to a mere cell cycle arrest is counterproductive for any therapy that relies on cell proliferation, such as radiation or conventional chemotherapy.

The p53 checkpoint is lost in many tumour cells, and thus the ability to undergo apoptosis in response to chemo- or radiotherapy. The restoration of this checkpoint could therefore sensitize many tumour cells to conventional therapies. Strategies along these lines involve the development of compounds that can reactivate mutant p53 or inhibit MDM2, or the use of gene therapeutic approaches for the reintroduction of functional *p53* genes (Nielsen and Maneval, 1998).

Other drug-based strategies aim at an inhibition of checkpoint control to improve the efficacy of existing therapies that rely on DNA damage, such as radiation or DNA-damaging chemotherapy, by means of minimizing the time available for DNA repair. Suitable targets in this context are the signalling pathways transmitting the damage signals to the checkpoint machinery. Prime candidates for the discovery of chemo- and radiosensitizing drugs are therefore proteins such as ATM, ATR, Chk1/2 and Plk1, which regulate checkpoints in G_1, G_2 and mitosis (see **Figure 7**). First results obtained with an inhibitor of checkpoint kinase Chk1, UCN-01 (Busby *et al*., 2000; Graves *et al*., 2000), suggest that this strategy may indeed be successful with p53-negative cells, although the consequences of eliminating checkpoint control in normal cells could remain a major problem with respect to therapy-induced side effects.

Perspective

These potential applications are just examples. Numerous other mechanisms controlling cell cycle progression have been discovered and approaches for therapeutic intervention are being developed, suggesting that targeting of the cell cycle has great potential for the development of new anticancer drugs. It can be anticipated that this new class of anticancer drugs will lead to important advances in clinical oncology.

Please see also the following chapters: *Regulation of the Cell Cycle*; *Inherited Predispositions to Cancer*; *Genetic Instability and DNA Repair*; *Apoptosis*; *Signalling by Cytokines*; *Signalling by Tyrosine Kinases*; *The Formation of DNA Adducts*; *Gene Knockouts in Cancer Research*; *Models for Drug Development and Drug Resistance*; *Mechanisms of Action of Cancer Therapeutic Agents: DNA-Interactive Alkylating Agents and Antitumour Platinum-Based Drugs*; *Drug Resistance and Reversal*; *Growth Factor Receptor Blockade*; *Signal Transduction Pathway Targeting*.

ACKNOWLEDGEMENTS

Our research has been supported by the Deutsche Forschungsgemeinschaft, the Dr Mildred Scheel Stiftung für Krebsforschung and the National Foundation for Cancer Research (USA) through Krebsforschung International eV (Germany).

REFERENCES

Achenbach, T. V., *et al*. (2000). Bcl-2 independence of flavopiridol-induced apoptosis: mitochondrial depolarization in the absence of cytochrome *c* release. *Journal of Biological Chemistry*, **275**, 32089–32097.

Agami, R. and Bernards, R. (2000). Distinct initiation and maintenance mechanisms cooperate to induce G_1 cell cycle arrest in response to DNA damage. *Cell*, **102**, 55–66.

Bible, K. C. and Kaufmann, S. H. (1997). Cytotoxic synergy between flavopiridol (NSC 649890, L86-8275) and various antineoplastic agents: the importance of sequence of administration. *Cancer Research*, **57**, 3375–3380.

Busby, E. C., *et al*. (2000). The radiosensitizing agent 7-hydroxystaurosporine (UCN-01) inhibits the DNA damage checkpoint kinase hChk1. *Cancer Research*, **60**, 2108–2112.

Buschfort, C., *et al*. (1997). DNA excision repair profiles of normal and leukaemic human lymphocytes: functional analysis at the single-cell level. *Cancer Research*, **57**, 651–658.

Chaney, S. G. and Sancar, A. (1996). DNA repair: enzymatic mechanisms and relevance to drug response. *Journal of the National Cancer Institute*, **88**, 1346–1360.

Eilers, M. (1999). Control of cell proliferation by Myc family genes. *Molecular Cells*, **9**, 1–6.

Friedberg, E. C., *et al*. (1995). *DNA Repair and Mutagenesis*. (ASM Press, Washington, DC).

Graves, P. R., *et al*. (2000). The Chk1 protein kinase and the Cdc25C regulatory pathways are targets of the anticancer agent UCN-01. *Journal of Biological Chemistry*, **275**, 5600–5605.

Harbour, J. W. and Dean, D. C. (2000). The Rb/E2F pathway: expanding roles and emerging paradigms. *Genes and Development*, **14**, 2393–2409.

Jiricny, J. (1998). Eukaryotic mismatch repair: an update. *Mutation Research*, **409**, 107–121.

Johnson, R. E., *et al*. (2000). Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature*, **406**, 1015–1019.

Kawate, H., *et al*. (1998). Separation of killing and tumorigenic effects of an alkylating agent in mice defective in two of the DNA repair genes. *Proceedings of the National Academy Sciences of the USA*, **95**, 5116–5120.

Le Page, F., *et al*. (2000). Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell*, **101**, 159–171.

Mailand, N., *et al*. (2000). Rapid destruction of human Cdc25A in response to DNA damage. *Science*, **288**, 1425–1429.

Muhua, L., *et al*. (1998). A cytokinesis checkpoint requiring the yeast homologue of an APC-binding protein. *Nature*, **393**, 487–491.

Nielsen, L. L. and Maneval, D. C. (1998). p53 tumor suppressor gene therapy for cancer. *Cancer Gene Therapy*, **5**, 52–63.

Pavletich, N. P. (1999). Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *Journal of Molecular Biology*, **287**, 821–828.

Pegg, A. E. (1990). Mammalian O^6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogens and therapeutic agents. *Cancer Research*, **50**, 6119–6129.

Pfeifer, G. P. (ed.) (1996). *Technologies for Detection of DNA Damage and Mutations*. (Plenum Press, New York).

Rajewsky, M. F., *et al*. (1998). Relevance of DNA repair to carcinogenesis and cancer therapy. *Recent Results in Cancer Research*, **154**, 128–146.

Reese, J. S., *et al*. (1996). Retroviral transduction of a mutant *MGMT* into human $CD34^+$ cells confers resistance to O^6-benzylguanine plus BCNU. *Proceedings of the National Academy of Sciences of the USA*, **93**, 14088–14093.

Rotman, G. and Shiloh, Y. (1999). ATM: a mediator of multiple responses to genotoxic stress. *Oncogene*, **18**, 6135–6144.

Scolnick, D. M. and Halazonetis, T. D. (2000). Chfr defines a mitotic stress checkpoint that delays entry into metaphase. *Nature*, **406**, 430–435.

Senderowicz, A. M. (1999). Flavopiridol: the first cyclin-dependent kinase inhibitor in human clinical trials. *Investigations of New Drugs*, **17**, 313–320.

Sielecki, T. M., *et al*. (2000). Cyclin-dependent kinase inhibitors: useful targets in cell cycle regulation. *Journal of Medicinal Chemistry*, **43**, 1–18.

Smits, V. A., *et al*. (2000). Polo-like kinase-1 is a target of the DNA damage checkpoint. *Nature Cell Biology*, **2**, 672–676.

Tanaka, H., *et al*. (2000). A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage. *Nature*, **404**, 42–49.

Weinert, T. (1998). DNA damage and checkpoint pathways: molecular anatomy and interactions with repair. *Cell*, **94**, 555–558.

Wood, R. D. (1996). DNA repair in eukaryotes. *Annual Review of Biochemistry*, **65**, 135–167.

FURTHER READING

Bartek, J., *et al*. (1999). Perspective: defects in cell cycle control and cancer. *Journal of Pathology*, **187**, 95–99.

D'Atri, S., *et al*. (1998). Involvement of the mismatch repair system in temozolomide-induced apoptosis. *Molecular Pharmacology*, **54**, 334–341.

Esteller, M., *et al*. (2001). A gene hypermethylation profile of human cancer. *Cancer Research*, **61**, 3225–3229.

Fink, D., *et al*. (1997). *In vitro* and *in vivo* resistance to cisplatin in cells that have lost DNA mismatch repair. *Cancer Research*, **57**, 1841–1845.

Koepp, D. M., *et al*. (1999). How the cyclin became a cyclin: regulated proteolysis in the cell cycle. *Cell*, **97**, 431–434.

Li, G.-M. (1999). The role of mismatch repair in DNA damage-induced apoptosis. *Oncology Research*, **11**, 393–400.

Loeb, L. A. (2001). A mutator phenotype in cancer. *Cancer Research*, **61**, 3230–3239.

Nature insight: Cancer (2001). *Nature*, **411**, 335–395.

O'Connell, M. J., *et al*. (2000). The G_2-phase DNA-damage checkpoint. *Trends in Cellular Biology*, **10**, 296–303.

Page, A. M. and Hieter, P. (1999). The anaphase-promoting complex: new subunits and regulators. *Annual Review of Biochemistry*, **68**, 583–609.

Sherr, C. J. and Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G_1-phase progression. *Genes and Development*, **13**, 1501–1512.

Sherr, C. J. and Weber, J. D. (2000). The ARF/p53 pathway. *Current Opinions in Genetics and Development*, **10**, 94–99.

Simon, J. A., *et al*. (2000). Differential toxicities of anticancer agents among DNA repair and checkpoint mutants of *Saccharomyces cerevisiae*. *Cancer Research*, **60**, 328–333.

Vogelstein, B., *et al*. (2000) Surfing the p53 network. *Nature*, **408**, 308–310.

Young, R. A. (2000). Biomedical discovery with DNA arrays. *Cell*, **102**, 9–15.

Zeng-Rong, N., *et al*. (1995). Elevated DNA repair capacity is associated with intrinsic resistance of lung cancer to chemotherapy. *Cancer Research*, **55**, 4760–4764.

Zhou, B.-B. S. and Elledge, S. J. (2000). The DNA damage response: putting checkpoints in perspective. *Nature*, **408**, 433–439.

Website

Cancer Research Encyclopedia: http://claim.springer.de/EncRef/CancerResearch/default.htm.

Chapter 98

Gene Therapy – Tumour-suppressor Gene Replacement/ Oncogene Suppression

Jack A. Roth

University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

Susan F. Grammer

Biotechwrite: Biomedical and Science Communications, Houston, TX, USA

CONTENTS

- Introduction to Gene Therapy
- Gene Therapy for Cancer
- Summary and Conclusions
- Acknowledgements

INTRODUCTION TO GENE THERAPY

Historical Perspective

Gene therapy technology emerged naturally from the knowledge that DNA is the blueprint for the functional program of every cell, and of the recognition that specific alterations in DNA give rise to specific diseases. The common goal of all gene therapy strategies is to alter gene expression to treat, cure or prevent diseases resulting from inherited or acquired genetic damage. Successful gene therapy requires that the desired gene sequence be transferred to appropriate cells and expressed, and that expression be regulated. Genes destined for use in gene therapy are usually transferred via either viral vectors, which exploit the ability of viruses to adsorb to and infect eukaryotic cells, or nonviral vectors, which exploit the membrane properties of cells. Most gene therapy strategies currently in development target only somatic cells; once technical and ethical issues are resolved, however, the exploitation of germ-line gene therapy to prevent inherited genetic diseases has great potential.

In order to be considered successful, a gene therapy approach must accomplish three things. First, the gene must be delivered to the cell that will express it; second, it must be expressed; and third, expression must be regulated. For some applications, such as cancer, the expression may be transient. During early gene therapy clinical trials, many technological obstacles to meeting these three basic requirements became apparent, making vector design and gene expression studies critical to the further evolution of gene therapy. Phase I and II clinical trials are the appropriate venue in which to identify these limitations, and are generally designed to answer the following questions: Is the vector safe? What is the maximum tolerated dose? Did the gene transfer into the patient's cells? Is the gene expressed? Is there an indication of a therapeutic response? Additional research addressing optimization of all aspects of the clinical protocols is then required, so that therapeutic responses can be accurately analysed in further clinical trials.

Along with the technological obstacles, which confront the field of gene therapy, scientists and clinicians also face extensive regulations and public scrutiny. Some scientists have suggested that, in order to temper some of the distrust of gene manipulation technology, gene therapy might be presented as a more efficient method of administering conventional protein drug therapy (Wilson, 1999). Rather than manufacturing a therapeutic protein (such as insulin, erythropoietin, growth hormone or cytokines) at a pharmaceutical house, the machinery to produce the protein – the new gene – is introduced into the patient's cells and the protein is then manufactured by the cell type which normally produces it. The gene essentially becomes a pharmaceutical.

Early gene therapy research was aimed at inherited monogenic diseases – those caused by loss or alteration of one gene. Initial targets of gene therapy appeared to require a relatively simple strategy – insert and express a normal copy of the mutated or deleted gene and cure the disease. It has not proved so simple.

Testing of the first generation of viral vectors for gene transfer in 1990 unmasked several limitations. The transfer of genes was inefficient with the vectors available at the

time, expression was not stable for long periods, and the viral vectors were immunogenic. Improvements in vector design became a critical prerequisite to further developments in gene therapy, and the biopharmaceutical industry became involved in development and testing in the 1990s.

The first human gene therapy trials began in 1990. In these trials, cells to be altered were harvested from patients and the genetic material was manipulated *ex vivo*, or outside the patient. After insertion of the desired gene sequence into cultured cells, modified cells were transplanted back into patients. Early trials included gene therapy for an inherited form of immunodeficiency and treatment of children and adults with very high serum cholesterol levels.

Gene therapy rapidly evolved to the level of *in vivo* trials, in which the vector is administered directly to the patient. Early *in vivo* trials utilizing attenuated adenovirus to treat cystic fibrosis, and phase I and II clinical trials of gene therapy for ADA, Gaucher disease, haemophilia, Duchenne muscular dystrophy and sickle cell anaemia, have since produced promising results (Blaese *et al*., 1995; Crystal, 1995; Kiem *et al*., 1995; Miller and Vile, 1995).

Following pioneering studies with inherited monogenic disorders, gene therapy research expanded into acquired diseases with a genetic component, including infectious diseases such as AIDS, cardiovascular diseases and cancer.

In his Presidential Address at the Second Annual Meeting of the American Society of Gene Therapy, James M. Wilson identified four issues critical to the success of gene therapy (Wilson, 1999). First was an 'ongoing commitment to, and investment in, basic research relevant to somatic cell transplantation and gene transfer.' Second, Wilson discussed the 'role of clinical trials,' acknowledging that 'human data are absolutely critical in driving this field forward.' The third issue which Wilson presented to the gene therapy community was 'the value of our community and how we relate to one another as colleagues and competitors.' Wilson noted that true success would be measured by 'proof-of-concept in humans' and 'the commercial development of effective gene therapy products.' Reaching these goals will require cooperation from all involved in addressing ethical and regulatory issues. And, because the fourth critical issue was 'the development of gene therapy as a business,' Wilson stated that all participants have a responsibility to 'identify and address any potential barriers to commercial success.'

The rapid evolution of gene therapy from the laboratory bench to clinical trials has been possible thanks to major developments in the science of molecular genetics. Gene therapy as a topic in peer-reviewed journals climbed from less than 100 articles per year in 1989 to over 1200 in 1998. Rapid progress in the mapping of the human genome has allowed for the identification of the genetic component of many more diseases, broadening the pool of potential gene therapy targets.

Tools of Gene Therapy

In order to understand fully the rapidly expanding pool of new data, it is helpful to understand the basic tools of gene therapy, developed thanks to recent advances in the technology for isolation and characterization of genes and transfer into animals. This chapter will present only a general overview of these enabling technologies; the reader is directed to the section on The Molecular Basis of Cell and Tissue Organisation of *The Cancer Handbook* for more thorough coverage of molecular biology basics.

Basic Genetics – From Gene to Protein

Genes are the blueprints for the assembly of proteins from free amino acids. Each gene has a unique nucleotide sequence and occupies a specific region of a chromosome. A gene carries not only the unique nucleotide sequence which codes for the gene product, but also numerous regulatory sequences which can be turned on or off by other proteins in the cell, dictating whether a gene will be expressed.

Nearly every cell in an organism contains a complete set of inherited genetic material which defines its genotype; each cell will only express a subset of this suite of genes, however, depending primarily on each gene's promoter sequence. (See chapters on *Cell and Tissue Organisation*; *Regulation of the Cell Cycle*.) Gene expression depends on regulatory elements which act at the level of transcription (DNA to RNA) or of translation (RNA to protein). The resulting gene expression profile confers on a cell its phenotype.

Gene mutations, which alter the DNA sequence and therefore the genotype of all daughter cells, may cause a gene not to be expressed, to be expressed inappropriately, or to produce a nonfunctional protein. Mutations can occur in both alleles of a gene or in only one. Some gene products can be expressed normally if one undamaged copy remains, but others require two normal alleles. The loss of one allele is a heterozygous mutation; the loss of both is a homozygous mutation and results in total loss of gene function. In some cases a mutation in one allele is combined with loss of the other allele; for example, a point mutation in one allele of the tumour-suppressor gene *p53* is often combined with the loss of the p arm of the human chromosome 17. This is referred to as loss of heterozygosity (LOH), which can be detected by analysis of tumour cell DNA using methods discussed below. LOH in chromosome 17p, chromosome 10 or a number of other chromosomes in tumour cells often correlates with the level of malignancy to which a tumour has progressed. (See chapters on *Overview of Oncogenesis*; *Genomic Instability and DNA Repair*.)

The Tools

Manipulation of DNA

Gene therapy is entirely dependent on the technology which permits the isolation of DNA from a cell and

sequencing of nucleotides. In addition, a suite of enzymes, called restriction endonucleases, can cut DNA at specific nucleotide sequences, allowing for the isolation of entire genes or sections of genes. Fragments of DNA are separated according to size on a gel matrix, and specific known sequences can be visualized by 'probing' with labelled oligonucleotides. Other enzymes, called ligases, allow the joining together or ligation of previously unrelated pieces of DNA to form new gene sequences or DNA constructs, often with new promoter or regulatory sequences. A third class of enzymes, DNA polymerases, allows the synthesis of large amounts of new DNA in the laboratory.

One type of recombinant DNA construct useful in gene therapy is assembled by splicing a tissue-specific promoter sequence from one gene together with the protein-coding region of a second gene, resulting in a transgene which can be transferred into cells in tissue culture or into living organisms, and expressed, at the direction of the altered promoter sequence, in a different cell type. Some DNA constructs alter the structure, and therefore the function, of the final protein, and others insert an entirely new sequence in place of the original. (See chapters on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*; *Gene Knockouts in Cancer Research*.)

Gene Transfer

Gene sequences must be transferred into cells for expression to occur. This can be accomplished by several methods.

Transfection is the process of transferring new DNA into cells growing in culture. Cells take up the DNA and a small amount becomes incorporated and expressed. Transfected cells are then propagated to generate progeny which all contain the altered genotype and hopefully express the desired phenotype.

Viral vectors are DNA constructs built into a viral genome. Gene delivery via viral vectors exploits a virus's innate ability to bind receptors on and actively infect cells. The viral genome is altered in the laboratory to reduce expression of sequences required for replication, while maintaining the viral genes which are responsible for infection and expression. Desired eukaryotic gene sequences are inserted into the viral genome and then enter the cells as passengers within the viral nucleic acid. Viral vectors often contain powerful promoters to drive gene expression. Current progress in the development of viral vectors for cancer gene therapy will be discussed in this chapter.

Non-viral vectors, designed to deliver therapeutic genes without the need for intact virus, have shown promise in initial studies. Liposomes (lipids complexed with DNA), molecular conjugates (which attempt to target passenger DNA to a specific cell type using synthetic ligands coupled to DNA) and naked DNA have all been delivered to cells by mechanical means.

Animal Models in Gene Technology

Technology allowing alteration of the genetic make-up of entire organisms has been developed and the resulting animal models are useful in evaluating the biological function of new genes. Generation of transgenic animals (see chapter on *Transgenic Technology in the Study of Oncogenes and Tumour-Suppressor Genes*) involves injection of a transgene, or new DNA sequence, into fertilized mouse eggs which are subsequently placed in the oviducts of pseudo-pregnant mice. Progeny contain the transgene in every cell in the body, but expression is restricted by the specific promoter and regulatory sequences incorporated into the transgene. Using this technology, DNA of one species can be expressed in cells of another species. An example is the production of human growth hormone and insulin by transgenic laboratory animals.

Another useful animal model is the knockout mouse (see chapter on *Gene Knockouts in Cancer Research*), which provides a definitive method for establishing *in vivo* gene function. By deleting a specific gene sequence, the actual morphological or physiological outcome of the loss' of a specific gene can be evaluated; subsequent attempts to replace the missing gene product with a pharmaceutical or the missing gene via gene therapy can be evaluated for return of functional activity.

A more recent development is the knock-in mouse, a more technically sophisticated model in which a particular gene sequence is replaced with a modified functional gene or an entirely new gene sequence.

As is apparent from these advanced technologies, recombinant DNA technology, once a tool found only on the bench in a select few laboratories, has rapidly progressed through the laboratory animal and to the patient's bedside.

GENE THERAPY FOR CANCER

Introduction – Where Are We Now?

Suspicions that faulty gene expression might be the aetiological agent of cancer were voiced as early as 1914, when T. Boveri wrote (Boveri, 1929):

> *The unlimited tendency to rapid proliferation in malignant tumor cells [could result] from a permanent predominance of the chromosomes that promote division . . .. Another possibility [to explain cancer] is the presence of definite chromosomes which inhibit division . . .. Cells of tumors with unlimited growth would arise if those 'inhibiting chromosomes' were eliminated . . .. [Since] each kind of chromosome is represented twice in the normal cell, the depression of only one of these two might pass unnoticed . . ..*
>
> (T. Boveri, 1914; translation published in 1929)

As recently as 1966, many scientists still denounced the theory that neoplasms were the result of mutations in somatic cells. Ironically, Peyton Rous, who spoke out strongly against a role for mutations in tumorigenesis, discovered the tumour virus which eventually resulted in the universally accepted genetic paradigm that has driven cancer research in recent years (Bishop, 1991). It is now generally accepted that cancer is caused by multiple mutations in DNA, which take place sequentially and ultimately lead to transformation of a cell to the malignant phenotype. (See relevant chapters in section on the *Causation and Prevention of Cancer*.)

Early strategies for treating cancer focused on methods for surgical removal of neoplastic tissue, along with radiation therapy and chemotherapy to kill surviving tumour cells. The last two strategies exercise their effect on the tumour cell by damaging DNA; because the tumour cells are rapidly dividing, they are killed more rapidly than adjacent normal cells. More recently, immunotherapeutic strategies have been added to the arsenal and are often used in combination with radiation and chemotherapy to increase specificity or enhance the immune response. All of these treatment strategies involve significant systemic side effects to the patient, and also significant bystander damage to normal tissues. Although recent developments in all of these approaches have benefitted many patients, lack of specificity for the offending tumour cell and toxicity to the patient continue to limit their ability to cure cancer.

As the aetiology of cancer became more apparent, it became possible to consider attacking the disease at its root – the gene. Rather than destroying all cells in the process of synthesizing DNA, a particular gene might be isolated and replaced or repaired, leaving normal cells unaffected by treatment. Gene therapy, as was discussed in the previous section, offers the potential to overcome the lack of specificity inherent in most cancer therapies with an acceptable level of toxicity to the patient. Although gene therapy first arose as an approach to treating monogenic diseases, cancer therapy has become its number one application in recent years.

Modern tools and techniques of gene therapy are being applied to cancer via two distinctly different approaches. One approach involves augmentation of the more conventional approaches to cancer – immunotherapy and chemotherapy – and has been called 'gene therapeutics' (Roth and Cristiano, 1997). Strategies encompassed within this 'augmentative' approach attempt both to confer specificity and to reduce toxicity, and include transfer of cytokine genes to augment immunotherapy (see chapter on *Antibodies and Recombinant Cytokines*), drug sensitization with genes for prodrug delivery and transduction of drug resistance genes into bone-marrow stem cells as a means of protecting normal bone-marrow cells during high-dose chemotherapy (see chapter on *Drug Resistance and Reversal*). These strategies are discussed in detail in other chapters.

The second general approach to the gene therapy of cancer, which is the subject of this chapter, targets the underlying genetic lesion in the cancer cell – the aetiology of the disease – and has been called 'gene replacement' therapy (Roth and Cristiano, 1997). The gene families which have been implicated as the aetiological agents of the most common cancers include dominant oncogenes [mutant alleles of a class of normal genes called proto-oncogenes (Bishop, 1991)] and tumour-suppressor genes (Weinberg, 1991). Because strategies to replace functional copies of these genes target the underlying genetic lesion responsible for the disease, they have the ultimate potential both to treat and to prevent cancer.

Numerous proto-oncogenes and tumour-suppressor genes are expressed in normal cells, and their protein products are critical components of the cellular machinery which drives the normal processes of signal transduction (see chapter on *Signal Transduction Pathway Targeting*), DNA transcription, cell division, cell cycle arrest and programmed cell death (apoptosis). (See chapter on *Apoptosis*.) Alterations in these genes lead to disruption of normal processes, in effect altering the functional programme of the cell. In addition, most tumours develop subsequent alterations in expression of genes encoding the proteins that govern such processes as angiogenesis and cell locomotion, or invasion. In general, proto-oncogenes participate in signal transduction and transcription. Tumour-suppressor genes appear critical in governing proliferation by regulating transcription and cell cycle control. (See also relevant chapters in the section on *The Treatment of Human Cancer*.)

Only a single dominant oncogene allele need be mutated (by point mutation, amplification, translocation or rearrangement) to accomplish malignant transformation. Historically, it was believed that, because only one functional tumour-suppressor gene allele is normally required for adequate regulation of proliferation, malignant transformation by tumour-suppressor mutation required homozygous loss of function, with both copies of the gene either mutated, deleted or both (Harris, 1996). More recent studies, however, have provided evidence for haploid insufficiency (Tang *et al*., 1998). Tumour-suppressor genes can undergo homozygous loss of function by mutation and/or deletion, or, in some cases, one faulty allele is inherited in the germ line, and the second genetic alteration occurs as a somatic mutation at a later time.

The application of gene therapy to cancer is complicated by the fact that cancer is associated with multiple genetic lesions within each malignant cell. Some lesions contribute to the aetiology of the disease, while others occur due to the underlying genetic instability of transformed cells in general (see chapter on *Genomic Instability and DNA Repair*). In some cases, a genetic defect is inherited and lies undetected for years, until an additional mutation is acquired which confers the malignant phenotype.

Historically, many clinicians believed that gene replacement therapy for cancer was impractical because of the impossibility of correcting multiple genetic abnormalities in each cancer cell. Recent observations, however, suggest that correction of a single lesion – for example, elimination of expression of a dominant oncogene or insertion of a normal copy of a tumour-suppressor gene – may be sufficient to negate the malignant potential of a tumour cell. In addition, studies conducted by the author's colleagues (Mukhopadhyay *et al*., 1991; Fujiwara *et al*., 1993), using human cancer cells with multiple genetic abnormalities, have demonstrated that transduction with retroviral or adenoviral vectors expressing either wild-type p53 tumour-suppressor (wt-p53) or antisense K-*ras* (see chapter on Antisense and Ribozyme Therapy), resulted in inhibition of cancer cell growth. A common mechanism of cell death in these cells was apoptosis, or programmed cell death (Fujiwara *et al*., 1993), emphasizing the critical role of apoptosis in tumour cell demise and its regulation by tumour-suppressor genes. As demonstrated in clinical trials to date – mainly in patients with very advanced cancers previously proved to be refractory to conventional treatment – gene replacement and gene inactivation therapy can, indeed, induce tumour regression with acceptably low toxicity (Roth and Cristiano, 1997).

Inefficient viral transduction was once considered a potential limitation for the application of gene therapy to cancer, but it has become apparent from laboratory studies and early clinical trials that transduced cells and their products confer a 'bystander effect,' or growth inhibition effect, on neighbouring nontransduced tumour cells (Nishizaki *et al*., 1999).

An additional commonly asked question has been, why replace the faulty gene and not just the missing gene product? Attempts to restore p53 gene product function with small molecule pharmaceutical therapy have failed owing to complexities in the structure of the p53 protein (Harris, 1996); it is proving to be more practical to restore the function of defective genes in the cancer cell by gene transfer.

Tumour behaviour encompasses several functional realms, such as tumorigenesis, anchorage-independent growth and invasion and angiogenesis, (Bookstein *et al*., 1990; Takahashi *et al*., 1992), making it unlikely that all cells of a large tumour mass could be killed by current gene therapy strategies. Current approaches, though, are suitably designed to aid in control of locoregional recurrence of tumours and in systemic control of micrometastases.

Cell Biology and Biochemistry of Oncogenes and Tumour-suppressor Genes

Oncogenes, first identified as transforming sequences in retroviruses which were also present in normal cells (suggesting that they were 'appropriated' by the virus), are the mutated forms of normal genes called proto-oncogenes. The publication in 1987 of data demonstrating that many human tumours contained a point mutation in the proto-oncogene *RAS* led to increased suspicions that mutagenesis might play a role in the genesis of cancer (Barbacid, 1987). Other proto-oncogenes identified early included *myc* (mutated by translocation in Burkitt lymphoma and by amplification in carcinoma of the lung, breast and cervix), H-*RAS*, K-*RAS* and N-*RAS* (altered by point mutations in numerous carcinomas – small cell lung, colon, pancreas, thyroid, genitourinary tract – and in melanoma) and *SRC* (altered in colon carcinoma). By 1989, there were already more than 60 oncogenes described. Bishop (1991) defined proto-oncogenes as follows: 'if manipulation of any sort engenders a transforming allele (or "oncogene"), convention admits the normal counterpart to the ranks of proto-oncogenes.' Proto-oncogenes have now been identified as critical members of many signal transduction pathways, and several are prominent factors in gene transcription.

Tumour-suppressor genes are most easily studied when they are missing from the genome, a fact that placed severe limitations on the study of these critical genes until recently. Developments in molecular genetics have produced new technologies for gene manipulation, resulting in a rapidly expanding literature base on tumour-suppressor genes. Some of first tumour-suppressor genes described were *RB* (retinoblastoma), incriminated in retinoblastoma, osteosarcoma and carcinomas of breast, bladder and lung *p53*, associated with astrocytoma, carcinoma of breast, bladder and lung and osteosarcoma, *WT1*, associated with Wilms' tumour, and *DCC*, associated with colon carcinoma. *RB* and *p53* have since been identified as integral players in regulation of the cell cycle, signal transduction pathways, apoptosis and other normal cellular functions.

In normal cells, the protein products of proto-oncogenes and tumour-suppressor genes feed into several signal transduction pathways. The relevant protein products include polypeptide hormones and hormone receptors that act at the cell surface, along with proteins that convey signals across the membrane, through the cytoplasm and to the nucleus. These gene classes also encode several nuclear factors that control the mix of positive and negative signals from a complex transcription factor network governing gene expression. The initial step in these biochemical cascades is commonly an interaction between a growth factor and a growth factor receptor at the cell surface (see chapters on *Growth Factor Receptor Blockade*; *Signalling by Steroid Receptors*; *Signalling by Cytokines*). Activation of surface receptors results in the expression of enzymatic activity inside the cell, which acts like a switch to catalyse further chemical reactions. Eventually, barring any difficulties associated with nonfunctional genes or gene products, an appropriate signal is transmitted through the nuclear membrane, where

activation of DNA binding transcription factors alters the expression of genes and therefore the functional programme of the cell.

Regulation of the cell cycle (see chapter on *Regulation of the Cell Cycle*) is also critical to normal cellular function, and several pathways interact to provide numerous opportunities for regulation of cell growth, differentiation and proliferation – all critical parameters in the control of tumorigenesis. Normal cells of an adult remain in either G_0 or G_1 most of the time, maintained in that state by the proper balance of positive and negative signals from two intimately interwoven pathways, the RB pathway and the p53 pathway. Protein products of several proto-oncogenes and tumour-suppressor genes play critical roles in both of these pathways, and are often dysfunctional in tumour cells.

Briefly, the expression of a number of growth factors, oncogenes, cyclins and cyclin-dependent kinases (CDKs) all drive the cell cycle towards proliferation. Expression of tumour-suppressor genes and inhibitors of CDKs induce arrest of the cell cycle when appropriate. The Rb G_1 arrest pathway is initiated by the protein p16, the product of the *p16INK4* gene, and culminates in the arrest of cells in G_1 by inhibiting phosphorylation of the product of the retinoblastoma gene, *RB*. Phosphorylation of the Rb protein is catalysed by cyclin-dependent kinase 4 (CDK4) in association with cyclin D1 which releases the transcription factor E2F; therefore, inhibitors of CDK4 can inhibit phosphorylation of Rb preventing progression through the cell cycle.

The other G_1 arrest pathway, often designated the p53 pathway, is tightly interwoven with the Rb pathway and is regulated at the protein level by several other tumour suppressors and oncogenes. The protein product of the tumour-suppressor gene *p19ARF*, which is encoded by the same gene locus as *p16INK4a*, but is read in a different reading frame (Kamijo *et al*., 1997), inhibits the protein Mdm2 (product of the proto-oncogene *MDM2*) which is a regulator of p53 at the protein level. Deletion of *p19ARF* results in increased levels of Mdm2 and inactivation of p53, allowing progression through the cell cycle at an inappropriate time. Proto-oncogenes, such as *MYC* and *RAS*, inhibit Mdm2 by inducing expression of *p19ARF*. Mutations in *MYC* and *RAS* result in oncogene expression, loss of inhibition of Mdm2 and, again, dysregulation of the cell cycle through p53. Several informative reviews on this unique locus, now designated *INK4A/ARF*, and its associated pathways are listed in the Further Reading section at the end of this chapter.

In summary, control of proliferation in mammalian cells is essentially a function of these two pathways, with the Rb protein regulating maintenance and release from the G_1 phase and the p53 protein effecting growth arrest and/or apoptosis in response to cellular stress and DNA damage. Both of these tumour-suppressor genes and many of the proto-oncogenes involved in these pathways are commonly altered in a variety of cancers.

Tumour-suppressor Gene Replacement

Rationale

The most common mutation seen in cancer is in the tumour-suppressor gene *p53* (also known as *TP53*), and the inactivation of this gene appears to contribute to tumorigenesis; it follows, then, that replacement of a copy of a 'wild-type' (wt; nonmutated) gene might be sufficient to restore normal growth and proliferation pathways. The *p53* gene product (p53) is a phosphoprotein consisting of 393 amino acids and capable of complexing with viral proteins, which can inactivate the p53. The p53 protein seems to be multifunctional – it contains major domains that transactivate other genes, can bind other proteins, can bind sequence-specific DNA and can form oligomers with other p53 proteins. Abnormalities in any of these critical functions have the potential to eliminate or severely reduce the tumour-suppressor function of the p53 gene product.

Because tumour-suppressor genes normally regulate gene transcription and cell proliferation, gene replacement therapy might potentially have its effect through induction of apoptosis, by induction of tumour dormancy or by prevention of transformation of premalignant cells to the malignant phenotype. p53 is also a regulatory element for the progression of a cell from G_1 to G_2 in the cell cycle. Normally, when DNA is damaged p53 will cause arrest at G_1 until the DNA is repaired or will trigger apoptosis.

Loss of cell cycle regulatory function or DNA repair capabilities, or untimely expression of molecules normally held in check by the p53 protein, may make the cell more susceptible to transformation events. Therefore, replacement of p53 function has the potential to suppress other genes which drive the cell toward uncontrolled cell growth and the potential to activate genes that will suppress pre-existing signals for uncontrolled cell growth and progression to the transformed phenotype.

Preclinical Studies

Preclinical studies, both *in vitro* and *in vivo*, have demonstrated that the restoration of p53 function can result in the induction of apoptosis in cancer cells. Laboratory studies have demonstrated suppression of tumour growth in an orthotopic human lung cancer model, when retroviral p53 expression vector was administered (Fujiwara *et al*., 1994a). Functional *p53* gene expression, restored via a retroviral p53 expression vector, has been shown to suppress the growth of some human lung cancer cell lines, but not others (Cai, *et al*., 1993). Cells transduced with the *p53* gene were also shown to reduce the growth rate of nontransduced cells in human lung cancer, indicating the existence of a bystander effect. Recent evidence suggests that the bystander effect may be mediated by the antiangiogenic effects of wild-type p53 and by release of proapoptotic factors such as Fas ligand and Fas (Owen-Schaub *et al*., 1995; Bouvet *et al*., 1998). (See also

chapters on *Extracellular Matrix*; *Invasion and Metastasis*; *Angiogenesis*.)

An adenovirus vector, developed to deliver the wild-type human *p53* gene to tumour cells, was found to be capable of inducing apoptosis in cancer cells with altered non-functional p53, but did not appreciably affect proliferation of normal cells (Wang *et al*., 1995). In addition, the same vector inhibited tumour growth in a mouse model of human orthotopic lung cancer (Georges *et al*., 1993). Products of other tumour-suppressor genes, including *RB* and *p16*, have also been found to suppress tumour growth in animal models (Jin *et al*., 1995; Xu *et al*., 1996).

Clinical Application

Taken together, data from preclinical studies led to the approval by the Recombinant DNA Advisory Committee of the National Institutes of Health (NIH) and the US Food and Drug Administration (FDA) of the rationale for the first clinical protocol involving injection of recombinant retrovirus expressing the wild-type (normal) *p53* gene. This protocol aimed to replace a defective *p53* gene with a retroviral vector expressing normal *p53* (Roth, 1996a,b) in patients with unresectable lung cancer, which had proved resistant to other interventions. Of nine patients enrolled in this protocol, eight had received radiation, five chemotherapy and five surgical resection of the primary lung tumour (three) or of a brain metastasis (two). All eight patients who completed the protocol showed some evidence of gene transfer. Three of seven evaluable patients showed evidence of tumour regression in treated lesions, while other untreated lesions progressed. In these seven evaluable patients, no toxic effects were observed which were directly attributable to the vectors, although there were complications related to the procedure used to administer the vectors.

Based on results of studies in animal models demonstrating tumour regression following intratumoural injection of an adenovirus vector containing p53 complementary DNA, a phase I clinical trial was initiated with 28 non-small cell lung cancer patients whose cancers had progressed with conventional treatments (Swisher *et al*., 1999). Polymerase chain reaction (PCR) analysis demonstrated the presence of vector DNA in 80% of the evaluable patients, indicating successful gene transfer; vector-specific p53 mRNA, an indicator of gene expression, was detected in 46% of patients. Apoptosis of tumour cells was demonstrated in all but one of the group of patients expressing the gene and vector-related toxicity was minimal, despite up to six repeated injections per patient. Two of the 25 evaluable patients (8%) exhibited partial responses, 16 patients (64%) exhibited disease stabilization ranging from 2 to 14 months, and the remaining seven patients (28%) exhibited disease progression. The results of this study demonstrated efficient gene transfer and expression, low toxicity and an indication of antitumour activity. Similar studies have been carried out by Clayman and colleagues with patients with recurrent head and neck squamous cell carcinomas, with similar findings (Clayman *et al*., 1998).

Oncogene Suppression

Rationale and Preclinical Studies

The *RAS* family of oncogenes, which includes H-*RAS*, N-*RAS* and K-*RAS*, each encodes an oncoprotein localized on the inner plasma membrane. These proteins have guanosine triphosphatase (GTPase) activity and play a role in signal transduction. Point mutations in the *RAS* genes alter the protein product p21, interrupting the many signal transduction cascades in which p21 plays a role. Gene therapy strategies aimed at correcting the defect at *RAS* involve antisense technology, in which the therapeutic gene introduced into the cell has a DNA sequence complementary to the RNA sequence of the oncogene. When the antisense sequences bind to the sense sequences, the production of mutant RAS protein is inhibited. As reviewed (Roth and Cristiano, 1997), the growth rate of human lung cancer cells *in vitro* and in animal models has been reduced by transduction with either an antisense K-*RAS* DNA plasmid or a retroviral construct. Using a similar strategy, transduction of murine lung cancer cells with an antisense cyclin D1 gene construct reduced proliferation of tumour cells.

Another approach to oncogene suppression is the transfer of a gene that is known to block a particular activated oncogene (reviewed by Swisher and Roth, 1999). An example of this approach is the HER-2/*neu* oncogene that encodes an epidermal growth factor (EGF)-related tyrosine kinase. Overexpression of HER-2/*neu* has been observed in many cancers and correlates with poor prognosis. Her-2/*neu*-mediated malignant transformation has been inhibited by the gene product of an adenovirus construct, E1A (Zhang *et al*., 1995).

Critical Areas for Future Research

Identification of Synergies Between Gene-based Agents and Other Cancer Therapies

An exciting area for future research is the evaluation of combination therapeutic strategies. The molecular mechanisms responsible for tumour cell death after chemotherapy and radiation therapy (see chapter on *Molecular Mechanisms of Radiotherapy*), may involve p53-induced apoptosis. Tumours with altered or deleted p53 are often resistant to chemotherapy- or radiation-induced apoptosis. *p53* delivered by adenoviral vectors has demonstrated enhanced apoptosis in both *in vitro* and *in vivo* studies (Nguyen *et al*., 1996; Spitz *et al*., 1996).

Additional preclinical studies have demonstrated a synergistic relationship between p53 replacement therapy and DNA-damaging chemotherapeutic agents such as cisplatin and etoposide. In addition, studies have suggested that gene therapy may increase sensitivity to radiation. One study of human p53-deficient lung tumours growing as subcutaneous tumours in laboratory mice demonstrated that cisplatin was unable to cause tumour regression or apoptosis when delivered alone; however, when delivered with or following administration of an adenoviral p53 vector, tumours regressed and apoptosis was observed. p53, together with DNA-damaging agents such as cisplatin (Platinol) and ionizing radiation, might cause the induction of apoptosis when neither is capable of doing so alone. Fujiwara *et al*. (1994b) enhanced the sensitivity of p53-deficient cancer cell lines to CDDP by replacing the wt-*p53* gene. Other researchers (Son and Huang, 1994) have shown that the treatment of CDDP-resistant tumour cell lines with CDDP increases their sensitivity to transduction by DNA-carrying liposome vectors. In another approach, Chen *et al*. (1995) demonstrated improved tumour killing by combining interleukin expression with transduction via an HSV-TK vector. Based on the positive results in these and other studies, it follows that conventional therapy combined with gene therapy will indeed lead to enhanced therapeutic effects.

Vector Design

The technology for the manipulation of DNA and the generation of new DNA constructs is well developed. The major technological limitation facing gene therapy is the vector, which carries the desired DNA sequences to the cells in which they are to be expressed. The efficiency of gene transfer, induction of gene expression, regulation of gene expression and safety of the vector (including toxicity and immunogenicity) all must be optimized to optimize potential therapeutic benefits to the patient.

Viral Vectors

Most of the vectors currently in use in clinical trials are based on viruses and take advantage of the virus's unique ability to target, adhere to and infect cells. The challenge of viral vector development is to remove the disease-causing portion of the virus's own genome while retaining infectivity, and also include the desired gene sequence that will, when expressed, lead to therapeutic benefits for the patient.

The most commonly used viral vectors in gene therapy are retroviral, adenoviral and herpes vectors. Retroviruses are single-stranded RNA viruses that integrate into the host genome, providing potential for stable expression over time. Adenoviral and herpes vectors are DNA viruses and are maintained episomally – separate from the host's chromosomes. They may be passed on to progeny through several cell divisions, but are eventually lost; consequently, continued gene expression requires subsequent doses of virus. In spite of this limitation, herpes and adenoviral vectors are able to infect a higher percentage of cells, and so have become popular tools for gene therapy. In fact, unstable expression is not necessarily a disadvantage in cancer patients, because prolonged expression is not required or even necessarily desirable after destruction of the tumour cells. Unlike retroviruses, which infect only dividing cells, adenoviral vectors have the advantage of infecting both dividing and nondividing cells. They are also easily manipulated *in vitro*; therefore, large-scale production is feasible. Because of the antigenicity of adenoviruses, it remains to be discovered how many repeat treatments a patient will be able to receive without adverse effects. Numerous studies are under way to determine the best course to take in combating the limitations of adenoviral vector transfer.

Herpes vectors have generated interest because of their ability to establish latent infections in the brain. They have been studied for gene delivery to neurons and have potential application in brain cancer. Other viruses, including vaccinia viruses, poxviruses and baculoviruses, are also under investigation for potential gene therapy applications.

As of 1997, most approved clinical protocols for gene therapy involved retroviral vectors. Because retroviral vectors tend to integrate more efficiently into replicating cells, these have been the most useful for stable integration of desired DNA sequences into tumour cells. Retroviruses, however, can transfer only small pieces of DNA and large-scale production of vector is difficult, making the treatment of large tumours unlikely. Adenoviral vectors, on the other hand, are readily produced in large batches, but have several potential toxicity and immunogenicity issues limiting their widespread use.

Nonviral Vectors and Naked DNA

Other gene delivery strategies under investigation were reviewed by Roth and Cristiano (1997) and include naked DNA, proteins, lipids and combinations formed into synthetic particles. Liposomes, which consist of lipids complexed with DNA, are relatively simple to prepare, but the system lacks the ability to target to specific cell types. Molecular conjugates, on the other hand, result in highly specific and efficient gene delivery *in vitro*, but the duration of gene expression is very short. This approach, however, provides the groundwork for the generation of 'synthetic viruses,' capable of efficient gene delivery without the inherent risks and complications of intact viruses. Injection of 'naked DNA' into tissues offers by far the simplest delivery system, but this approach suffers from limitations such as lack of targeting, inability to reach a large number of cells and, in some cases, the need for surgical access to tissue.

Vector Targeting and Specificity

Enhancement of specificity of gene expression is being addressed through investigation of vector targeting,

tissue-specific promoters, route of delivery and modulation of immune receptors. The following is a brief summary of some current approaches; a more complete review can be found in Roth and Cristiano (1997).

Some viral vectors have a certain degree of built in specificity because of tropism for particular tissues. Adenoviruses, for example, normally infect lung epithelium and are also efficient at transducing lung epithelium. Herpes simplex virus efficiently infects neuronal cells. Although retroviruses exhibit no tissue tropism, the genes delivered are integrated and expressed only in dividing cells, making them especially suited to cancer gene therapy. Studies have shown that complexing viral vectors to a specific ligand for a cell surface receptor can enhance targeting to a specific cell type expressing that receptor. This approach provides potential to have developed targeted retroviral vectors and also the possibility of altering adenoviral-binding properties to limit or expand tissue tropism. Several ligands have been conjugated to nonviral vectors to increase targeting of these genes. Folate has been used to target ovarian carcinoma cells, which overexpress folate receptor, and asialoorosomucoid has been studied for its specificity for liver parenchyma. Additional experiments have demonstrated that overexpression of EGF receptors by some tumour cells has lead to specific uptake of EGF–DNA complexes. One obstacle to this approach has been DNA degradation when the receptor–ligand–DNA complex is internalized by receptor mediated endocytosis.

Another approach to enhancing specificity is the inclusion of tissue- or cell-specific promoters in the therapeutic gene construct. For example, the cytomegalovirus (CMV) promoter, the strongest promoter yet identified, is active primarily in rapidly dividing cells. A number of other specificity-enhancing sequences, including the carcinoembryonic antigen (CEA) promoter, and the regulatory sequence for a portion of the tyrosinase gene encoding the surfactant A protein, are under investigation. A truly novel approach to targeted gene therapy is the use of specific immune receptors on tumour-infiltrating lymphocytes to carry the therapeutic gene to the tumour. Single-chain antibodies to a renal cell carcinoma antigen have already been demonstrated to be capable of targeting tumours and lysing tumour cells. Unfortunately, the likelihood of identifying a universal cancer promoter sequence for use in gene therapy is slim, owing to the heterogeneity of tissues affected by cancer.

Integration, Expression and Regulation of Expression

In addition to reaching a target cell efficiently, a vector must be able to deliver the gene for successful integration into the cell nucleus and the gene must be expressed long enough to do its job, but not so strongly as to cause toxicity. As discussed previously, some viral vectors such as retroviruses and some adeno-associated viruses integrate into the host genome, resulting in long-term expression, which may also contribute to vector-associated toxicity. Short-term expression may be sufficient for most cancer therapy applications, so those genes which are episomally expressed provide excellent potential. Nonviral vectors, however, may provide the best potential of all.

Owing to the use of particularly strong promoters, such as CMV, it is also important to be able to regulate the promoter. Several inducible promoter sequences, including one which will increase gene expression after exposure to radiation, are under investigation. In addition, the incorporation of sequences from ‘temperature-sensitive’ viral mutants is being examined as a method to regulate viral gene expression.

Vector Immunogenicity and Toxicity

It is theoretically possible that there may be long-term vector-generated side effects associated with cancer gene therapy. No major side effects of retroviral delivery have been observed during clinical trials, but adenoviral vectors have two potential difficulties associated with their use. Systemic or local immune responses may develop owing to the need for repeated doses. Many toxic effects are expected to be overcome by improved vector targeting, elimination of some endogenous viral genes and the use of immunomodulators to ablate the immune response to vector preparations.

SUMMARY AND CONCLUSIONS

By June 1996, 13 clinical protocols using viral vectors to replace defective tumour-suppressor genes or inactivate oncogenes were in place. Of the 78 patients entered in those trials, 26 reports were evaluable in June 1996 and six had reported tumour regression. In spite of limitations to the gene therapeutic approach to cancer at this time, the current delivery strategies have been shown to have application in clinical situations and several facts stand out from the results of studies to date (Roth and Cristiano, 1997).

First, viral gene transfer has been demonstrated to be more efficient in cancer cells than was originally predicted from studies of gene transfer into normal tissues. Second, viral vectors appear to spread readily through a tumour mass and mediate cell death through apoptosis. A third and very critical fact – critical because of the low efficiency of transduction attainable with even the most potent vectors – transduced cells have proved to be capable of mediating bystander killing of nontransduced cells. Finally, earlier concerns about the potential applicability of gene therapy to cancer have been unfounded; in spite of multiple genetic alterations in tumour cells, correction of a single genetic lesion has yielded significant tumour regressions. As has

been observed in leukaemia, patients can undergo lengthy remissions in spite of the continued presence of malignant cells. It follows, then, that even partial irradication of a tumour would be clinically beneficial.

For the specialist and nonspecialist alike, and for the general public, several factors should be considered while attempting to evaluate the rapidly expanding collection of data in the scientific literature and in the public sector. Gene therapy is in its infancy. Strategies and new technologies which will revolutionize the field may be months or years from discovery, or they may be awaiting statistical analysis or in the hands of a reviewer as this sentence is being written or read. Maintenance of the current climate of cooperation between scientists and clinicians from multiple specialties and from all relevant communities – academic, healthcare and pharmaceutical industry – is necessary to continue the trend of taking these developments rapidly from the molecular biology laboratory bench to the patient's bedside.

In addition, public opinion of gene therapy technology will be shaped by the layman's perception of responsibility shown by those active in the field. As was suggested by James Wilson in his Presidential Address to the American Society of Gene Therapy (Wilson, 1999), much of the fear of gene therapy might be resolved by a widespread understanding that it is the machinery for the natural manufacture of a necessary protein that is being replaced in these strategies.

Current limitations of gene therapy, including efficiency of gene transfer and expression, along with control of potential side effects from the use of viral vectors, will be best battled if researchers and clinicians conform to strict guidelines for reporting complete data from clinical trials. Details on study design, patient population, prior treatment, vector design, gene expression and toxic side effects, as well as on therapeutic benefits observed, should be available to all researchers and clinicians involved in the development of gene therapy strategies. It is through the cooperation of scientists and clinicians in all fields that cancer will be controlled and eventually cured through the use of gene transfer technology.

ACKNOWLEDGEMENTS

This work was funded by grants from the National Cancer Institute and the National Institutes of Health (P01 CA78778-01A1) (J.A.R.), by a grant for a Specialized Program of Research Excellence (SPORE) in Lung Cancer (P50-CA70907), by gifts to the Division of Surgery and Anesthesiology from Tenneco and Exxon for the Core Laboratory Facility, by the UT M. D. Anderson Cancer Center Support Core Grant (CA16672) and by a sponsored research agreement with Introgen Therapeutics, Inc.

REFERENCES

Barbacid, M. (1987). Ras genes. *Annual Review of Biochemistry*, **56**, 779–779.

Bishop, J. M. (1991). Molecular themes in oncogenesis. *Cell*, **64**, 235–248.

Blaese, R. M., *et al*. (1995). T lymphocyte-directed gene therapy for ADA(−) SCID: initial trial results after 4 years. *Science*, **270**, 475–480.

Bookstein, R., *et al*. (1990). Suppression of tumorigenicity of human prostate carcinoma cells by replacing a mutated RB gene. *Science*, **247**, 712–715.

Bouvet, M., *et al*. (1998). Adenovirus-mediated wild-type *p53* gene transfer downregulates vascular endothelial growth factor expression and inhibits angiogenesis in human colon cancer. *Cancer Research*, **58**, 2288–2292.

Boveri, T. (1929). *The Origin of Malignant Tumors*. (Williams and Wilkins, Baltimore).

Cai, D. W., *et al*. (1993). Stable expression of the wild-type *p53* gene in human lung cancer cells after retrovirus-mediated gene transfer. *Human Gene Therapy*, **4**, 617–624.

Chen, S. H., *et al*. (1995). Combination gene therapy for liver metastasis of colon carcinoma *in vivo*. *Proceedings of the National Academy of Sciences of the USA*, **92**, 2577–2581.

Clayman, G. L., *et al*. (1998). Adenovirus-mediated *p53* gene transfer in patients with advanced recurrent head and neck squamous cell carcinoma. *Journal of Clinical Oncology*, **16**, 2221–2232.

Crystal, R. G. (1995). Transfer of genes to humans: early lessons and obstacles to success. *Science*, **270**, 404–410.

Fujiwara, T., *et al*. (1993). A retroviral wild-type p53 expression vector penetrates human lung cancer spheroids and inhibits growth by inducing apoptosis. *Cancer Research*, **53**, 4129–4133.

Fujiwara, T., *et al*. (1994a). Therapeutic effect of a retroviral wild-type p53 expression vector in an orthotopic lung cancer model. *Journal of the National Cancer Institute*, **86**, 1458–1462.

Fujiwara, T., *et al*. (1994b). Induction of chemosensitivity in human lung cancer cells *in vivo* by adenoviral-mediated transfer of the wild-type *p53* gene. *Cancer Research*, **54**, 2287–2291.

Georges, R. N., *et al*. (1993). Prevention of orthotopic human lung cancer growth by intratracheal instillation of a retroviral antisense K-ras construct. *Cancer Research*, **53**, 1743–1746.

Harris, C. C. (1996). Structure and function of the p53 tumor suppressor gene: clues for rational cancer therapeutic strategies. *Journal of the National Cancer Institute*, **88**, 1442–1455.

Jin, X., *et al*. (1995). Cell cycle arrest and inhibition of tumor cell proliferation by the $p16^{INK4}$ gene mediated by an adenovirus vector. *Cancer Research*, **55**, 3250–3253.

Kamijo, T., *et al*. (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell*, **91**, 649–659.

Kiem, H. P., *et al*. (1995). Gene therapy and bone marrow transplantation. *Current Opinions in Oncology*, **7**, 107–114.

Miller, N. and Vile, R. (1995). Targeted vectors for gene therapy. *FASEB Journal*, **9**, 190–199.

Mukhopadhyay, T., *et al*. (1991). Specific inhibition of K-ras expression and tumorigenicity of lung cancer cells by antisense RNA. *Cancer Research*, **51**, 1744–1748.

Nguyen, D. M., *et al*. (1996). Gene therapy for lung cancer: enhancement of tumor suppression by a combination of sequential systemic cisplatin and adenovirus-mediated p53 gene transfer. *Journal of Thoracic and Cardiovascular Surgery*, **112**, 1372–1377.

Nishizaki, M., *et al*. (1999). Recombinant adenovirus expressing wild-type *p53* is antiangiogenic: a proposed mechanism for bystander effects. *Clinical Cancer Research*, **5**, 1015–1023.

Owen-Schaub, L. B., *et al*. (1995). Wild-type human p53 and a temperature-sensitive mutant induce Fas/APO-1 expression. *Molecular and Cellular Biology*, **15**, 3032–3040.

Roth, J. A. (1996a) Clinical protocol: modification of mutant K-ras gene expression in non-small cell lung cancer (NSCLC). *Human Gene Therapy*, **7**, 875–889.

Roth, J. A. (1996b) Clinical protocol: modification of tumor suppressor gene expression and induction of apoptosis in non-small cell lung cancer (NSCLC) with an adenovirus vector expressing wildtype p53 and cisplatin. *Human Gene Therapy*, **7**, 1013–1030.

Roth, J. A. and Cristiano, R. J. (1997). Gene therapy for cancer: what have we done and where are we going? (Review). *Journal of the National Cancer Institute*, **89**, 21–39.

Son, K. and Huang, L. (1994). Exposure of human ovarian carcinoma to cisplatin transiently sensitizes the tumor cells for liposome-mediated gene transfer. *Proceedings of the National Academy of Sciences of the USA*, **91**, 12669–12672.

Spitz, F. R., *et al*. (1996). Adenoviral-mediated wild-type p53 gene expression sensitizes colorectal cancer cells to ionizing radiation. *Clinical Cancer Research*, **2**, 1665–1671.

Swisher, S. G., *et al*. (1999). Adenoviral-mediated p53 gene transfer in advanced non-small cell lung cancer. *Journal of the National Cancer Institute*, **91**, 763–771.

Swisher, S. G., *et al*. (2000). Gene therapy in lung cancer. *Current Oncology Reports*, **2**, 64–70.

Takahashi, T., *et al*. (1992). Wild-type but not mutant p53 suppresses the growth of human lung cancer cells bearing multiple genetic lesions. *Cancer Research*, **52**, 2340–2343.

Tang, B., *et al*. (1998). Transforming growth factor-b1 is a new form of tumor suppressor with true haploid insufficiency. *Nature Medicine*, **4**, 802–807.

Wang, J., *et al*. (1995). Apoptosis induced in human osteosarcoma cells is one of the mechanisms for the cytocidal effect of Ad5CMV-*p53*. *Cancer Gene Therapy*, **2**, 9–17.

Weinberg, R. A. (1991). Tumor suppressor genes. *Science*, **254**, 1138–1146.

Wilson, J. (1999). *1999 American Society of Gene Therapy: Presidential Address.* 2nd Annual Meeting of the American Society of Gene Therapy.

Xu, H. J., *et al*. (1996). Enhanced tumor suppressor gene therapy via replication-deficient adenovirus vectors expressing an N-terminal truncated retinoblastoma protein. *Cancer Research*, **56**, 2245–2249.

Zhang, Y., *et al*. (1995). *Her-2/neu*-targeting cancer therapy via adenovirus-mediated E1A delivery in an animal model. *Oncogene*, **10**, 1947–1954.

FURTHER READING

Bishop, J. M. (1991). Molecular themes in oncogenesis. *Cell*, **64**, 235–248.

Haber, D. A. (1997) Splicing into senescence: the curious case of p16 and p19ARF (Review) *Cell*, **91**, 555–558.

Harris, C. C. (1996). Structure and function of the p53 tumor suppressor gene: clues for rational cancer therapeutic strategies. *Journal of the National Cancer Institute*, **88**, 1442–1455.

Hunter, T. (1991). Cooperation between oncogenes. *Cell*, **64**, 249–270.

Kamijo, T., *et al*. (1997) Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell*, **91**, 649–659.

Pomerantz, J., *et al*. (1998). The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell*, **92**, 713–723.

Serrano, M., *et al*. (1996). Role of the INK4a locus in tumor suppression and cell mortality. *Cell*, **85**, 27–37.

Sherr, C. J. (1996). Cancer cell cycles (Review). *Science*, **274**, 1672–1677.

Tang, B., *et al*. (1998). Transforming growth factor-b1 is a new form of tumor suppressor with true haploid insufficiency. *Nature Medicine*, **4**, 802–807.

Weinberg, R. A. (1991). Tumor suppressor genes. *Science*, **254**, 1138–1146.

Websites

www.geocities.com/CollegePark/Lab/1580. An excellent summary of cell cycle regulatory pathways and apoptosis.

M. D. Anderson Cancer Center Department of Gene Therapy, www.mdanderson.org:80/depts/genetherapy.

Institute for Human Gene Therapy, www.med.upenn.edu/ihgt.

Chapter 99

Genetic Prodrug Activation Therapy (GPAT)

Iain A. McNeish and Nicholas R. Lemoine
Imperial College School of Medicine, London, UK

CONTENTS

- General Introduction
- Genetic Prodrug Activation Therapy (GPAT)
- Discussion

GENERAL INTRODUCTION

Few areas of research have raised as much interest and introspection, not to mention false optimism, as gene therapy. Indeed, few subjects can have produced quite so many philosophical editorials. The possibility of using DNA as a therapeutic tool first became theoretically feasible with the isolation and cloning of the genes responsible for inherited monogenetic disorders such as cystic fibrosis, Lesch–Nyhan syndrome and adenosine deaminase (ADA) deficiency associated with severe combined immune deficiency (SCID). These disorders could theoretically be cured by the introduction and expression of a normal functional copy of the faulty gene in the appropriate tissue, although the practical requirements for such gene therapy are formidable. After the gene has been isolated, its regulatory sequences must also be identified to ensure that expression of the transgene occurs in the appropriate tissue and at the appropriate time. Second, practical ways have to be found of delivering the gene to the requisite organ and, finally, once expression is achieved, it must continue indefinitely, to obviate the requirement for multiple and repeated treatments. Perhaps not surprisingly, initial clinical trials of gene therapy in cystic fibrosis and ADA-deficient SCID have met with very limited success.

At a superficial level, cancer is an even less attractive candidate for gene therapy than either cystic fibrosis or ADA-deficient SCID: the progression from normal tissue to invasive malignancy may involve up to six or more separate genetic events, all of which, in theory, would have to be corrected to reverse the malignant phenotype. Similarly, it would also be necessary to restore normal gene function to 100% of the cells within a tumour population, which is impractical with current vector technology.

However, novel gene therapy strategies for cancer have evolved that do not rely on gene complementation, and which thus circumvent some, but by no means all, of the difficulties listed above. One such strategy is genetic prodrug activation therapy.

GENETIC PRODRUG ACTIVATION THERAPY (GPAT)

GPAT, which is represented schematically in **Figure 1**, has two fundamental steps. The first is the delivery to tumour cells, usually via a recombinant virus, of the gene encoding a nonhuman enzyme. The enzyme has been selected for its ability to convert a nontoxic prodrug into a cytotoxic species, which will then kill the cell in which it has been formed. The production of active drug by the tumour cells themselves should allow the generation of higher drug concentrations within the tumour microenvironment than could be achieved by systemic administration alone, and thus lead to a higher therapeutic index.

The concept of GPAT arose from studies in antibody-directed enzyme–prodrug therapy (ADEPT) where the enzyme itself (rather than its gene) was delivered to tumour cells by means of an antibody directed against a tumour-specific antigen. The advantages of GPAT over ADEPT include the fact that the enzyme is generated within the tumour cell itself, rather than being present in the interstitium, thus generating higher concentrations of the active species inside the cell and allowing ready access to any cofactors that may be required for enzyme activity. Also, the inherent immunogenicity of antibody–enzyme conjugates may preclude multiple administrations.

The second important feature of GPAT is the so-called bystander effect. This describes the ability of active species generated in one cell to kill neighbouring, non-transduced cells, thereby eliminating the need to transduce every tumour cell, which is clearly impractical with current vector technology.

The ideal enzyme would consist of a single polypeptide species of reasonably low molecular mass and independent of post-translational modifications. Fundamentally, the catalytic activity or substrate specificity must be distinct from any human enzyme and so those in use are derived predominantly from bacteria and viruses. The desirable parameters of potential enzyme–prodrug combinations are a high differential toxicity of the active species relative to

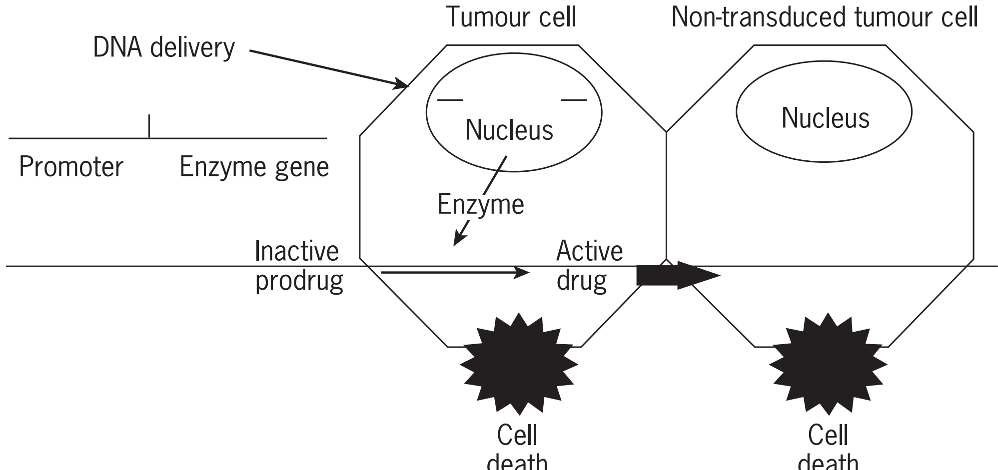


Figure 1 Genetic prodrug activation therapy.

the prodrug and a low K_M and a high K_{cat}, which will maximize production of the active species at any given concentration of prodrug and enzyme. The physicochemical properties of the prodrug and its cytotoxic species, such as lipophilicity, will influence tissue distribution, cellular uptake and bystander killing. The half-life of the activated cytotoxic agent will also affect its distribution and efficacy: a longer half-life should allow a more homogeneous distribution within a tumour, but this may be offset by a greater potential to diffuse into the vascular compartment, increasing the systemic concentration of the cytotoxic agent.

Many different potential enzyme–prodrug combinations have been described, but the general principles can be illustrated by consideration of the following systems.

Thymidine Kinase–Ganciclovir

In 1986, Moolten described the successful *in vitro* and *in vivo* sensitization of murine tumour cells to the anti-herpes drug ganciclovir (GCV), via expression of the herpes simplex virus 1 (HSV) thymidine kinase (*tk*) gene (Moolten, 1986). The HSV, but not mammalian, tk enzyme is capable of phosphorylating various nucleoside analogues, including ganciclovir and its relative acyclovir. Treatment of HSV tk-positive cells with ganciclovir leads to the formation of ganciclovir triphosphate (**Figure 2**), a potent antimetabolite that interrupts DNA synthesis by erroneous incorporation as a false nucleotide. Thus, in Moolten's original work, there was inhibition of growth of tk-positive cells following 3 days of incubation with GCV at concentrations of 10^{-7} mol/L, whereas tk-negative cells were not inhibited until concentrations of GCV reached 10^{-4} mol/L.

Since the original description of the effectiveness of HSV tk/GCV, there have been many reports of its use in the treatment, both *in vitro* and *in vivo*, of multiple tumour types, including breast, glioma, pancreatic, mesothelioma and colon, with the *tk* gene delivered by retrovirus, adenovirus, naked DNA and liposomes.

Ganciclovir
↓ Thymidine kinase
Ganciclovir monophosphate
↓ Guanylate kinase
Ganciclovir diphosphate
↓ Cellular kinases
Ganciclovir triphosphate

Figure 2 The activation of ganciclovir by herpes simplex virus thymidine kinase.

It was noticed in the original description of the effect of HSV tk–GCV combination that the mixing of tk-positive and tk-negative cells at high density in a ratio of 1:9

resulted in almost complete eradication of the tk-negative cells upon treatment with GCV at a concentration of 4×10^{-6} mol/L. In contrast, when the cells were plated at the same ratio at low density, there was almost complete survival of the tk-negative cells, implying that direct cell to cell contact between the enzyme expressing and non-expressing cells was capable of transferring sensitivity to GCV between cells, a process initially described as 'metabolic cooperation' (Moolten, 1986). It was subsequently shown in subcutaneous adenocarcinoma and fibrosarcoma models in mice that only 10% tk-positive cells were required for marked growth inhibition and that the normal tissue adjoining and overlying the tumours was unaffected by administration of GCV (Culver *et al*., 1992). In the same study, the delivery of HSV *tk* gene via injection of retroviral producer cells into experimental rat glioma produced complete regression in 11/14 animals on exposure to GCV. It has also been noted that a bystander effect can exist within the peritoneal cavity. Injecting tk-positive and negative sarcoma cells at a ratio of 1:9 into the peritoneum of mice resulted in significant extension of survival following treatment with GCV compared with those mice that received 100% tk-negative cells alone. If the ratio was 50:50, some of the mice survived long term (>70 days) (Freeman *et al*., 1993). Interestingly, Freeman *et al*. also noted that it was possible to prolong the survival of pre-existing tk-negative intraperitoneal tumours by injecting tk-positive tumour cells and GCV, without any attempt at gene transfer. However, other groups have shown that injecting tk-positive 3T3 fibroblasts into experimental gliomas in rats does not produce any bystander effect.

Samejima's group confirmed that cell–cell contact, or at least close proximity, was required for bystander killing in rodent fibroblasts expressing tk and that the process could be inhibited by coadministration of forskolin, the activator of adenylate cyclase. They also demonstrated that there was evidence of apoptosis in the cells killed by the bystander effect and suggested that phagocytosis by nonenzyme-expressing cells of apoptotic bodies released from dying cells could explain the mechanism of the bystander effect (Samejima and Merulo, 1995). However, it has become evident from more recent work that gap junctional activity is of fundamental importance to the bystander effect in the HSV tk–GCV enzyme–prodrug combination. It was demonstrated that the bystander effect seen in some tumour cell lines, such as SKHep J hepatocellular carcinoma cells (Elshami *et al*., 1996), which have little gap junctional activity, increased dramatically when transfected with connexin 43 or connexin 32, which are major gap junctional proteins. This finding could explain why nontransduced normal tissue immediately adjacent to tumours with little gap junctional connection with those tumour cells, was not affected by administration of GCV, whereas the tumour cells were rapidly eliminated.

There appears to be an additional mechanism which can give rise to a bystander effect *in vivo*. In a model of pulmonary metastases of the murine melanoma B16 line, it was noticed that the antitumour effect of HSV tk–GCV was markedly greater in immunocompetent mice than in immunodeficient animals. This effect was not due to the immunogenicity of tk itself and subsequent rechallenge of GCV-treated mice showed that some protection against wildtype tumour had been generated in the immunocompetent mice. This implies that the death of tumour cells following exposure to GCV results in the establishment of a systemic tumour-specific immunity (Vile *et al*., 1994). These findings were confirmed in murine colorectal carcinoma and rat hepatoma models and have also been shown for the cytosine deaminase–5-fluorocytosine enzyme–prodrug combination (see below). Vile *et al*. have suggested further that the protective immunity seen in the B16 Murine melanoma model is due to a T_{h1}-type response (Vile *et al*., 1997), that the induction of heat-shock protein 70 expression is a requirement for this immunity and that immunity is established after necrotic but not apoptotic cell death (Melcher *et al*., 1998).

More recently, it has been suggested that there is an opposite, 'good samaritan,' effect whereby the tk-positive cells are protected from the effects of the toxic metabolites of GCV via gap junctional activity, thereby allowing a more prolonged production of those toxic metabolites. The presumed mechanism is that the gap junctions allow rapid passage of the activated GCV away from the tk-positive cell, thereby sparing that cell (Wygoda *et al*., 1997).

Despite there being at least 20 clinical trials of HSV tk–GCV approved by the beginning of 1997, only a very small number have been published (Izquierdo *et al*., 1996; Ram *et al*., 1997; Sterman *et al*., 1998). In the first (Izquierdo *et al*., 1996), producer cells releasing tk-encoding retroviruses were implanted stereotactically into recurrent primary brain tumours in five patients, who then received 14 days of intravenous GCV at a dose of $5\,\mathrm{mg\,kg^{-1}}$ twice daily. One tumour in one patient showed a partial response, with progressive disease in all other injected tumours. The second trial protocol was very similar (Ram *et al*., 1997), with retroviral producer cells releasing tk-encoding virions injected into 19 tumours in 15 patients with recurrent primary and metastatic intracerebral tumours. Two tumours were resected 7 days after producer cell injection and *in situ* hybridization revealed only very small clusters of tk-positive cells. Accordingly, in the remaining patients who were treated with GCV twice daily for 14 days, there was a complete response in only two nodules (both in the same patient), a partial response in three nodules with no response in the remainder and no responses in any uninjected lesion. All the nodules that responded were under 2 mL in volume prior to treatment. On a more positive note, there were no severe adverse events. A similar, but larger scale multicentre trial has been published more recently (Shand *et al*., 1999). Forty-eight patients with

recurrent glioblastoma multiforme (GBM) were again injected with producer cells releasing tk-encoding retroviruses following resection of the recurrent tumour. GCV administration followed 14 days later at a dose of $5\,mg\,kg^{-1}$ intravenously twice daily for another 14 days. Fourteen of 48 patients had demonstrable vector DNA in peripheral blood (as assessed by polymerase chain reaction), thought to be secondary to transduction of lymphocytes within the resected tumour bed, which then migrated into the peripheral circulation. Disappointingly, the median survival for the patients was only 8.6 months, which is no better than would be expected in a population of patients who had surgical resection of recurrent GBM without any other treatment. Ten patients had tk-positive residual tumours present at post mortem examination, indicating that, whilst there had been tumour cell transduction, subsequent GCV treatment had failed to kill them, either because GCV penetrated poorly into the CNS or because the cells were not in the correct phase of the cell cycle at the time of treatment (see later).

In the first trial reported of this enzyme–prodrug combination in non-CNS tumours, an adenovirus encoding HSV tk was administered at doses of between 1×10^9 and 1×10^{12} plaque-forming units into the pleural cavity of 20 patients with pleural mesothelioma, followed by 14 days of GCV again at a dose of $5\,mg\,kg^{-1}$ twice daily (Sterman *et al*., 1998). Although no formal attempts were made to establish whether there were any clinical responses, there was evidence of gene transfer in 11 of 20 patients in a dose-related manner, again with minimal toxicity.

There has now been a reported phase I trial in recurrent localized prostate cancer (Herman *et al*., 1999) in which an adenovirus encoding HSV tk was injected directly into the prostate of 18 men at doses of between 1×10^8 and 1×10^{11} infectious units. After 24 h, 14 days of intravenous GCV (dose as in previous trials) was commenced. Three patients had an objective response (fall in serum prostate-specific antigen of at least 50%), the longest lasting 12 months.

There were only two severe adverse reactions, both at the highest dose level. One was thrombocytopenia, which resolved after 5 days, and the other was hepatotoxicity, with transient elevation of bilirubin, transaminases and alkaline phosphatase. There has also been a report of severe hepatotoxicity in a rat colorectal carcinoma model following portal vein delivery of adenoviruses encoding the *tk* gene and GCV administration. This implies that activated GCV may be toxic to noncycling hepatocytes, as this toxicity was not seen in rats that received the adenoviral tk without GCV or in those rats that received a control adenovirus expressing β-galactosidase, followed by GCV (van der Eb *et al*., 1998). Another potential safety concern was highlighted recently in a syngeneic rat glioma model (Dewey *et al*., 1999). Following adenovirally-delivered HSV tk and treatment with GCV, the brains of long-term (>3 months) surviving rats demonstrated extensive inflammation, with microglial and T lymphocyte infiltration, and also widespread demyelination. There was also evidence of persistent tk expression 3 months after a single injection of 2×10^7 infectious units of the adenovirus, implying both that the duration of transgene expression after adenoviral delivery may be much longer *in vivo* than previously imagined and also that the administration of GCV in clinical trials of tk–GCV should, perhaps, continue for much longer than the 14 days that most trials currently utilize.

Following reports of incomplete eradication of tumours following *tk* gene transfer due partially to poor gene transfer and poor GCV penetration, and given the probable benefits of immune system involvement in bystander killing *in vivo*, there have been recent attempts to increase the potency of the HSV tk–GCV system by combining it with immunotherapy. There are reports that IL-2, IL-4 and IL-7 can combine successfully with tk–GCV to increase the elimination of colorectal metastases, gliomas and non-small cell lung cancer models, respectively. The significance of these results is that prodrug-mediated killing of one tumour nodule may allow the elimination of distant nodules that do not express the prodrug-converting enzyme. There have also been reports of combined cytosine deaminase–5-fluorocytosine and HSV *tk*–GCV gene therapy being more effective than either alone in eliminating murine mammary adenocarcinoma tumours.

Others have attempted to improve the efficacy of tk–GCV by utilizing the thymidine kinase gene of other herpes virus species (Loubiere *et al*., 1999). There is some evidence that equine herpes virus tk mediates more rapid phosphorylation of GCV than that of HSV. Similarly, the thymidine kinase and phosphotransferase homologues of human herpesvirus-8 are both capable of phosphorylating GCV and may have potential for use in GPAT (Cannon *et al*., 1999). Other enhancements include the use of alternative prodrugs such as famciclovir and penciclovir that can also be activated by HSV tk, but which are orally active. This would facilitate a much longer duration of treatment than is currently possible with intravenous GCV, which would permit more cells to enter S phase and thereby become sensitive to the effects of the activated prodrug.

Cytosine Deaminase–5-Fluorocytosine

The second most commonly employed enzyme–prodrug combination utilizes the cytosine deaminase (*cd*) gene, which can be isolated from fungi and bacteria, and deaminates the antifungal drug 5-fluorocytosine (5-FC) into 5-fluorouracil (**Figure 3**), the main chemotherapy drug used in the treatment of gastrointestinal malignancies for 30 years.

5-FU itself is converted into 5-fluorouridine 5′-triphosphate and 5-fluoro-2′-deoxyuridine 5′-monophosphate, which function as potent inhibitors of both DNA and RNA synthesis via inhibition of thymidylate synthase and

$$\text{5-Fluorocytosine} + H_2O \xrightarrow{\text{Cytosine deaminase}} \text{5-Fluorouracil} + NH_3$$

Figure 3 The conversion of 5-fluorocytosine into 5-fluorouracil by the action of cytosine deaminase.

erroneous incorporation as false bases. That there is some inhibition of RNA synthesis as well as DNA synthesis should also produce some toxicity to quiescent cells, although the toxicity of 5-FU still depends critically upon the length of S phase. Initial reports in transfected or retrovirally transduced murine fibroblast lines demonstrated effective killing of cd-positive cells on exposure to 5-FC at a concentration of 5–10 $\mu g\,mL^{-1}$, whilst the parental cd-negative cells were unaffected by concentrations in excess of 100 $\mu g\,mL^{-1}$. However, in cell mixing experiments, there appeared to be selective killing of the cd-positive cells, with sparing of the cd-negative cells, implying no bystander effect (Mullen *et al*., 1992). A group that had independently cloned the *cd* gene from *Escherichia coli* confirmed the sensitization of cd-positive WiDR human colorectal carcinoma cells to 5-FC via its conversion to 5-FU (Huber *et al*., 1993), but showed that there was marked bystander killing *in vitro* when 33% of cells were cd positive and that this effect did not rely upon cell–cell contact. Also, there was a significant reduction in growth rates of mixed cd-positive and -negative subcutaneous WiDR xenografts in nude mice on treatment with 5-FC. It was possible to demonstrate this effect with only 2% cd-positive cells, implying a very powerful bystander effect *in vivo* with this enzyme–prodrug combination (Huber *et al*., 1994).

The same study also demonstrated that it was possible to generate a concentration of 5-FU of $>400\,\mu mol\,L^{-1}$ within cd-positive tumours following systemic 5-FC administration. There is no apparent explanation for the discrepant results between the two groups in terms of bystander activity, although other groups have subsequently recorded bystander activity for cd–5-FC. Equally, it has also been reported that *cd* gene delivery to normal liver can generate sufficient 5-FU to cause the regression of adjacent cd-negative colorectal metastases, without significant toxicity to those hepatocytes (Topf *et al*., 1998), which implies both a powerful bystander effect and lack of toxicity of 5-FU to noncycling cells.

An attempt was made to compare the efficacy of tk–GCV and cd–5-FC in WiDR cells and it was demonstrated that, on exposure to GCV, there was no reduction in size of subcutaneous xenografts that contained 10% tk-positive cells, yet there was a marked antitumour effect in cd-positive tumours on exposure to 5-FC with only 4% of cells enzyme positive (Trinh *et al*., 1995). However, the poor bystander effect for the tk tumours can be explained by the electron microscope finding that WiDR tumours do not express gap junctions.

Although the protocols of several clinical trials of cd–5-FC have been approved, only one trial has actually been completed (Pandha *et al*., 1999), in which plasmid DNA encoding cd under the control of the c-*erb*B-2 proximal promoter was injected directly into cutaneous metastases of 12 women with recurrent breast cancer positive for c-*erb*B-2. Overexpression of c-*erb*B-2 is seen in approximately 20% of breast carcinomas and is associated with poor prognosis. There was evidence of cd expression in 11 of 12 tumour nodules injected and transgene expression was limited to c-*erb*B-2 positive cells, indicating that transcriptional targeting using this promoter is feasible in human patients. Only eight of the 12 patients received 5-FC (200 $mg\,kg^{-1}$ per day as a 48-h infusion) and there was evidence of some local tumour response in two of these eight patients. There were also two minor responses in patients who received the *cd* gene but no 5-FC, which may reflect an immune response to the injected plasmid DNA.

As with HSV tk–GCV, there are now reports of combining cd–5-FC with immunotherapy. One recent report demonstrated that adenoviral GM-CSF and adenoviral *cd* gene transfer to melanomas in mice resulted in greater growth inhibition than with either cd or GM-CSF alone and that the combination also resulted in greater protective immunity to rechallenge with parental cells (Cao *et al*., 1998). The majority of published data on cd–5-FC has come from use of the *E. coli cd* gene. However, the cytosine deaminase of *Saccharomyces cerevisiae* is capable of deaminating 5-FC at least eight times faster than the *E. coli* enzyme and has a K_M 44-fold lower, suggesting that it would produce 5-FU much more efficiently than the bacterial enzyme (Hamstra *et al*., 1999).

Nitroreductase–CB1954

The active species of both tk–GCV and cd–5-FC are antimetabolites, which means that they will be predominantly toxic to cells that are replicating their DNA. Even within a rapidly growing tumour nodule, the proportion of cells that will be actively dividing may be as low as 6% (Tubiana and Malaise, 1976). The large proportion of cells resting in G_0 has been proposed to be a major factor underlying resistance to GCV. In one study, tumours outgrew 30 days of continuous GCV treatment, but remained sensitive on retreatment, indicating that cells can remain in G_0 for long periods and that acquired resistance to the prodrug was not the cause of the original regrowth (Golumbek *et al*., 1992). This potential shortcoming and the realization that enzyme-catalysed activation of alkylating agents from nontoxic prodrugs could provide some advantages over antimetabolites have prompted the search for alternative enzyme–prodrug combinations, one of the

most promising of which involves the prodrug 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB1954). CB1954 is a weak monofunctional alkylating agent which originally provoked interest in the 1960s following the discovery that Walker rat carcinosarcoma cells were extremely sensitive to CB1954-mediated killing (Cobb *et al*., 1969). The origin of the Walker cell sensitivity was later found to lie in the expression of DT diaphorase, a FAD-containing dehydrogenase that employs either NADH or NADPH as a cofactor, and which catalyses the bioreduction of CB1954 to its 4-hydroxylamino derivative, which is then converted by thioesters such as acetyl coenzyme A into a powerful bifunctional alkylating agent (Knox *et al*., 1988). The activated cytotoxic species is not phase specific and can kill noncycling cells (Bridgewater *et al*., 1995). Human DT diaphorase performs this reduction much less efficiently than the rat enzyme, which may explain the lack of efficacy of CB1954 in human subjects in an unpublished pilot clinical trial in the early 1970s. Recent studies have indicated that the molecular basis of the difference between the ability of human and rat DT diaphorase to reduce CB1954 may lie at residue 104, tyrosine in the rat enzyme and glycine in the human. Cells with the capacity to bioreduce CB1954 have been demonstrated to be 10^5 times more sensitive on a dose basis than those which do not (Roberts *et al*., 1986). An *E. coli* nitroreductase (NTR) was identified that is capable of reducing CB1954 and was shown to be a monomeric FMN-containing flavoprotein with a molecular mass of 24 kDa that also employed either NADH or NADPH as a cofactor. By comparison of amino acid sequences, it was also shown that NTR has homology to the 'classical nitroreductase' of *Salmonella typhimurium* and a nitroreductase in *Enterobacter cloacae*. It was further demonstrated that NTR reduced CB1954 up to 60 times more rapidly than rat DT diaphorase (Knox *et al*., 1992), although both the 4-hydroxylamino and the less toxic 2-hydroxylamino derivatives are generated in equal proportions (**Figure 4**).

NTR was selected for use in conjunction with CB1954, initially for antibody-directed enzyme–prodrug therapy, and subsequently for GPAT once the gene encoding NTR, *nfnB*, was identified and cloned from the B strain of *E. coli*.

Figure 4 The bioactivation of CB1954 by *E. coli* nitroreductase.

Studies by several groups have confirmed the prediction that mammalian cells expressing NTR would be sensitized to CB1954 and have demonstrated the existence of a bystander effect with this system (Bridgewater *et al*., 1995), although other studies have reported little or no bystander effect (Clark *et al*., 1997). Initial *in vivo* work was limited to two studies in transgenic mice. In one, NTR was expressed selectively in secretory epithelial cells of the mammary gland from an ovine β-lactoglobin promoter (Clark *et al*., 1997) and treatment with CB1954 resulted in ablation of the mammary epithelial tissue, with sparing of the adjacent myoepithelial cells, which were NTR negative, implying no bystander effect. However, in the other study (Drabek *et al*., 1997), in which NTR was expressed in T cells under the control of elements from the human *CD2* locus, CB1954 administration resulted in severe reduction in total cell numbers within the thymus with evidence of extensive apoptosis. There was also extensive cell reduction within the spleen, which contained a large number of non-T cells, implying a degree of bystander killing *in vivo*. This study also demonstrated that high doses of CB1954 (50 mg kg^{-1} day^{-1} for 5 days) were also very toxic to Balb C and C3H/He mice and there have been some attempts to discover less toxic alternative prodrugs that can also be activated by NTR.

Recently, the first demonstration of the efficacy of NTR/CB1954 in murine tumour models was published (McNeish *et al*., 1998), in which bearing NTR-expressing subcutaneous and intraperitoneal tumour cells were cured following exposure to two bolus doses of CB1954. The same study also indicated that a cisplatin-resistant ovarian tumour cell line remains as susceptible to the NTR-dependent cytotoxicity of CB1954 as parental cells, which has importance in any potential clinical application of NTR–CB1954, especially in the treatment of ovarian cancer.

As with tk–GCV and cd–5-FC, there is evidence that CB1954-mediated killing of NTR-positive tumours *in vivo* leads to the establishment of antitumour immunity. Again, this effect is not dependent upon the immunogenicity of the enzyme and is cell-type specific. There is, however, convincing evidence that NTR-positive cells killed following exposure to CB1954 undergo an apoptotic rather than necrotic death (Cui *et al*., 1999), implying that the previous observation with tk–GCV that necrotic rather than apoptotic death is required for the development of antitumour immunity following prodrug delivery may not be a universal phenomenon.

Data on the delivery of the *NTR* gene to murine tumour models using recombinant viruses are limited. However, one as yet unpublished study demonstrates that the delivery of a recombinant adenovirus encoding the *NTR* gene to nude mice bearing intraperitoneal SUIT2 pancreatic tumour cells followed by two doses of CB1954 leads to a significant increase in animal survival (Weedon *et al*., 2000). Interestingly, this study shows that delivery of the adenovirus itself into the peritoneal cavity of nude mice without prodrug administration leads to a small, but significant, increase in animal survival. A similar result was noted when recombinant adenoviruses were administered into the pleural cavity of nude mice in a model of malignant mesothelioma (Lan *et al*., 1997). However, these findings were not confirmed in models utilizing SCID mice, implying that natural killer (NK) cells, present in nude but not SCID mice, may be responsible for the observed antitumour effect of recombinant adenoviruses alone.

Others

There are many other potential enzyme–prodrug systems, some of which evolved from ADEPT strategies. Some of the more promising ones are listed in **Table 1**.

The human thymidine phosphorylase–5′-deoxy-5-fluorouridine, VZV thymidine kinase–araM and *E. coli* purine nucleoside phosphorylase–9-(β-D-2-deoxyerythropentofuranosyl)-6-methylpurine combinations will all generate toxic antimetabolite agents which are likely to meet with the same potential hurdles as tk–GCV and cd–5-FC. However, the other combinations all generate novel active species that are toxic to both cycling and noncycling cells and merit further discussion.

One possible advantage that the carboxypeptidase G2 (CPG2)-mediated activation of 4-[(2-chloroethyl)(2-mesyloxyethyl)amino] benzoylglutamic acid (CMDA) has over other combinations is that the active species produced by CPG2 requires no further modification by cellular enzymes, which contrasts with 5-FU and both activated CB1954 and GCV: if the enzymes required for further drug modification become deficient in tumour cells, this could lead to resistance to the activated drug. Like activated CB1954, the active species of CMDA is an alkylating agent, toxic to noncycling cells and less likely to induce resistance. One potential disadvantage of CPG2 is that the enzyme in its natural configuration is secreted and removal of the signal peptide to is required to prevent this. There is *in vitro* evidence that expression of this altered CPG2 is capable of sensitizing a range of ovarian and colorectal carcinoma cells to CMDA and that a modest bystander effect exists (Marais *et al*., 1996).

Cytochrome P450 isoenzyme 4B1 (CYP4B1) was first identified as an activator of the mould toxin 4-ipomeanol following a mysterious cattle pulmonary illness. It has also been shown to activate 2-aminoanthracene, although the precise nature of both active species is unknown. However, glioma cells stably expressing the rabbit isoform of the enzyme are sensitized both *in vitro* and *in vivo* to both prodrugs, with evidence of single-strand DNA breaks and an adequate bystander effect (Rainov *et al*., 1998).

Another cytochrome P450 enzyme, the human isoenzyme 1A2 (CYP1A2), is capable of activating the analgesic drug acetaminophen (paracetamol), which is used world-wide as an analgesic drug, but can be hepatotoxic when taken in overdose. There are several pathways

Table 1 Other potential enzyme–prodrug combinations for use in GPAT

Enzyme	Prodrug	Active species	Reference
Human thymidine phosphorylase	5′-Deoxy-5-fluorouridine	5-FU	Evrard *et al*. (1999)
Varicella zoster virus (VZV) thymidine kinase	9-(β-D-Arabinofuranosyl)-6-methoxy-9*H*-purine (araM)	AraATP	Huber *et al*. (1991)
E. coli purine nucleoside phosphorylase	9-(β-D-2-Deoxy erythropentofuranosyl) -6-methylpurine	6-Methylpurine	Sorscher *et al*. (1994); Hughes *et al*. (1995)
Carboxypeptidase G2	4-[2-Chloroethyl)(2-mesyloxyethyl)amino] benzoylglutamic acid (CMDA)	4-[2-Chloroethyl)(2-mesyloxyethyl) amino]benzoic acid	Marais *et al*. (1996)
Rabbit cytochrome	2-Aminoanthracene	Unknown	Rainov *et al*. (1998)
P450 isoenzyme 4B1	4-Ipomeanol	Unknown	
Human cytochrome P450 isoenzyme 1A2	Acetaminophen (paracetamol)	*N*-Acetyl benzoquinonimine (NABQI)	Thatcher *et al*. (1999)
Human carboxylesterase	7-Ethyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothecin (CPT-11; irinotecan)	7-Ethyl-10-hydroxy camptothecin (SN38)	Kojima *et al*. (1998)
Linamarase	2-Hydroxyisobutyronitrile-β-D-glucopyranoside (linamarin)	Cyanide	Cortés *et al*. (1998)

of acetaminophen elimination, one of which involves CYP1A2, which oxidizes the drug into the toxic metabolite *N*-acetylbenzoquinonimine (NABQI). Normally, NABQI is detoxified by direct conjugation with reduced glutathione, but in overdose, however, the glutathione conjugation pathway is saturated, leading to excess NABQI and direct hepatic toxicity. An early report indicates that the expression of CYP1A2 in Chinese hamster V79 cells dramatically increases their sensitivity to acetaminophen-mediated toxicity (Thatcher *et al*., 1999). There is evidence of bystander killing of non-CYP1A2-expressing V79 cells at low concentrations of acetaminophen when mixed with only 5% of CYP1A2-positive cells. This effect is not cell-type specific, with evidence of killing of human ovarian carcinoma SKOV3 and colon carcinoma HCT116 cells when mixed with CYP1A2-expressing V79 cells.

The chemotherapeutic drug irinotecan, also known as CPT-11 {7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin}, is active against lung, cervical, ovarian and colorectal carcinomas. As with 5-FU, it should strictly be classified as a prodrug as its piperidino side chain is cleaved by the enzyme carboxylesterase (CE) to reveal the active species, SN-38 (7-ethyl-10-hydroxy-camptothecin). Uniquely amongst the other active species mentioned above, SN-38 functions as an inhibitor of mammalian DNA topoisomerase I, an enzyme whose functions include releasing the torsional stress of supercoiled DNA. However, SN-38 is insoluble and is therefore impractical as a chemotherapeutic agent in its own right. Administration of irinotecan is frequently associated with severe, but unpredictable, toxicity, especially diarrhoea. This may be due to a genetically determined person-to-person variability in the expression of the activating CE enzyme resulting in highly variable levels of SN-38 production. Expression of high levels of CE by tumour cells would permit much lower doses of irinotecan to be administered, with resultant reduction in toxicity. It has been shown that adenoviral delivery of the human *CE* gene to A549 lung carcinoma cells increases both SN-38 production and sensitivity to irinotecan, and can produce growth delay in subcutaneous A549 xenografts in nude mice (Kojima *et al*., 1998).

The enzyme–prodrug combination linamarase and 2-hydroxyisobutyronitrile-β-D-glucopyranoside (linamarin) is of interest because the enzyme derives not from a virus or fungus, but from the cassava plant, *Manihot esculenta* Crantz. Linamarase is a β-glucosidase that hydrolyses linamarin into glucose and acetone cyanohydrin. The latter is unstable at pH >6 and spontaneously degrades into acetone and hydrogen cyanide. The hydrolysis of linamarin does not take place in normal mammalian tissue and the prodrug is excreted unchanged in the urine. Expression of linamarase in range of human and rodent cells has been shown to increase both their sensitivity to linamarin and the production of cyanide.

The cyanide ion (CN^-) is freely diffusible and HCN is a gas, which would suggest that any bystander effect of this system would not rely upon gap junctional activity or

cell–cell contact. There is *in vitro* evidence that a 50 : 50 mixture of linamarase-expressing and nonexpressing murine ψCRIP cells is completely eradicated on exposure to concentrations of linamarin that are nontoxic to the parental cells (Cortés *et al*., 1998).

Given the toxicity of cyanide and the diffusible nature of the active species, there would be considerable concern that the use of this enzyme–prodrug combination *in vivo* would be associated with extensive toxicity. In a rat glioma model, direct infusion of linamarin into the brains of rats bearing linamarase-positive tumours resulted in complete tumour eradication with no discernible toxicity to the animals. However, there was no attempt to deliver the linamarin systematically, as there is evidence that bacteria present in normal gut flora of both rodents and humans contain β-glucosidases that are capable of releasing cyanide from linamarin. This problem would have to be overcome safely if this enzyme–prodrug combination were ever to reach clinical trials.

DISCUSSION

Genetic prodrug activation therapy has great potential as a gene therapy strategy for the treatment of malignant disease. The two most studied combinations (herpes simplex virus thymidine kinase–ganciclovir and cytosine deaminase–fluorocytosine) have produced much promising *in vitro* and preliminary animal data. However, early clinical trials have been disappointing, with low rates of gene transfer and few objective clinical responses. These disappointments, in addition to the theoretical shortcomings of antimetabolites as active species, have driven the investigation into alternative enzyme–prodrug combinations. However, none of the novel combinations has yet been subjected to the reality of a clinical trial, where trial shortcomings may also be exposed.

In a defence of gene therapy, early clinical trials of any new treatment involve patients with recurrent and metastatic disease that is, by definition, refractory to current conventional therapy. It is, therefore, little surprise that there have been few clinical responses in published trials. However, the degree of prior expectation has ensured that these results are heralded as a wholesale failure of gene therapy. Gene therapy, especially as a treatment for malignant disease, is a technique in its infancy (a excuse that has been used for over a decade, but surely cannot be valid indefinitely) and it is very unlikely that any one gene therapy strategy will ever produce dramatic cures for all types of malignant disease. Much more likely, combined gene therapy strategies will be used in conjunction with existing modalities in the treatment of residual postsurgical disease. To this end, there is ongoing research using enzyme–prodrug therapy as a radiosensitizer (Szary *et al*., 1997) and even double enzyme–prodrug therapy (tk–GCV and cd–5-FC) with the genes delivered by a selectively replicating adenovirus and used in conjunction with radiotherapy (Freytag *et al*., 1998).

One of the other persistent criticisms of gene therapy, as highlighted by the Varmus report, has been the poor design of clinical trials, which has reduced the amount of useful information that can be gleaned from any negative results. Therefore, future trials of gene therapy must be sophisticated in design with clear and distinct scientific endpoints if the potential of this existing technology is to realized.

See also the chapters on *Gene Therapy Models*; *Genetic and Cellular Vaccines*; *Antisense and Ribozyme Therapy*; *Gene Therapy – Tumour Suppressor Replacement/Oncogene Suppression*; *Translational Research*.

REFERENCES

Bridgewater, J., *et al*. (1995). Expression of the bacterial nitroreductase enzyme in mammalian cells renders them selectively sensitive to killing by the prodrug CB1954. *European Journal of Cancer*, **31A**, 2362–2370.

Cannon, J., *et al*. (1999). Human herpesvirus 8-encoded thymidine kinase and phosphotransferase homologues confer sensitivity of ganciclovir. *Journal of Virology*, **73**, 4786–4793.

Cao, X., *et al*. (1998). Adenovirus-mediated GM-CSF gene and cytosine deaminase gene transfer followed by 5-fluorocytosine administration elicit more potent antitumor response in tumor-bearing mice. *Gene Therapy*, **5**, 1130–1136.

Clark, A., *et al*. (1997). Selective cell ablation in transgenic mice expressing *E. coli* nitroreductase. *Gene Therapy*, **4**, 101–110.

Cobb, L., *et al*. (1969). 2,4-Dinitro-5-ethyleneiminobenzamide (CB1954): a potent and selective inhibitor of growth of the Walker carcinoma 256. *Biochemical Pharmacology*, **18**, 1519–1527.

Cortés, M., *et al*. (1998). Successful use of a plant gene in the treatment of cancer *in vivo*. *Gene Therapy*, **5**, 1499–1507.

Cui, W., *et al*. (1999). Nitroreductase-mediated cell ablation is very rapid and mediated by a p53-independent apoptotic pathway. *Gene Therapy*, **6**, 764–770.

Culver, K., *et al*. (1992). *In vivo* gene transfer with retroviral vector-producer cells for treatment of experimental brain tumours. *Science*, **256**, 1550–1552.

Dewey, R., *et al*. (1999). Chronic brain inflammation and persistent herpes simplex virus thymidine kinase expression in survivors of syngeneic glioma treated by adenovirus-mediated gene therapy: implications for clinical trials. *Nature Medicine*, **5**, 1256–1263.

Drabek, D., *et al*. (1997). The expression of bacterial nitroreductase in transgenic mice results in specific cell killing by the prodrug CB1954. *Gene Therapy*, **4**, 93–100.

Elshami, A., *et al*. (1996). Gap junctions play a role in the bystander effect of the herpes simplex virus thymidine kinase/ganciclovir system *in vitro*. *Gene Therapy*, **3**, 85–92.

Evrard, A., *et al*. (1999). Increased cytotoxicity and bystander effect of 5-fluorouracil and 5′-deoxyuridine-5-fluorouridine

in human colorectal cancer cells transfected with thymidine phosphorylase. *British Journal of Cancer*, **80**, 1726–1733.

Freeman, S., *et al*. (1993). The bystander effect: tumour regression when a fraction of the tumour mass is genetically modified. *Cancer Research*, **53**, 5274–5283.

Freytag, S., *et al*. (1998). A novel three-pronged approach to kill cancer cells selectively: concomitant viral, double suicide gene and radiotherapy. *Human Gene Therapy*, **9**, 1323–1333.

Golumbek, P., *et al*. (1992). Herpes-simplex 1 virus thymidine kinase is unable to completely eliminate live, nonimmunogenic tumor cell vaccines. *Journal of Immunotherapy*, **12**, 224–230.

Hamstra, D., *et al*. (1999). Enzyme/prodrug therapy for head and neck cancer using a catalytically superior cytosine deaminase. *Human Gene Therapy*, **10**, 1993–2003.

Herman, J., *et al*. (1999). *In situ* gene therapy for adenocarcinoma of the prostate: a phase I clinical trial. *Human Gene Therapy*, **10**, 1239–1249.

Huber, B., *et al*. (1991). Retroviral-mediated gene therapy for the treatment of hepatocellular carcinoma: an innovative approach for cancer therapy. *Proceedings of the Natural Academy of Sciences of the USA*, **88**, 8039–8043.

Huber, B., *et al*. (1993) *In vivo* antitumor activity of 5-fluorocytosine on human colorectal carcinoma cells genetically modified to express cytosine deaminase. *Cancer Research*, **53**, 4619–4626.

Huber, B., *et al*. (1994). Metabolism of 5-fluorocytosine to 5-fluorouracil in human colorectal tumor-cells transduced with the cytosine deaminase gene – significant antitumor effects when only a small percentage of tumor-cells express cytosine deaminase. *Proceedings of the Natural Academy of Sciences of the USA*, **91**, 8302–8306.

Hughes, B., *et al*. (1995). Bystander killing of melanoma cells using the human tyrosinase promoter to express the *Escherichia coli* purine nucleoside phosphorylase gene. *Cancer Research*, **55**, 3339–3345.

Izquierdo, M., *et al*. (1996). Human malignant brain tumor response to herpes simplex thymidine kinase (HSVtk)/ ganciclovir gene therapy.*Gene Therapy*, **3**, 491–495.

Knox, R., *et al*. (1988). The nitroreductase enzyme in Walker cells that activates 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB 1954) to 5-(aziridin-1-yl)-4-hydroxylamino-2-nitrobenzamide is a form of NAD(P)H dehydrogenase (quinone) (EC 1.6.99.2). *Biochemical Pharmacology*, **37**, 4671–4677.

Knox, R., *et al*. (1992). The bioactivation of 5-(aziridin-1-yl)-2,4-dinitrobenzamide (CB 1954). II: a comparison of an *E. coli* nitroreductase and Walker DT diaphorase. *Biochemical Pharmacology*, **44**, 2297–2301.

Kojima, A., *et al*. (1998). *In vivo* human carboxylesterase cDNA gene transfer to activate the prodrug CPT-11 for local treatment of solid tumors. *Journal of Clinical Investigation*, **101**, 1789–1796.

Lan, K., *et al*. (1997). *In vivo* selective gene expression and therapy mediated by adenoviral vectors for human carcinoembryonic antigen-producing gastric carcinoma. *Cancer Research*, **57**, 4729–4284.

Loubiere, L., *et al*. (1996). The equine herpes virus 4 thymidine kinase leads to a superior ganciclovir cell killing than the human herpes virus 1 thymidine kinase. *Gene Therapy*, **6**, 1638–1642.

Marais, R., *et al*. (1996). Gene-directed enzyme prodrug therapy with a mustard prodrug/carboxypeptidase G2 combination. *Cancer Research*, **56**, 4735–4742.

McNeish, I., *et al*. (1998). Virus directed enzyme prodrug therapy for ovarian and pancreatic cancer using retrovirally delivered *E. coli* nitroreductase and CB1954. *Gene Therapy*, **5**, 1061–1069.

Melcher, A., *et al*. (1998). Tumor immunogenicity is determined by the mechanism of cell death via induction of heat shock protein expression. *Nature Medicine*, **4**, 581–587.

Moolten, F. (1986). Tumor chemosensitivity conferred by inserted herpes thymidine kinase genes: paradigm for a prospective cancer control strategy. *Cancer Research*, **46**, 5276–5281.

Mullen, C., *et al*. (1992). Transfer of the bacterial gene for cytosine deaminase to mammalian cells confers lethal sensitivity to 5-fluorocytosine: a negative selection system. *Proceedings of the Natural Academy of Sciences of the USA*, **89**, 33–37.

Pandha, H., *et al*. (1999). Genetic prodrug activation therapy for breast cancer: a phase I clinical trial of *erb*B-2-directed suicide gene expression. *Journal of Clinical Oncology*, **17**, 2180–2189.

Rainov, N., *et al*. (1998). New prodrug activation gene therapy for cancer using cytochrome P450 4B1 and 2-aminoanthracene/ 4-ipomeanol. *Human Gene Therapy*, **9**, 1261–1273.

Ram, Z., *et al*. (1997). Therapy of malignant brain tumours by intratumoral implantation of retroviral vector-producer cells. *Nature Medicine*, **4**, 1354–1361.

Roberts, J., *et al*. (1986). CB1954 (2,4-dinitro-5-azirinylbenzamide) becomes a DNA interstrand crosslinking agent in Walker tumour cells. *Biochemical and Biophysical Research Communications*, **140**, 1073–1078.

Samejima, Y. and Merulo, D. (1995). 'Bystander killing' induces apoptosis and is inhibited by forskolin. *Gene Therapy*, **2**, 50–58.

Shand, N., *et al*. (1999). A phase 1–2 clinical trial of gene therapy for recurrent glioblastoma multiforme by tumor transduction with the herpes simplex thymidine kinase gene followed by ganciclovir. *Human Gene Therapy*, **10**, 2325–2335.

Sorscher, E., *et al*. (1994). Tumor cell bystander killing in colonic carcinoma utilizing the *Escherichia coli* DeoD gene to generate toxic purines. *Gene Therapy*, **1**, 233–238.

Sterman, D., *et al*. (1998). Adenovirus-mediated herpes simplex virus thymidine kinase/ganciclovir gene therapy in patients with localised malignancy: results of a phase I clinical trial in malignant mesothelioma. *Human Gene Therapy*, **9**, 1083–1092.

Szary, J., *et al*. (1997). Characteristics of cytosine deaminase–5-fluorocytosine system: enhancement of radiation cytotoxicity and bystander effect. *Cancer Gene Therapy*, **4**, 307–308.

Thatcher, N., *et al*. (1999). The potential of acetaminophen as a prodrug in gene directed enzyme prodrug therapy. *Cancer Gene Therapy*, **7**, 521–525.

Topf, N., *et al*. (1998). Regional ‘pro-drug’ gene therapy: intravenous administration of an adenoviral vector expressing the *E. coli* cytosine deaminase gene and systemic administration of 5-fluorocytosine suppresses growth of hepatic metastases of colon carcinoma. *Gene Therapy*, **5**, 507–513.

Trinh, Q., *et al*. (1995). Enzyme/prodrug gene therapy: comparison of cytosine deaminase/5-fluorocytosine *versus* thymidine kinase/ganciclovir enzyme prodrug systems in a human colorectal carcinoma line. *Cancer Research*, **55**, 4808–4812.

Tubiana, M. and Malaise, E. (1976). Growth rate and cell kinetics in human tumours: some prognostic and therapeutic implications. In: Symington, T. and Carter, R. (eds), *Scientific Foundations of Oncology*. 126–136 (Heinemann, London).

van der Eb, M., *et al*. (1998). Severe hepatic dysfunction after adenovirus-mediated transfer of the herpes simplex virus thymidine kinase gene and ganciclovir administration. *Gene Therapy*, **5**, 451–458.

Vile, R., *et al*. (1994). Systemic gene therapy of murine melanoma using tissue specific expression of the *HSVtk* gene involves an immune component. *Cancer Research*, **54**, 6228–6234.

Vile, R., *et al*. (1997). Generation of an anti-tumour immune response in a non-immunogenic tumour: HSVtk killing *in vivo* stimulates a mononuclear cell infiltrate and a Th1-like profile of intratumoural cytokine expression. *International Journal of Cancer*, **71**, 267–274.

Weedon, S., *et al*. (2000). Sensitisation of human carcinoma cells to the prodrug CB1954 by adenovirus vector-mediated expression of *E. coli* nitroreductase. *International Journal of Cancer*, **86**, 848–854.

Wygoda, M., *et al*. (1997). Protection of herpes simplex virus thymidine kinase transduced cells from ganciclovir mediated cytotoxicity by bystander cells: the ‘Good Samaritan’ effect. *Cancer Research*, **57**, 1699–1703.

FURTHER READING

Editorial (2000). Gene therapy – cautious optimism. *Nature Medicine*, **6**, 717. Editorial reviewing the progress of gene therapy.

Human Gene Marker/Therapy Clinical Protocols (2000). *Human Gene Therapy*, **11**, 2543–2619. Complete listing of all gene therapy protocols approved by the NIH Recombinant DNA Advisory Committee (RAC) up to the end of 2000.

Moolten, F. (1986). Tumor chemosensitivity conferred by inserted herpes thymidine kinase genes: paradigm for a prospective cancer control strategy. *Cancer Research*, **46**, 5276–5281. The original description of thymidine kinase/ganciclovir gene therapy.

Niculescu-Duvaz, I. and Springer, C. (1997). Antibody-directed enzyme prodrug therapy (ADEPT): a review. *Advanced Drug Development Reviews*, **26**, 151–172. Review on antibody-directed enzyme–prodrug therapy, the forerunner of GPAT.

Roth, J. and Cristiano, R. (1997). Gene therapy for cancer: what have we done and where are we going? *Journal of the National Cancer Institute*, **89**, 21–39. Review of all cancer gene therapy clinical trials approved up to the end of 1997.

Websites

American Society of Gene Therapy, www.asgt.org.

European Society of Gene Therapy, www.biosci.ki.se/esgt/.

Chapter 100

Stem Cell Transplantation

Arnon Nagler
Chaim Sheba Medical Center, Tel Hashomer, Israel

CONTENTS

- Introduction
- Autologous Stem Cell Transplantation (ASCT)
- Allogeneic SCT
- Indications for Allogeneic SCT
- Chronic GVHD
- Unrelated SCT
- Cord Blood Transplantation
- Haploidentical SCT
- Adoptive Immunotherapy and Donor Lymphocyte Infusion
- Cytokine-mediated Immunotherapy
- Low-intensity and Nonmyeloablative SCT
- SCT for Autoimmune Diseases

INTRODUCTION

Stem cell transplantation (SCT) is recognized as a conventional, occasionally even front-line, mode of therapy for an increasing number of haematological malignancies including leukaemias and lymphomas and certain solid tumours, not curable by other therapeutic modalities (Slavin and Nagler, 1991). In addition, allogeneic SCT is the treatment of choice for (1) clinical syndromes associated with the life-threatening deficiency of marrow stem cells (e.g. severe aplastic anaemia) or stem cell products (e.g. T lymphocytes in severe combined immunodeficiency, osteoclasts in osteopetrosis, granulocytes in Kostmann's syndrome) and (2) certain genetic disorders leading to production of abnormal marrow products (e.g. β-thalassaemia major) or enzyme-deficiency disorders (e.g. metachromatic leukodystrophy, Gaucher's disease) (Slavin and Nagler, 1991). The history of SCT started in the late 1950s when attempts to treat human patients with total body irradiation or chemotherapy and marrow infusion were reported (Thomas *et al*., 1957). Mathe *et al*. achieved the first persistent allogeneic marrow graft in a patient with leukaemia (Mathe *et al*., 1965). In 1977, Thomas *et al*. reported 100 patients with advanced acute leukaemia who were transplanted from a human leucocyte antigen (HLA) – matched siblings, eight of them still alive and well (Thomas *et al*., 1977). In the past few years, major progress has been made in several aspects of SCT including alternative sources of haematopoietic stem cells (PBSC) and cord blood-derived stem cells (Gluckman *et al*., 1997; Schmitz *et al*., 1998), unrelated and mismatched family related including haploidentical PBSC transplantation (Aversa *et al*., 1998; Hansen *et al*., 1998), donor lymphocyte infusion (DLI) and adoptive immunotherapy and cytokines mediated immunotherapy for prevention and treatment of relapse post-BMT (Slavin *et al*., 1996; Nagler *et al*., 1997). Finally, transplantation using less toxic preparative regimens to induce mixed chimaerism (Slavin *et al*., 1998) makes transplants today applicable both to elderly and to heavily treated high-risk patients traditionally considered ineligible for allogeneic transplants using conventional preparative regimens. In addition, these types of modern transplants make possible an application to other disease categories including autoimmune diseases (Van Bekkum, 1998).

AUTOLOGOUS STEM CELL TRANSPLANTATION (ASCT)

In the last decade, autologous stem cell transplantation has increasingly become the alternative treatment modality for patients with haematological malignancies such as leukaemias and lymphomas, as well as with solid tumours (Varadi and Nagler, 1994) (**Table 1**). This therapy involves the use of high-dosage or dose-intensive chemoradiotherapy in conjunction with autologous marrow or peripheral stem cell rescue. According to the International Bone Marrow Transplantation Registry, the number of autologous transplantations is increasing at a higher rate than that of allogeneic transplantations. For treatment of leukaemias, autologous transplantation has several advantages

Table 1 Indications for autologous stem cell transplantation (ASCT)

Disease	Comment
Acute myeloid leukaemia	Best results from SCT performed as consolidation therapy during first complete remission, or with stored marrow during first relapse
Acute lymphoid leukaemia	High relapse rate; early detection of relapse important
Chronic myeloid leukaemia	Best results obtained after induction of Philadelphia chromosome negativity in marrow or stem cells
Non-Hodgkin lymphoma	Consider early ASCT in patients with high-risk lymphomas
Hodgkin disease	ASCT best performed during second relapse; consider early SCT for patients with high-risk disease
Multiple myeloma	Substantial responses also in older patients and those with responsive but recurrent disorder
Breast cancer	Greatest benefit in patients with lymph node involvement responsive to standard chemotherapy

over allogeneic transplantation, as follows: SCT is available for more patients, since it can be used for patients who do not have HLA identical donors and for adults up to age 60 years who achieve initial remission and are in good clinical condition. Intensive immunosuppression is not required for ASCT and therefore transplant-related toxicity and mortality are significantly lower. The preparative regimen provides the maximum dose-intensive therapy with the goal of eradicating the malignancy. Patients receiving ASCT can tolerate more intensive conditioning regimens, since there is no graft versus host reaction. In addition, the patient's immune reconstitution is much faster and therefore there is a decreased risk of infectious complications such as cytomegalovirus. There are also theoretical disadvantages for ASCT: the graft versus leukaemia effect observed with allogeneic transplantation is lacking. There is an additional concern of malignant cell contamination of the harvested autologous bone marrow and peripheral blood stem cells (Brenner, 1995).

Regarding the pre-ASCT conditioning regiments, there appear to be no significant differences between the various conditioning regimens for ASCT in terms of disease-free survival and cure rates, and it is likely that the current chemoradiotherapeutic conditioning regimens are at the limit of multiorgan toxicity. Overall, the relationship between the dosage of chemoradiotherapy and tumoricidal effect is logarithmic, and substantial dosage increments would be needed for total eradication of malignancy (Varadi and Nagler, 1994). Recombinant haematopoietic growth factors and autologous peripheral stem cell reduced substantially the morbidity and mortality associated with chemoradiotherapy and autologous transplantation, (Gianni *et al*., 1989). However, alternative approaches such as cell-mediated and cytokine-mediated immunotherapy will be used to control residual clonogenic tumour cells that have escaped the conditioning procedures.

ALLOGENEIC SCT

Bone marrow has been the major source of haematopoeitic stem cells (HSCs) used for allogeneic transplantation until it was discovered that growth factors were able to mobilize HSC into the peripheral blood (Gianni *et al*., 1989). The optimum number of nucleated cells needed for long-term engraftment is estimated at 2×10^8 kg^{-1}. The identifications of markers of progenitor stem cells such as the CD34 antigen has allowed the development of methods for the isolation and purification of immature progenitors. The critical number of transplanted T cells for induction of acute graft versus host disease (GVHD) is within the range 1×10^5–1×10^6 kg^{-1}. In order to prevent GVHD, several methods of T cell depletion have been used, including negative selection with monoclonal antibodies, positive selection of CD34+ cells with immunomagnetic beads or biotin–avidin columns (Aversa *et al*., 1998; Slavin and Nagler, 1991). This manipulation gives an enrichment of CD34+ cells with a 2–3 logarithmic T cell depletion. These methods of T cell depletion are mostly used in family HLA mismatched or in unrelated transplantation. It has been shown that T cell depletion increases the risk of rejection and leukaemic relapse. This complication can be overcome by increasing the dose of conditioning and the number of cells infused (Aversa *et al*., 1998). Other protocols investigate selective T cell depletion or addition of donor peripheral blood lymphocytes after transplant (Nagler *et al*., 1997). The incidence of chronic GVHD seems to be higher following allogeneic G-CSF mobilized PBSCT (Schmitz *et al*., 1998).

INDICATIONS FOR ALLOGENEIC SCT

Indications for allogeneic SCT include malignant and nonmalignant disorders. Most of the allogeneic SCT (about 80%) are performed for haematological malignancies which include in decreasing order: chronic myeloid leukaemia (CML) mainly in chronic phase but also in more advanced disease, acute myeloid leukaemia (AML) (two-thirds are performed in first complete remission and about third in more advanced disease), acute lymphatic leukaemia (ALL) (about half are performed in first complete remission and the other half in more advanced

disease), myelodysplastic syndrome (MDS) and recently also in chronic lymphatic leukaemia (CLL) in increasing numbers. In other haematological malignancies such as non-Hodgkin lymphoma (NHL), multiple myeloma (MM) and Hodgkin lymphoma (HD) autologous SCT is the first choice. However, in patients who either failed autoSCT or suffer from primary refractory or resistant disease, alloSCT is an legitimate alternative. About 10% of the alloSCT are performed in these disease categories. AlloSCT in solid tumours including renal cancer, breast cancer, ovarian cancer, melanoma, central nervous system malignancies, neuroblastoma, etc., have been reported, but they should be performed only in the context of experimental clinical protocols.

As for nonmalignant disorders, severe aplastic anaemia and Fanconi anaemia are the leading indication for alloSCT, followed by haemoglobinopathies (mainly thalassaemia major), inborn errors and immunodeficiency syndromes. About 10% of the alloSCT are performed in nonmalignant disorders. Recently, several alloSCT have been performed in autoimmune diseases.

Chronic Myeloid Leukaemia (CML)

Patients with CML under the age of 60 years may be treated by allogeneic SCT from sibling donors on a routine basis. The transplant should ideally be performed in chronic phase within 1 year from diagnosis, but patients in advanced phases may also be offered transplants on an individual basis. The timing of transplant for patients with compatible donors who achieve a major or complete cytogenetic response to interferon-α remains uncertain. The number of transplants using unrelated donors is increasing. The effect that the recent compound ST1571 will have on the algorithm of transplantation in CML has yet to be seen.

Acute Myeloid Leukaemia

Patients with AML in first remission may be treated by allogeneic SCT on an individual basis or within the context of a clinical study. Patients who fail to achieve complete remission after two courses of induction chemotherapy may be treated by allogeneic SCT with an HLA-identical sibling. Patients with AML in early relapse or in second or later remission may also be treated by allogeneic SCT. Patients in established relapse are not generally recommended for allogeneic transplant. Transplants involving unrelated donors for AML in remission should proceed largely in the context of a clinical research protocol.

Acute Lymphoblastic Leukaemia (ALL)

Selected patients with ALL, especially those with poor prognostic features (e.g. adults, patients of any age with Philadelphia chromosome-positive ALL), are currently considered for treatment by allogeneic SCT in first remission if they have a sibling donor. The same stipulations apply to a patient with standard-risk ALL treated with chemotherapy only in first remission who relapses and is then restored to second remission with further chemotherapy. As with AML, ALL patients who fail to achieve complete remission after two courses of induction chemotherapy may be treated by allogeneic SCT with an HLA-identical sibling or, if time permits, by allografting from a matched unrelated donor.

Mylelodysplastic Syndromes (MDS)

Allogeneic SCT is considered the treatment of choice for patients with MDS or secondary AML (sAML) offering a good chance of long-term disease-free survival if the transplant is performed in an early stage of the disease (i.e. refractory anaemia or refractory anaemia with excess of blasts) or if the patient is transplanted in complete remission after chemotherapy. The transplant is limited to patients aged less than 55 years with an HLA-identical sibling. Allogeneic SCT may also be considered for patients aged less than 45 years who have a fully matched unrelated donor.

Chronic Lymphocytic Leukaemia (CLL)

Selected patients under the age of 55 years with HLA-identical sibling donors may be treated by allogeneic transplantaion in the context of a clinical research protocol. Such patients will usually be adults with poor prognostic features at diagnosis who have responded to conventional therapy pretransplant.

Multiple Myeloma (MM)

Allogeneic SCT in MM is preferably carried out in patients up to the age of 55 years who have responded to first-line treatment, or before second-line treatment if the patient has not responded to first-line treatment. Allogeneic transplantation with sibling donors may also be considered for selected patients who are nonresponsive or who have already received several lines of treatment. Transplant with unrelated donors should only be considered in the context of a clinical research protocol.

Malignant Lymphoma (NHL and HD)

Allogeneic SCT for lymphoblastic lymphoma might be considered for young adults in first remission. AlloSCT may be considered for patients with NHL or HD with an HLA-identical sibling donor who relapse postautografting or patients with primary refractory or resistant relapse in the setting of an experimental protocol.

Nonmalignant Disorders

Severe Aplastic Anaemia (SAA)

Allogeneic SCT from an HLA-identical sibling is the treatment of choice for patients with acquired SAA under the age of 45 years. The conditioning regimen should probably not include irradiation because of the high risk of secondary tumours. In older patients, or in the absence of a matched sibling, an initial course of immunosuppressive treatment is recommended. Unrelated donor transplants are still associated with significant morbidity and mortality and should be undertaken within clinical research protocols.

Immunodeficiency Syndromes, Inborn Errors and Haemoglobinopathies

For severe immunodeficiency syndromes such as severe combined immunodeficiency (SCID), allogeneic SCT from a matched donor or a haploidentical SCT from a parent is recommended. For all severe homozygous thalassaemias and also for sickle cell disease, allogeneic SCT from an HLA-identical sibling is the only chance of cure.

CHRONIC GVHD

Chronic GVHD (cGVHD) is a major complication occurring in 25–45% of patients 100 days after allogeneic SCT (Sullivan *et al*., 1981). It resembles connective tissue–autoimmune like immunological disorder characterized by cutaneous and mucosal manifestations (Sullivan *et al*., 1981). The hallmark of the disease is lichenoid or sclerodermoid lesions of the skin (Sullivan *et al*., 1981).

Autoreactive T cells specific for the common determinant of MHC class II molecules and production of unusual patterns of cytokines including IL-4 are most probably involved (de Gast *et al*., 1987). However, the precise mechanism still has to be elucidated, and a broad array of effector cells and cytokines are believed to participate. Current treatment options for cGVHD are limited and consist of immunosuppressive agents including methylprednisolone, cyclosporin and azathioprine, thalidomide and total lymph node irradiation, with a limited success rate (Neudorf *et al*., 1984). Halofuginone, an inhibitor of collagen type $\alpha 1$ (I) synthesis and gene expression is a new potential therapy for cGVHD (Nagler and Pines, 1999). Halofuginone is a plant alkaloid known to inhibit specifically collagen $\alpha 1$ (I) gene expression and collagen synthesis (Nagler and Pines, 1999). We have previously demonstrated in two murine models of cGVHD, (1) B10.D2 $\rightarrow$ Balbc and (2) tight skin (TSK) mice, that halofuginone abrogated in a dose-dependent manner the increase in collagen $\alpha 1$ (I) gene expression and the increase in skin collagen and prevented thickening of the dermis and loss of subdermal fat, all of which are characteristics of cGVHD (Levi-Schaffer *et al*., 1996; Halevy *et al*., 1996). Recently, we treated a cGVHD patient by topical application of halofuginone. Halofuginone-containing ointment was applied daily on the left side of the neck and shoulder of the cGVHD patient. Collagen $\alpha 1$ (I) gene expression and collagen content in skin biopsy specimens were evaluated by *in situ* hybridization and sirius red staining, respectively. After 3 and 6 months, a marked reduction in skin collagen synthesis was observed, accompanied with increased neck rotation on the treated side. After cessation of treatment, the sclerosis, skin tightness and collagen $\alpha 1$ (I) gene expression returned to the baseline level. No adverse effects were observed, and no plasma levels of halofuginone could be detected. Halofuginone may thus provide a promising novel and safe therapy for cGVHD patients (Nagler and Pines, 1999).

UNRELATED SCT

Only 30–40% of the patients in the Western hemisphere who are in need of SCT have an HLA-identical sibling who can serve as an allogeneic marrow donor (Beatty *et al*., 1988). Patients who do not have a compatible donor must make do with a donor who is a partially matched sibling, a family member, or unrelated (Kernan *et al*., 1993). The major obstacles for successful allogeneic BMT are GVHD and graft rejection (Slavin and Nagler, 1991). These complications were found to be directly correlated with the degree of HLA disparity between the patient and the stem cell donor (Anasetti *et al*., 1989). Furthermore, in order to increase the success of SCT across a major HLA barrier, the intensity of conditioning must be increased to achieve maximum immunosuppression and a large number of cells must be administered (Aversa *et al*., 1998). This superintensive conditioning is liable to result in higher transplant-related morbidity and mortality. Until recently, donor selection was based mainly on identity of serologically defined HLA class I and class II antigens. At present, molecular analysis of HLA class II alleles has also become mandatory for unrelated donor selection (Petersdorf *et al*., 1991). HLA A–B and DR antigens are the three loci considered to be crucial for donor selection. However, recent publications and circumstantial evidence indicate the possible importance of additional HLA loci, particularly HLA-C of class I, which have so far not been included in the criteria for selecting unrelated bone marrow donors (Nagler *et al*., 1996). We recently analysed the impact of molecular HLA-C disparity on postallogeneic unrelated SCT outcome and complication and were able to demonstrate that the GVHD and graft rejection were significantly higher and the actuarial survival and disease-free survival significantly lower in molecular HLA-C mismatched than matched patients (Nagler *et al*., 1996). We conclude, therefore, that a mismatch in locus C may be

detrimental to SCT outcome and should therefore be included as a risk factor in routine pre-SCT HLA phenotyping.

CORD BLOOD TRANSPLANTATION

Cord blood (CB) is an alternative source of haematopoietic progenitors of allogeneic transplantation of patients lacking an HLA-matched marrow donor (Varadi *et al*., 1995). The demonstration of the presence of HSC in CB suggested the use of these cells for transplantation and the first successful CB transplantation (CBT) was performed by Gluckman *et al*. and reported in 1989 (Gluckman *et al*., 1989). Since then more than 2000 patients all over the world have received related or unrelated CB transplants for a variety of haematological and genetic diseases (Gluckman *et al*., 1997; Rubinstein *et al*., 1998). Most CB recipients have been children with an average weight of up to 44 kg. The progression-free survival rates reported thus far are comparable to results achieved following allogeneic bone transplantation, with a suggestion that CB may produce less GVHD (Rocha *et al*., 2000).

The benefits of using umbilical cord blood cells for transplantation are immediate availability of cells, absence of donor risks and low risk of transmitting infectious diseases (Varadi *et al*., 1995). Other potential advantages of CBT include the ability to increase markedly the number of allografts available, and thus the number of patients who could be transplanted, given the availability and ease of collecting CB from placental veins prior to disposal of the placenta compared with collecting bone marrow from an unrelated living donor. This new source of haematopoietic progenitors will allow the transplant community to target collection of CB from normal pregnancies in minority populations to treat the high incidence of genetic diseases. Additionally, the efficiency of gene transfer into CB progenitors appears to be higher than that reported into marrow or peripheral blood progenitors cells (PBPCs) (Zhov *et al*., 1994). Banks of cryopreserved and HLA typed CB have now been established worldwide (Rubinstein *et al*., 1995). The most important advantage of CBT is less GVHD, which is the main obstacle for successful allogeneic SCT (Rocha *et al*., 2000). Rocha *et al*. (2000) studied 113 HLA identical sibling umbilical CB transplants reported to the International Bone Marrow Registry and Eurocord Cord Blood Transplant Group between 1990 and 1997, and compared them with 2052 HLA identical sibling bone marrow transplants performed during the same time period. Children ($\leq$15 years old) were included in the study. Multivariate analysis demonstrated lower risks of grade II–IV acute GVHD ($p = 0.001$) and chronic GVHD ($p = 0.02$) in umbilical CBT recipients. Since acute GVHD results from activation, clonal expansion and proliferation of donor-derived T lymphocytes that recognize alloantigen presented by either host or donor antigen-presenting cells, the lower GVHD risk after CBT transplantation might be due to impairment of these functions in CB cells. It has been shown that CB lymphoid cells are naive, immature and have different cytokine requirements than bone marrow cells. Data from *in vitro* and *in vivo* studies demonstrated that umbilical cord blood lymphocytes (1) are either functionally or phenotypically naive compared with adult blood lymphocytes, (2) have a unique cytokine profile (producing fewer cytokines and expressing mRNA transcripts for interferon-γ, IL-4 and IL-10, but very little IL-2), (3) have a fully constituted polyclonal T cell repertoire, (4) could be protected from apoptosis due to low levels of CD95 and (5) have functions that are inducible through *in vitro* or *in vivo* activation. Consequently, early NK and T cell cytotoxicity is impaired, but secondary activation can occur. All these immunological properties might result in a reduced capacity of transplanted CB T cells to induce GVHD (Rocha *et al*., 2000).

The main disadvantage of CBT is delayed neutrophil and platelet recovery (Gluckman *et al*., 1997; Rubinstein *et al*., 1998; Rocha *et al*., 2000). Expansion of CB progenitors *ex vivo* prior to infusion could potentially ameliorate the slower haematopoietic recovery by generating a higher number of haematopoietic progenitors. One way to obtain *ex vivo* expanded CB derived haematopoietic progenitor cells is by using copper chelators (Peled *et al*., 1999). We have previously demonstrated that metallic ions, including copper, have a regulatory role in proliferation and differentiation of HPC. Depletion of copper by polyamine chelators such as tetraethylenepentamine (TEPA) transiently blocks differentiation, thus allowing increased and prolonged cytokine supported expansion of HPC with minimal cell differentiation. In a recent study, CD34+ cell-enriched populations derived from human CB were cultured in Flt-3, TPO, SCF and IL-6. TEPA-containing cultures produced clonogenic and CD34+ cells continuously for at least 11 weeks (**Table 2**). During this period, total cells increased by up to 1×10^6-fold and CFU and CD34+ cells by up to 1×10^4-fold each. A significant subset of the population remained

Table 2 *Ex vivo* expansion of CB-derived HPC

Weeks	+TEPA			−TEPA		
	CFU[a]	CD34(%)	CD34[b]	CFU[a]	CD34%	CD34[b]
4	352	7.6	60	288	3.9	29
5	ND	7.2	179	ND	3.2	72
7	7936	6.0	422	947	2.0	43
10	11264	6.0	7620	0	<0.5	<5
11	ND	5.0	5000	0	<0.5	<5

Cultures initiated with 1×10^4 CB-derived CD34+ cells mL^{-1} were supplemented with SCF, Flt-3, TPO and IL-6 (50 mg mL^{-1} each) with or without TEPA (15 μM).
[a] $\times 10^3$.
[b] Total CD34 cells $\times 10^4$.

morphologically undifferentiated. Removal of the chelator resulted in cell differentiation. In the absence of TEPA, clonogenic and CD34+ cells expanded for 4–8 weeks; thereafter, their numbers declined rapidly as they differentiated. These results demonstrate that reducing copper availability delays differentiation and maintains self-renewal of a subset of CB-derived HPC and thereby promotes their long-term expansion. These effects are transient and reversible, indicating that chelator-treated cells maintain the potential for normal development. This method could be utilized for *ex vivo* manipulation of pluripotent stem cells and committed progenitor cells for clinical applications. A prephase I clinical trial using TEPA for *ex vivo* expansion of CB-HPC for transplantation is under way (Peled *et al*., 1999).

HAPLOIDENTICAL SCT

The application of SCT for the treatment of patients with haematological malignancies and other diseases is hampered by the lack of availability of suitable major histocompatibility complex (MHC)-matched donors (Beatty *et al*., 1988; Anasetti *et al*., 1989; Petersdorf *et al*., 1991; Kernan *et al*., 1993; Nagler *et al*., 1996; Aversa *et al*., 1998). Less than 30% of patients who might benefit from SCT have HLA-identical siblings, and only 3–5% have a relative with only a single HLA locus mismatch. The recent establishment of large registries of HLA typed individuals has led to a substantial increase in transplants from unrelated donors (Hansen *et al*., 1998). Although 40–50% of Caucasian patients in the USA are successful in locating an HLA-A, B DR matched, unrelated donor, other ethnic groups have a much lower probability of finding donors owing to marked polymorphism; in general, many patients fail to find an appropriate (related or unrelated) donor. By contrast, nearly all patients have an HLA-haploidentical relative (parent, child, sibling) who could serve as a donor.

During the 1980s, transplantation of bone marrow from family donors who were not fully histocompatible with the recipients was unsuccessful because of graft failure and severe GVHD, at times affecting as many as 90% of recipients. Thorough depletion of T cells from the donor's bone marrow succeeded in preventing GVHD in children with severe combined immunodeficiency disease, but the results of this procedure were disappointing in patients with leukaemia, because the benefit of preventing GVHD was offset by graft failure.

Recently, the Reisner group has demonstrated that a megadose of T cell-depleted bone marrow can overcome MHC barriers in sublethally irradiated mice (Bachar-Lustin *et al*., 1995). They suggested that this facilitating activity is mediated by cells within the CD34+ population, endowed with potent veto activity (Gur *et al*., 1999). Applying this concept to patients Aversa *et al*. were able to achieve high rate of engraftment in recipients of T cell-depleted mismatched transplants who were also given high numbers of haematopoietic stem cells from bone marrow and peripheral blood (Aversa *et al*., 1998). Subsequently the same group transplanted 43 patients with high-risk acute leukaemia with PBSC following CD34 positive selection with no GVHD prophylaxis (Aversa *et al*., 1998). In all the patients, full donor type engraftment was achieved. In none of the patients who could be evaluated did acute or chronic GVHD develop. Regimen-related toxicity was minimal. Transplantation-related mortality was 40%. After a median follow-up of 18 months (range 8–30 months), 12 of the 43 patients were alive and free of disease.

Similarly, Henslee-Downey *et al*. reported 88% engraftment at 32 days, 16% grade II–IV acute GVHD and 8% severe cGVHD in 72 patients transplanted from partially mismatched related donors using the T10 B9 monoclonal antibody for T cell depletion (Henslee-Downey *et al*., 1997). Recently, additional novel approaches for haploidentical transplants were developed. Guinan *et al*. transplanted 12 patients from mismatched donors following induction of alloantigen specific anergy using the CTLA-4 immunoglobulin an agent that inhibits B7:C28-mediated co-stimulation (Guinan *et al*., 1999). Sykes *et al*. performed HLA mismatched transplants by induction of mixed lymphohaematopoietic chimaerism and tolerance (Sykes *et al*., 1999).

ADOPTIVE IMMUNOTHERAPY AND DONOR LYMPHOCYTE INFUSION

Allogeneic SCT is the most effective modality to date to eradicate haematological malignancies in patients at high risk of relapse or resistant to conventional doses of chemo-radiotherapy. High-dose, myeloablative chemo-radiotherapy may also be supported by autologous SCT; however, the rate of anticipated relapse is much higher owing to the lack of graft versus leukaemia (GVL) effects mediated by alloreactive donor-derived T cells (Horowitz *et al*., 1990). A number of approaches including SCT are available to improve antitumour effects. However, these approaches have already been proven to be hazardous since recipients of fully matched marrow allografts, not perfectly matched allografts or matched unrelated allografts, may develop GVHD, shown to be consistently lethal. Allogeneic cell-mediated immunotherapy (alloCT) by donor lymphocyte infusion (DLI) in graded increments of donor-derived peripheral blood lymphocytes is much safer and allows control of GVHD and induction of the GVL effect (Naparstek *et al*., 1995; Slavin *et al*., 1996). Furthermore, we have recently documented that host-type tumour cells resistant to DLI may still respond to unstimulated or *in vitro* activated lymphocytes supported *in vivo* by a short

course of well-tolerated doses of IL-2 (Slavin *et al*., 1996). Our data suggest that alloCT, especially when antileukaemia effector cells are activated by IL-2, may develop into a very effective modality for both treatment and prevention of relapse in patients with resistant disease (Slavin *et al*., 1996) or at high risk of relapse (Slavin *et al.*, 1996). The marked therapeutic benefits of alloCT induced by DLI always carry the risk of GVHD, with an incidence and severity that are unpredictable (Kolb *et al*., 1995; Slavin *et al*., 1996). New approaches to limit the lifespan of donor-derived T cells in the case of uncontrolled GVHD are currently under development. The most promising modality for controlling GVHD, and its incidence after discontinuation of anti-GVHD prophylaxis, is the use of donor T cells transduced with the herpes simplex virus thymidine kinase gene (Bonini *et al*., 1997). Genetically modified T cells of donor origin still retain their GVL capacity. Hence, in the event of uncontrolled GVHD, these antitumour effector cells can be successfully eliminated by administration of conventional doses of ganciclovir (Bonini *et al*., 1997).

Recently there have been some preliminary data that may support the existence of a graft versus tumour (GVT) effect similar to the GVL effect. We have documented possible antitumour responses in six breast cancer patients with documented metastatic breast cancer relapsing following autologous SCT transplantation and 13 patients with advanced malignant lymphoma treated with DLI obtained from an HLA matched sibling. Donor lymphocytes were activated with IL-2 *in vitro* and *in vivo*. One of the patients treated with no evidence of disease at the time of alloCT is still event and disease free over 5 years past therapy (Or *et al*., 1998a,b). Similarly, recently, Porter *et al*. described 18 patients with diverse malignant disorders including Hodgkin disease (HD), melanoma, renal cell carcinoma and non-Hodgkin lymphoma (NHL) who received DLI for relapsing disease postautologous SCT (Porter *et al*., 1999). Four patients developed acute GVHD. Three of the four patients with acute GVHD responded, one with durable complete remission (Porter *et al*., 1999).

CYTOKINE-MEDIATED IMMUNOTHERAPY

Although both HD and NHL are responsive to conventional doses of chemotherapy, in a substantial proportion of the patients, particularly with NHL, relapse cannot be avoided. Patients with primary resistant disease or with relapse following front-line remission induction protocols due to residual tumour cells acquiring resistance to subsequent doses of chemotherapy are unlikely to be cured unless they respond to high, myeloablative doses of chemoradiotherapy supported by autologous SCT (Varadi and Nagler, 1994). However, the relapse rate following autologous SCT (autoSCT) using maximum tolerated doses of chemoradiotherapy is still very high owing to minimum residual disease (MRD) which cannot be eliminated with any of the available modalities (Varadi and Nagler, 1994). A variety of recombinant cytokines may potentially be used to prevent or treat relapse following autoSCT. Some of these cytokines may cause direct antitumour effects, while others may facilitate immunological recognition or activate antitumour effector mechanisms following autoSCT. Immune suppression, including depressed absolute number of CD4+ T cells, decreased T cell response to mitogens, antigens or allogeneic stimulation and profound impairment of IL-2 production, has been observed for up to 1 year following autoSCT, hence it seems important to restore the immunocompetence of the recipient that may be impaired by high-dose chemotherapy required to reduce the tumour load.

In view of the above and in order to reduce relapse rates following autoSCT by induction of lymphokine mediated antitumour effects, we conducted a phase IIb clinical trial on 56 malignant lymphoma (ML) patients with MRD post-autoSCT utilizing a combination of IL-2 and IFN-α subcutaneously (s.c.) in an outpatient setting, and compared the results with 61 matched historical controls (Nagler *et al*., 1997).

The overall survival of ML patients who received immunotherapy was significantly higher than that of ML patients who did not. Survival at 48 months was 90% for the immunotherapy patients and 46% for the historical controls ($p < 0.01$). Similarly, the overall survival was significantly higher for the HD and NHL patients who received immunotherapy when compared with the historical controls. The survival rates at 48 months were 100 and 80% versus 57 and 42%, respectively ($p < 0.02$).

The overall disease-free survival (DFS) of ML patients who received immunotherapy was significantly higher than that of comparable ML patients in the historical control group who did not receive immunotherapy. The actuarial DFS at 48 months was 70 and 48%, respectively ($p < 0.01$). Similarly, the actuarial DFS was significantly higher for the NHL and HD patients after immunotherapy than for the historical controls. The actuarial DFS for NHL patients receiving immunotherapy at 48 months was 64 and 41% for patients who did not receive immunotherapy ($p < 0.01$). The actuarial DFS for patients with HD receiving immunotherapy at 48 months was 88 and 60% for patients who did not receive immunotherapy ($p < 0.042$).

The relapse rate was significantly lower for ML patients who received immunotherapy than for a similar cohort of patients belonging to the historical controls. Of the 56 patients who received immunotherapy, 11 (20%) relapsed (eight NHL and three HD patients), whereas of the 61 patients who did not receive immunotherapy 29 (46%) relapsed (21 NHL and eight HD patients) ($p < 0.01$).

We subsequently wanted to see whether modification of the immunotherapy schedule by reducing rIL-2 to 1 week

($3–6 \times 10^6$ IU/m^2 day^{-1}), followed by combined rIL-2 6×10^6 IU/m^2/day interferon-α 3×10^6 IU/day for 1 month and extending interferon-α to 6 months (3×10^6 IU day^{-1} $\times$ 3 per week) would improve efficacy and/or tolerability. Thirty-eight ML patients were enrolled in the (plus IL-2 followed by IFN-α maintenance immunotherapy) protocol. The results were similar to those obtained with the original protocol (Slavin and Nagler, 1998). We are currently conducting a multicentre prospective randomized trial, investigating our newest rIL-2/interferon-α combination for intermediate-and high-grade lymphoma and HD in an attempt to confirm the benefit of cytokine mediated immunotherapy in the setting of minimal residual disease.

LOW-INTENSITY AND NONMYELOABLATIVE SCT

Considering the role of alloreactive donor lymphocytes in mediating GVL and GVT effects, including in patients who are resistant to conventional anticancer modalities, we have introduced the concept of using the bone marrow transplantation procedure as a platform for induction of host versus graft transplantation tolerance rather than as a means of eradicating all tumour cells. Safe and stable transplantation tolerance can best be accomplished by induction of mixed chimaerism, which can be achieved using nonmyeloablative conditioning. Once host versus graft tolerance allows consistent and durable engraftment of donor immunohaematopoietic cells, donor lymphocytes can be added if needed for induction of GVL and GVT effects to displace residual malignant or genetically abnormal host cells. Patient age in the first cohort that entered our protocol ranged between 1 and 64 years (median, 38 years) (Slavin *et al.*, 1998). Conditioning included immunosuppressive treatment with six daily infusions of fludarabine (Fludara, Schering), 30 mg m^{-2} (days -10 to -5), oral busulfan, 4 mg kg^{-1} day^{-1} for two consecutive days (days -6 to -5), and anti-T lymphocyte globulin (ATG), 10 mg kg^{-1} day^{-1} for four consecutive days (days -4 to -1). Organ toxicity was minimal (Slavin *et al.*, 1998). The protocol was much better tolerated in comparison with the anticipated side effects following a standard myeloablative regimen. No cases of grade 3 or 4 toxicity (World Health Organisation criteria) were observed, and grade 2 mucositis was documented in only a few cases. In one-third of the patients, ANC did not decrease below 0.1×10^9 L^{-1}; some patients never experienced an ANC lower than 0.5×10^9 L^{-1}, and in some the platelet counts did not decrease below 20×10^9 L^{-1}, thus requiring no platelet support at all.

Engraftment was documented in all patients by increasing blood counts and using either amelogenin polymerase chain reaction (PCR) for detection of residual male cells in a recipient of female cells or using VNTR-PCR in sex-matched donor recipient pairs. Severe GVHD (grade 3 or 4) was the single major complication diagnosed and was the only cause of mortality in 13% of the patients, most of whom developed the first signs of disease while off cyclosporin A (CSA). The incidence of relapse was not higher than anticipated following conventional alloSCT. Minimal residual disease was consistently reversed by discontinuation of CSA or using DLI as soon as patients were off CSA. A response was observed in approximately half of the patients who experienced relapse with no GVHD (Slavin *et al.*, 1998). Overall this protocol was given to 70 patients with haematological malignancies (CML ($n = 19$), AML ($n = 17$), ALL ($n = 10$), NHL ($n = 15$), MDS and second leukaemia ($n = 6$), HD ($n = 2$) and MM ($n = 1$)), 16 patients with Matched Unrelated (MUD) and nine patients receiving an allograft following failure of autoBMT. Patients age ranged between 18 and 52 years (median, 34 years). They received fully matched ($n = 66$) or single locus mismatched ($n = 4$) stem cells on day 0. The preliminary results were holding. The protocol was very well tolerated by patients of all age groups. Day 100 mortality was 4%, 0% in patients with nonmalignant diseases and 7% in patients with malignancy. Fast and durable engraftment was observed in all patients with a matched sibling. Persistent evidence of disease or recurrent disease in mixed chimaeras was treated by discontinuation of CSA or by graded increments of DLI with 10 of 15 patients responding. After 3 years, with an observation period of 3–39 months (median, 24 months) the actuarial probability of survival was 68% and disease-free survival was 48%. Corresponding numbers for MUD at 12 months were 75%, and for recipients of second BMT at 18 months 48%, respectively. However, GVHD remains the single major problem (Slavin *et al.*, 1998).

Anderson's group performed a pilot trial of purine analogue-containing nonmyeloablative therapy for 25 patients with AML or MDS considered ineligible for myeloablative therapy and allogeneic transplantation because of either age or medical conditions (Giralt *et al.*, 1997). They used the Flag-ida (fludarabine + idarubicine + Ara-c) or the 2CDA + Ara-c protocol. Twenty patients had neutrophil recovery a median of 11 days after transplant (range 9–21 days), and 17 achieved platelet transfusion independence a median of 15 days after transplant (range 8–78 days). One patient died from infectious complications secondary to graft failure. The 1-year survival for all patients was 23%; disease-free survival was 15% (Giralt *et al.*, 1997). McSweeney *et al.* reported a preparative regimen consisted of TBI 200 cGy as a single fraction, followed by post-transplant immunosuppression with CSA from day -1 to $+35$ and mycophenolate mofetil (MMF) (McSweeney *et al.*, 1998). Eight patients with haematological malignancies (CLL, 2; AML-CR, 3; myeloma, 2; RAEB, 1) were treated. Myelosuppression was minimal, with only one patient developing neutrophil counts of $< 0.5 \times 10^9$ L^{-1}. Some degree of donor T cell engraftment was observed in all

patients: 15–100% donor by day 28 and 5–100% donor by day 56 (McSweeney *et al.*, 1998). Induction of mixed chimaerism by the nonmyeloablative approach with combination of less cytotoxic regimens, intense immunosuppression and high number of HSC is particularly attractive for nonmalignant diseases such as genetic diseases (Slavin *et al.*, 1998) or autoimmune diseases.

SCT FOR AUTOIMMUNE DISEASES

Autoimmune diseases result from self-reactive T-lymphocytes and autoantibodies, produced most likely in cooperation with T cell-dependent B cells. Until recently, nonspecific suppression of self-reactive lymphocytes or the inflammatory process mediated by the ongoing anti-self-reactivity represented the main goal of therapy, but in most cases neither cure nor remission could be obtained (Van Bekkum, 1998). Prior data for experimental animals and humans indicated that high-dose chemotherapy, especially myeloablative chemoradiotherapy supported by autologous SCT (Karussis *et al.*, 1992) and allogeneic SCT, can result in effective control of autoimmune diseases (Van Bekkum *et al.*, 1989). Fassas *et al.* (1997) performed autoSCT in patients with multiple sclerosis (MS) in a progressive phase. The patients were conditioned with the BEAM regimen and ATG. Improvement in the neurological scale was observed in 35% of the patients. One patient died of invasive aspergillosis (Fassas *et al.*, 1997). SCT may help overcome lack of response to self-antigens since the conditioning prior to the transplantation procedure normally involves myeloablative treatment that results in elimination of host-type immunohaematopoietic cells, T cells included, followed by stem cell rescue which results in regeneration of new T cells tolerant to self-antigens. Hence, if the autograft is T cell depleted, newly regenerating T cells are likely to become tolerant to self antigens since self-reactive T cells, certainly high-affinity self-reactive T cells, are likely to undergo apoptosis *in status nascendi* in the thymus. Similarly, following the use of T cell-depleted stem cell allograft, which can also be used for rescue of the myeloablated recipient, it seems very reasonable that both anti-self and anti-host reactivity will be abolished by the aforementioned procedure. The technical possibility of performing alloSCT and inducing a state of mixed chimaerism and tolerance with very low toxicity by using nonmyeloablative conditioning is thus very attractive for autoimmune diseases.

REFERENCES

Anasetti, C., *et al.* (1989). Effect of HLA compatibility on engraftment of bone marrow transplants in patients with leukemia or lymphoma. *New England Journal of Medicine*, **320**, 197–204.

Aversa, F., *et al.* (1998). Treatment of high risk acute leukemia with T cell depleted stem cells from related donors with one fully mismatched HLA haplotype. *New England Journal of Medicine*, **339**, 1186–1193.

Bachar-Lustin, E., *et al.* (1995). Megadose of T cell depleted bone marrow overcomes MHC barriers in sublethally irradiated mice. *Nature Medicine*, **1**, 1268–1273.

Beatty, P. G., *et al.* (1988). Probability of finding HLA matched unrelated marrow donors. *Transplantation*, **45**, 714–718.

Bonini, C., *et al.* (1997). HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft versus leukemia. *Science*, **276**, 1719–1724.

Brenner, M. K. (1995). The contribution of marker gene studies to hemopoietic stem cell therapies. *Stem Cells*, **13**, 453–461.

de Gast, G. C., *et al.* (1987). The multifactorial etiology of graft versus host disease. *Immunology Today*, **8**, 209–212.

Fassas, A., *et al.* (1997). Peripheral blood stem cell transplantation in the treatment of progressive multiple sclerosis: first results of a pilot study. *Bone Marrow Transplantation*, **20**, 631–638.

Gianni, A. M., *et al.* (1989). Granulocyte macrophage colony stimulating factor to harvest circulating haemopoietic stem cells for autotransplantation. *Lancet*, **2**, 580–585.

Giralt, S., *et al.* (1997). Engraftment of allogeneic hematopoietic progenitor cells with purine analog containing chemotherapy: Harnessing graft versus leukemia without myeloablative therapy. *Blood*, **89**, 4531–4536.

Gluckman, E., *et al.* (1997). Outcome of cord blood transplantation from related and unrelated donors. *New England Journal of Medicine*, **337**, 373–381.

Gluckman, E., *et al.* (1989). Hematopoietic reconstitution in a patient with Fanconi's anemia by means of umbilical cord blood from an HLA-identical sibling. *New England Journal of Medicine*, **321**, 1174–1178.

Guinan, E. C., *et al.* (1999). Transplantation of anergic histoincompatible bone marrow allografts. *New England Journal of Medicine*, **340**, 1704–1714.

Gur, H., *et al.* (1999). Specific *in-vitro* inactivation of host anti-donor CTL-p by human CD34 progenitor cells: evidence against anergy based mechanisms. *Blood*, **94**, 3919.

Halevy, O. (1996). Inhibition of collagen type I synthesis by skin fibroblasts of graft versus host disease and scleroderma patients: effect of halofuginone. *Biochemical Pharmacology*, **52**, 1057.

Hansen, J. A. (1998). Bone marrow transplants from unrelated donors for patients with chronic myeloid leukemia. *New England Journal of Medicine*, **338**, 962–968.

Henslee-Downey, P. J., *et al.* (1997). Use of partially mismatched related donors extends access to allogeneic marrow transplant. *Blood*, **89**, 3864–3872.

Horowitz, M. M., *et al.* (1990). Graft versus leukemia reactions after bone marrow transplantation. *Blood*, **75**, 555–562.

Karussis, D. M., *et al.* (1992). Prevention of experimental autoimmune encephalomyelitis and induction of tolerance with

acute immunosuppression followed by syngeneic bone marrow transplantation. *Journal of Immunology*, **148**, 1693–1698.

Kernan, N. A., *et al*. (1993). Analysis of 462 transplantations from unrelated donors facilitated by the National Marrow Donor Program. *New England Journal of Medicine*, **328**, 593–602.

Kolb, H. J., *et al*. (1995). Graft versus leukemia effect of donor lymphocyte transfusions in marrow grafted patients. *Blood*, **86**, 2041–2050.

Levi-Schaffer, F., *et al*. (1996). Inhibition of collagen synthesis and changes in skin morphology in murine graft versus host disease and tight skin mice: effect of halofuginone. *Journal of Investigative Dermatology*, **84**, 106.

Mathe, G., *et al*. (1965). Adoptive immunotherapy of acute leukemia: Experimental and clinical results. *Cancer Research*, **25**, 1525–1531.

McSweeney, P., (1998). Outpatient PBSC allografts using immunosuppression with low dose TBI before and cyclosporine and mycophenolate mofetil after transplant. *Blood*, **92**, 519a.

Nagler, A., (1997). Immunotherapy with recombinant human interleukin-2 and recombinant interferon-α in lymphoma patients post autologous marrow or stem cell transplantation. *Blood*, **89**, 3951–3959.

Nagler, A., *et al*. (1996). Bone marrow transplantation using unrelated and family related donors: the impact of HLA-C disparity. *Bone Marrow Transplantation*, **18**, 891–897.

Nagler, A. and Pines, M. (1999). Topical treatment of cutaneous chronic graft versus host disease with halofuginone: a novel inhibitor of collagen type I systhesis. *Transplantation*, **68**, 1–4.

Naparstek, E., *et al*. (1995). T-cell depleted allogeneic bone marrow transplantation for acute leukemia using Campath-1 antibodies and post transplant administration of donor's peripheral blood lymphocytes for prevention of relapse. *British Journal of Haematology*, **89**, 506–515.

Neudorf, S., *et al*. (1984). Prevention and treatment of acute graft versus host disease. *Seminars in Hematology*, **21**, 91–100.

Or, R., *et al*. (1998a). Allogeneic cell mediated immunotherapy for breast cancer after autologous stem cell transplantation. A clinical pilot study. *Cytokines Cellular and Molecular Therapy*, **4**, 1–6.

Or, R., *et al*. (1998b). Allogeneic cell mediated immunotherapy at the minimal residual disease stage following autologous stem cells transplantation for malignant lymphoma. *Immunotherapy*, **21**, 447–453.

Peled, T., *et al*. (1999). Long term expansion of hematopoietic progenitor (HPC) from human umbilical cord blood by copper chelators. *Blood*, **94**, 7099.

Petersdorf, E. W., *et al*. (1991). Polymorphism of HLA-DRw52 associated DRB1 genes as defined by sequence specific oligonucleotide probe hybridization and sequencing. *Tissue Antigens*, **38**, 169–177.

Porter, D., *et al*. (1999). Graft versus tumor induction with donor leukocyte infusions as primary therapy for patients with malignancies. *Journal of Clinical Oncology*, **7**, 1234–1243.

Rocha, V., *et al*. (2000). Comparison of graft versus host disease in children transplanted with HLA identical sibling umbilical cord blood versus bone marrow hematopoietic stem cells. *New England Journal of Medicine*, **342**, 1846–1854.

Rubinstein, P., *et al*. (1995). Processing and cryopreservation of placental/umbilical cord blood for unrelated bone marrow reconstitution. *Proceedings of the National Academy of Sciences of the USA*, **92**, 10119–10122.

Rubinstein, P., *et al*. (1998). Outcomes among 562 recipients of placental blood transplants from unrelated donors. *New England Journal of Medicine*, **339**, 1565–1577.

Schmitz, N., *et al*. (1998). Allogeneic bone marrow transplantation vs filgrastim mobilized peripheral blood progenitor cell transplantation in patients with early leukemia: first results of a randomized multicenter trial of the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplantation*, **21**, 995–1003.

Slavin, S. and Nagler, A. (1991). Developments in bone marrow transplantation. *Current Opinion in Oncology*, **3**, 254–271.

Slavin, S., *et al*. (1996). Allogeneic cell therapy with donor peripheral blood cells and recombinant human interleukin-2 to treat leukemia relapse after allogeneic bone marrow transplantation. *Blood*, **86**, 2195–2204.

Slavin, S. and Nagler, A. (1998). Immunotherapy in conjunction with autologous and allogeneic blood or marrow transplantation in lymphoma. *Annals of Oncology*, **9**, S31–S39.

Slavin, S., *et al*. (1998). Nonmyeloablative stem cell transplantation and cell therapy as an alternative to conventional bone marrow transplantation with lethal cytoreduction for the treatment of malignant and nonmalignant hematologic diseases. *Blood*, **91**, 756–763.

Sullivan, K. M., *et al*. (1981). Chronic graft versus host disease in 52 patients: adverse natural course and successful treatment with combination immunosuppression. *Blood*, **57**, 267–276.

Sykes, M., *et al*. (1999). Mixed lymphohaematopoietic chimerism and graft versus lymphoma effects after non-myeloablative therapy and HLA-mismatched bone marrow transplantation. *Lancet*, **353**, 1755–1759.

Thomas, E. D., *et al*. (1957). Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. *New England Journal of Medicine*, **257**, 491–496.

Thomas, E. D., *et al*. (1977). Cure of leukemia by marrow transplantation. *Leukemia Research*, **1**, 67–70.

Van Bekkum, D. W., *et al*. (1989). Regression of adjuvant induced arthritis in rats following bone marrow transplantation. *Proceedings of the National Academy of Sciences of the USA*, **86**, 10090–10099.

Van Bekkum, D. W. (1998). Autologous stem cell therapy for treatment of autoimmune diseases. *Experimental Hematology*, **26**, 831–834.

Varadi, G. and Nagler, A. (1994). Conditioning regimens in autologous bone marrow transplantation. *Clinical Immunotherapy*, **2**, 342–351.

Varadi, G., *et al*. (1995). Umbilical cord blood for use in transplantation. *Obstetrical and Gynecological Survey*, **50**, 611–617.

Zhou, S. Z., *et al*. (1994). Adeno associated virus 2-mediated high efficiency gene transfer into immature and mature subsets of hematopoietic progenitor cells in human umbilical cord blood. *Journal of Experimental Medicine*, **179**, 186–194.

FURTHER READING

Anasetti (1989). Unrelated stem cell transplantation. *New England Journal of Medicine*, **320**, 197–207.

Aversa (1998). Haploidentical Transplants. *New England Journal of Medicine*, **339**, 1186–1193.

Fassas (1997) Transplantation for autoimmune diseases. *Bone Marrow Transplantation*, **20**, 631–638.

Gianni (1989) Mobilized peripheral blood stem cells. *Lancet*, **2**, 580–588.

Gluckman (1989). Cord blood transplants. *New England Journal of Medicine*, **321**, 1176–1178.

O'Reilly (1983). Allogeneic stem cell transplantation. *Blood*, **62**, 941–949.

Slavin (1996). Adoptive immunotherapy and donor lymphocyte infusion. *Blood*, **86**, 2195–2204.

Slavin (1998). Low-intensity and non-myeloablative stem cell transplantation. *Blood*, **91**, 756–763.

Varadi (1994). Autologous stem cell transplantation. *Clinical Immunotherapy*, **2**, 342–351.

Waldmann (1984). Graft versus host disease. *Lancet*, **ii**, 483–489.

Chapter 101

Novel Surgical Strategies in the Management of Cancer

Ara Darzi and Paul Ziprin
Imperial College School of Medicine, London, UK

CONTENTS

- Introduction
- Principles of Surgical Oncology
- Laparoscopy
- Colonic Stenting
- Thoracic Surgery
- Sentinel Lymph Node Mapping
- Image-Guided Surgery
- The Future
- Conclusion

INTRODUCTION

Surgery is still the only potentially curative treatment that can be offered for many cancers, in particular gastrointestinal malignancies. This is despite extensive research into alternative treatments, including immunotherapy and gene therapy, at a high pecuniary cost.

Probably the greatest advance that has been made in surgery over the last decade has been the development of minimal access technology, which has led to a new philosophy in surgery: that of minimizing trauma to the patient. This has been made possible due the advances that have been made in fibre optics and computer-enhanced imaging. This has, therefore, given rise to many developments in all fields of surgery including the introduction of endoscopic surgery and image-guided surgery such as stereotactic and radioimmunoguided surgery.

PRINCIPLES OF SURGICAL ONCOLOGY

Surgery has evolved from the previous dogma of radical resection of the primary tumour and surrounding structures as surgeons appreciate the systemic nature of cancer and develop strategies that minimize locoregional recurrence. This is illustrated by the replacement of Halstead's radical mastectomy in the management of breast cancer with a more conservative approach: the use of wide local excision or segmentectomy coupled with radiotherapy or the use of the ABBI system (see the section Stereotactic Breast Surgery) in the management of suitable breast tumours. However, it is important that this more conservative approach must not compromise the oncological success of the surgery.

Although the exact management of different malignancies will vary, there are some basic principles that must be followed during surgical resection. An in-depth knowledge of the anatomy with special regard to the lymph node drainage and the vasculature remains paramount.

Curative or Palliative Therapy

The treatment of solid tumours can be considered to be either potentially curative or palliative and, therefore, the type of surgical therapy used will depend on both the patient and tumour's characteristics.

If the patient's health is such that major surgery is likely to be detrimental, then a more conservative or palliative approach is warranted. Likewise, if the primary cancer is unresectable and/or metastases are present, curative treatment will usually be unsuccessful and palliative therapy should be instigated. Palliation may include surgery, for example, the bypassing of an obstructing, inoperable tumour for symptomatic relief. Medical management such as pain control is discussed in more detail elsewhere in this book.

The mainstay of curative treatment of most solid tumours is surgical resection of the primary growth. Exceptions include radiotherapy or topical 5-fluorouracil in the management of suitable squamous cell carcinomas of the skin or radical radiotherapy for prostate cancer. Surgical therapies are now part of a multidisciplinary

approach to cancer management and can be used in combination with preoperative (neoadjuvant) or postoperative (adjuvant) radiotherapy or chemotherapy.

Clinical Staging

In producing the treatment plan for the patient, it is important to stage the tumour clinically. This allows the identification of those patients with (1) metastatic disease for whom extensive surgery will be unnecessary and (2) tumours that require neoadjuvant therapy, e.g. patients with locally advanced rectal carcinoma benefit from preoperative radiochemotherapy. Knowledge of the common sites to which a particular tumour metastasizes will aid in deciding on the appropriate investigations to be ordered. For example, radioisotope bone scanning is used in screening for metastases in prostate cancer and liver imaging is used for gastrointestinal tumours.

The different methods used for staging are summarized in **Table 1**. Symptoms elicited from the patient's history such as weight loss or bone pain are suggestive of metastatic disease. Clinical examination may reveal hepatomegally compatible with liver metastases or can be used to assess the primary tumour itself, such as assessing the fixity of a rectal cancer during a rectal examination.

Radiological investigations commonly used include ultrasonography, computer tomography, magnetic resonance imaging and less commonly positron emission tomography (PET), all of which can be used for imaging the primary tumour to assess local disease as well as the liver and other organs in the search for metastases. Other radiological techniques include radioisotope scanning, which can be used for staging patients with prostate, breast or thyroid cancer.

Ultrasonography probes have been modified for use intraluminally such as the use of transrectal ultrasound in the local staging of rectal tumours and for staging and to aid in the histological diagnosis of prostate cancer. Fibreoptic endoscopes have also been adapted to incorporate an ultrasound probe and have been used to increase the accuracy of staging of oesophageal, gastric and pancreaticobiliary tumours.

The use of virtual colonoscopy, which can accurately detect lesions as small as 1 cm in size, may with further development assessment become a part of the radiologists' regular armament. Further developments include the use of magnetic resonance (MR) endoscopy in which the coil is incorporated into the tip of the fibre-optic scope.

More invasive techniques are now being developed with the increasing use of staging laparoscopy and thoracoscopy in the management of a variety of tumours. These advances will be discussed in more detail later in the chapter. Although now rare with the recent development of imaging techniques, patients occasionally require a laparotomy, especially if the diagnosis is in doubt.

Table 1 Methods for clinical staging

Method	
Clinical	
History, e.g. cachexia, bone pain	
Clinical examination	
Radiological	For the assessment of the primary tumour and metastatic disease, e.g. liver secondaries from gastrointestinal malignancies
Ultrasonography	
Computed tomography	
Magnetic resonance imaging	
PET scanner	
Radioisotope scan	E.g. for bone metastases from breast, prostate, kidney, thyroid and bladder cancers
Endoscopy	
Endoscopic ultrasound	
Surgical	
Staging laparoscopy/thoracoscopy	
Laparoscopic ultrasonography	
Sentinel lymph node mapping	
Laparotomy	

Operative Principles

Prevention of locoregional recurrence remains the goal in surgical oncology. The surgeon should resect all of the malignant tissue from the primary organ and any other structure involved locally, together with the lymphatics, vascular structures and tissue through which the tumour is likely to spread. For example, in colonic surgery, the vessels should be ligated at their origin, allowing removal of the mesentery, which also contains the draining lymph nodes.

During surgery the tumour should be handled as little as possible to avoid tumour spillage. This problem may account for the reports of early port-site recurrences following minimal access surgery for malignancy during the infancy of advanced laparoscopic surgery (see the section Disadvantages of Laparoscopic Surgery). It is also imperative that the tumour is excised with adequate margins to reduce the risks of local recurrence. It has been shown, for example, that wedge resection of non-small cell carcinoma of the lung results in high recurrence rates compared with lobectomy. Problems arise particularly in oesophageal surgery where the tumour infiltrates submucosally, so making it difficult to identify whether the margins are microscopically clear at the time of the operation. Other strategies such as early ligation of vessels in colonic surgery to prevent tumour dissemination during surgical mobilisation have not been proved to be beneficial.

Although removal of the draining lymphatics remains an important principle in oncological surgery, there has recently been a change in practice. For some cancer operations, lymphadenectomy may be associated with a

high degree of morbidity, such as surgery for malignant melanoma and breast carcinoma. This has led to the development of sentinel lymph node mapping, which identifies those patients who do not require more extensive surgery (see the section Sentinel Lymph Node Mapping).

Summary

Although total excision of the primary tumour in continuation with its lymphatics and locally involved tissue remains the aim in oncological surgery, curative resection may not be possible in some situations. Debulking surgery together with radiotherapy or chemotherapy may achieve some degree of control of the disease and may prolong survival, but it is important to remember that the aim of surgery is also to improve the quality of life.

LAPAROSCOPY

Staging of Malignancy

In 1911, Bernheim reported the use of organoscopy in a patient with pancreatic cancer to 'reveal general metastases or a secondary nodule in the liver, thus rendering further procedures unnecessary' (Bernheim, 1911). Staging laparoscopy is used, therefore, for patients with a known diagnosis of cancer to ascertain the tumour's suitability for resection and so help in the surgical management of these patients.

Although ultrasound scanning, computed tomography (CT) and magnetic resonance imaging (MRI) can detect liver metastases with a sensitivity and specificity of 75–85%, these imaging modalities are poor at detecting subcapsular liver deposits and peritoneal disease.

Staging laparoscopy has gained increasing favour more recently in the management of cancers where the majority of patients present with advanced disease that is not curable, such as gastro-oesophageal, pancreatic and biliary malignancies. It not only allows direct visualization of the intrabdominal organs, but also enables the surgeon to perform biopsies of unsuspected lesions and cytological examination of peritoneal lavage samples. This process allowed the detection of occult metastases in over 30% of patients with pancreatic cancer in a recent study (Jimenez *et al.*, 2000).

The accuracy of staging has been further enhanced by the use of both endoscopic and laparoscopic ultrasound (**Figure 1**). Both have been shown to increase the accuracy of staging of upper gastrointestinal malignancies especially in the management of pancreatico-biliary cancers where laparoscopic and endoscopic ultrasound can detect nodal disease and early invasion of the portal vein and thereby identify patients with unresectable disease (van Dijkum *et al.*, 1999). Further developments include the use of endoscopic magnetic resonance imaging, which has the coil situated in the tip of the endoscope, and this is currently undergoing evaluation in the authors' institution.

Laparoscopic pelvic lymph node dissection is also used in the staging of prostate adenocarcinoma. It is used for patients at high risk of nodal metastases, who would otherwise be suitable for potentially curative treatment such as radical radiotherapy or radical prostatectomy. At-risk patients have been defined as those with a prostate-specific antigen level $>20\ \mathrm{ng\,mL^{-1}}$, a Gleason score of ≥ 7 and stage T2b–T3a disease (Stone *et al.*, 1997).

Laparoscopic Surgery

Laparoscopic and laparoscopic-assisted surgery are both well described now for most abdominal and pelvic malignancies, including gastro-oesophageal, pancreatic, gynaecological and urological cancers, with laparoscopic colorectal cancer surgery being the most commonly performed oncological procedure.

For laparoscopic surgery to be an accepted method of the surgical management of cancers, a number of parameters must be considered, as must be when any new treatment is used. It must be as safe as or safer than conventional surgery as assessed by operative morbidity and mortality, and offer benefits in terms of length of hospital stay, postoperative pain and cost. Also, sound oncological principles must not be compromised.

Oncological Principles and Laparoscopic Surgery

Parameters that have been used to assess the adequacy of oncological procedures include number of lymph nodes harvested and adequacy of resection margins.

With regard to colorectal cancer, it was first shown on cadavers that an adequate lymphadenectomy could be performed laparoscopically. Since then, comparative studies have shown no differences in the number of lymph nodes harvested by open or laparoscopic techniques.

The length of specimen retrieved and proximal and distal margins have also been studied. These parameters are comparable between laparoscopic and open colectomy. In rectal cancer surgery, the radial margin is most important in predicting local recurrence. In a series of 12 patients undergoing laparoscopic abdomino-perineal resection, our unit showed that there was not a significant difference in the radial margins when compared with a similar number of patients undergoing conventional surgery (Monson *et al.*, 1992).

Advantages of Laparoscopic Surgery

Laparoscopic colorectal surgery has been shown, when compared with open surgery, to result in a more rapid return of gastrointestinal function, less postoperative pain, a reduction in length of stay and an earlier return to normal activities. This is probably related to reduced

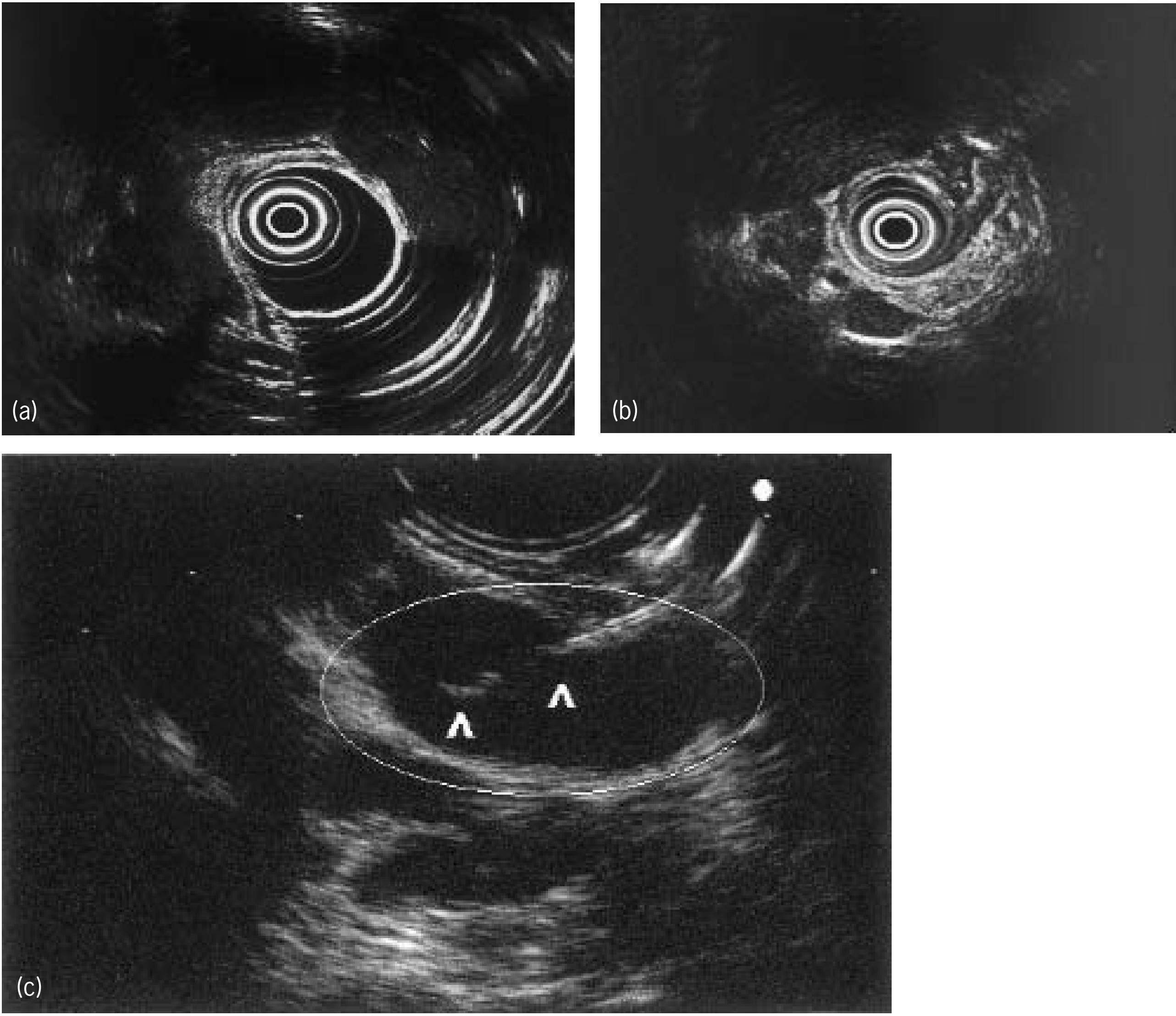


Figure 1 (a) Endoscopic ultrasound showing a carcinoma of the head of pancreas with sonographical pathological perihepatic lymph nodes (b) (courtesy of Mr David Bowrey, Cardiff, UK). (c) Endoscopic ultrasound image showing a transgastric ultrasound-guided fine needle aspiration (arrows) of a 2 × 1 cm coeliac axis lymph node (outlined).

access-related trauma, as seen by a reduced cytokine response following laparoscopic colorectal resection (Schwenk *et al*., 2000).

Although the actual cost of the laparoscopic procedure is higher than open surgery, there may be more long-term cost savings with regard to shorter convalescence and earlier return to work.

Disadvantages of Laparoscopic Surgery

Laparoscopy lacks the manual dexterity afforded by open surgery, which may give rise to some theoretical problems. It has been reported that synchronous tumours have gone unrecognized during laparoscopic surgery and there are other reports of healthy bowel having been resected under the belief it contained small tumours while the diseased bowel was left *in situ*. Poor tactile feedback may also make it more difficult to identify important structures such as the ureter during sigmoid colon mobilization.

One of the major controversies since the introduction of laparoscopic surgery in the treatment of cancer has been the phenomenon of port-site metastases (Wexner and Cohen, 1995). Port-site recurrences have been reported in a number of different cancers such as colorectal, gall-bladder, gynaecological, pancreatic, lung and urological malignancies. In colorectal surgery the incidence varies between 2 and 23% and these metastases appear to develop not only in patients with advanced disease, but also following surgery for early lesions such as Dukes A colorectal carcinoma.

There may be a number of reasons for this occurrence. First, it may be due to surgical inexperience during the initial development of laparoscopic surgery, resulting in the inappropriate handling of the tumour, resulting in cancer dissemination. It has been shown that the instruments can become contaminated with tumour cells. Extraction of the specimen through small trocar sites may also cause exfoliation of cells into the wound. The evidence that surgical skill plays an important role is twofold. First, a recent study showed that following laparoscopic splenectomy in an animal tumour model, the incidence of port-site recurrences decreased with increasing experience (Lee *et al*., 1998). Second, early results from ongoing randomized controlled studies have shown a much lower incidence.

Other causes may be technical problems associated with laparoscopic surgery. These include aerosolization of viable tumour cells by the pneumoperitoneum and the 'chimney effect,' which involves the drawing up of cells to the port sites by microscopic leakages around the laparoscopic instruments.

The pneumoperitoneum itself may play a role. A number of animal studies have shown that a carbon dioxide pneumoperitoneum causes an enhancement in tumour growth and peritoneal dissemination compared with air or gasless laparoscopy (Bouvy *et al*., 1998) and work from the author's unit has shown increased tumour cell invasiveness following exposure to an *in vitro* CO_2 pneumoperitoneum.

Only the results from ongoing prospective randomized, controlled trials will address both the short- and long-term outcomes following endoscopic surgery for malignancy. At present laparoscopic oncological surgery should be reserved for patients entered into trials or for those patients requiring palliative surgery.

Hand-assist Laparoscopic Surgery

As already mentioned, one of the drawbacks of laparoscopic surgery is the lack of tactile feedback afforded by open surgery and poor depth perception with two-dimensional display. In addition, advanced laparoscopic procedures have not been adopted owing to technical difficulties and to limitations of the instrumentation. To overcome these problems, hand-assist laparoscopic surgery (HALS) has been developed.

With this procedure, the surgeon inserts a hand into the peritoneal cavity to aid in the laparoscopic dissection while the pneumoperitoneum is maintained. This approach utilizes, from the start of surgery, an incision that would be necessary for the retrieval of the specimen later during the operation (**Figure 2; see colour plate section**).

The Handport System (Smith and Nephew, MA, USA) consists of base retractor, a bracelet and a sleeve. The bracelet is attached to the surgeon's wrist, which is usually the nondominant hand, before the sleeve is placed over the forearm and fastened to the bracelet. An incision equivalent in centimetres to the surgeon's glove size is made in the abdominal wall and the inner ring of the base retractor, is inserted into the abdominal cavity after the creation of the pneumoperitoneum. This is secured against the abdominal wall by inflating the base retractor, producing a seal that will maintain the pneumoperitoneum after the hand is introduced into the abdomen and the sleeve attached to the outer ring of the retractor.

A review of the use of HALS has been published recently (Litwin *et al*., 2000). Included in the review were a number of cases of colonic cancer requiring right hemicolectomy, sigmoid colectomy and in one case transverse colectomy. HALS compares favourably in terms of length of operation, postoperative pain and length of hospital stay with previously published data on comparable conventional laparoscopic procedures. Problems encountered with HALS include leakage of the pneumoperitoneum around the Handport device, hand fatigue, which may be due to poor placement leading to bad ergonomics, and slipping of the base retractor in obese patients as the abdominal wall may be too thick for the device.

Colorectal disease lends itself well to this procedure as it is a complex procedure, difficult to perform purely laparoscopically and it also requires an incision to remove the specimen. Other procedures that have utilized this technique include splenectomy, gastric surgery, adrenalectomy and cryoablation of liver metastases.

HALS, therefore, may help surgeons in advanced laparoscopic procedures by offering tactile feedback, safe hand retraction and the ability to use blunt dissection.

Endoscopic Surgery

Until recently, radical surgery had been the method of choice in the treatment of cancers. Now there is an increasing move to less radical surgery in the treatment of early cancers as seen in breast surgical oncology. Local resection requires accurate staging of the primary tumour prior to surgery, with, for example, endoanal ultrasound in the management of rectal cancer, as there is a increased propensity of lymph node metastases with more advanced cancers. Local resection would not be appropriate if there was a high risk of lymph node metastases.

Transanal Endoscopic Microsurgery

Transanal endoscopic microsurgery (TEM) was first described by Buess in 1985 in the resection of rectal tumours and enables the surgeon to resect tumours and suture inside the rectal cavity up to 20 cm from the anal margin (Buess *et al*., 1985). It requires the use of a 40-mm diameter rectoscope, videoscopy and gas insufflation to produce an operating space in the rectum (**Figure 3; see**

colour plate section). The rectoscope also has access for the surgeon to use up to four surgical instruments. Full thickness resections are possible, but can only be performed in the extraperitoneal part of the rectum (**Figure 4; see colour plate section**).

TEM has an obvious role in the surgical management of benign rectal adenomas, but it is now known that its use results in similar local recurrence rates and 5-year survival figures compared with anterior resection in the management of low-risk early (T1) rectal cancers. However there is unlikely to be a role for TEM in the curative treatment of high-risk T1 lesions (i.e. those with early lymphatic invasion) and T2 lesions, both of which are associated with an increased risk of lymph node disease.

Endoscopic Resection of Upper Gastrointestinal Malignancies

Although the majority of patients in the Western world presenting with gastric cancer have advanced disease, a small number are diagnosed with early gastric cancer (EGC). Endoscopic resection of EGC is now well described, especially by Japanese surgeons where the proportion of patients with EGC is much higher.

It has been suggested that tumours suitable for endoscopic resection are those cancers that are confined to the mucosa and that measure no more than 1cm in size (Tis or T1 tumours), as tumours of this size have a low incidence of lymph node disease (Nakamura *et al*., 1999). Results of endoscopic resection compare favourably from an oncological perspective with those for radical gastrectomy when used in these selected patients.

COLONIC STENTING

Stents have for a long time been used in the palliation of dysphagia in patients with oesophageal cancer, using celestin tubes and more recently expandable metal stents. The use of these metallic stents in colonic surgery has recently increased, particularly in the management of malignant large bowel obstruction.

Classically, patients were treated with a three-stage procedure: formation of a defunctioning loop colostomy to relieve the obstruction, then excision of the tumour and anastomosis and finally reversal of the loop colostomy. This method was used as full colonic preparation is not possible and anastomosing unprepared colon greatly increases the risk of a leak. Alternatives to this are to resect the tumour with the formation of an end colostomy (as in a Hartman's procedure) or to resect the diseased bowel, perform on-table colonic lavage and primary anastomosis (one-stage procedure). However, following the former over 60% of patients do not have their colostomy reversed and the latter is not suitable for high-risk cases, as are many of these patients. Mortality rates following emergency surgery for malignant large bowel obstruction can be as high as 20%.

Expandable metal stents such as the Enteral Wallstent™ (Microvasive Endoscopy, Boston Scientific, Boston, MA, USA) are sufficiently strong and yet flexible to conform to the curvature of the large bowel. The stents are positioned under fluoroscopic control alone or in combination with endoscopy. A guidewire is passed through the stricture and the position is checked radiologically using water-soluble contrast. The two ends of the stricture are marked with radio-opaque skin markers for accurate placement of the stent. Dilatation of the stricture may be performed prior to inserting the stent, but this is rarely necessary. An alternative to dilatation is laser ablation of the tumour to widen the lumen.

The stent is used for palliation if the patient is not curable, but is also a treatment option in the management of acute malignant large bowel obstruction. This will therefore relieve the obstruction allowing the patient to be rehydrated and fed, the bowel to be prepared, colonoscopic examination for synchronous tumours and full oncological staging prior to surgery. This therefore permits the surgeon to perform a one-stage procedure. Although the procedure can be performed in an outpatient setting and stents are successfully placed in 95% of cases, there are as yet no data to support the benefit of stenting in this situation (Lo, 1999).

This procedure, however, does have a number of adverse consequences with a morbidity and mortality rate of 18% and 2% respectively. Documented complications include stent migration, bowel perforation, bleeding and pain that may be abdominal or rectal in origin.

Therefore, metallic stenting can provide quick symptomatic relief in the management of malignant large bowel obstruction prior to surgery and also in the palliative setting, although the associated morbidity needs to be recognized.

THORACIC SURGERY

Thoracoscopy

Jacobaeus first described thoracoscopy in 1910 when he used a cystoscope to examine the thoracic cavity to help in the diagnosis and treatment of tuberculosis (Jacobaeus, 1910). Advances in technology have allowed surgeons to revisit this procedure and develop video-assisted thoracoscopic surgery (VATS).

Diagnostic Thoracoscopy

The most generally accepted use of thoracoscopy in oncology is as an aid to diagnosis, allowing the surgeon to visualize and biopsy pleural, mediastinal and parenchymal lesions.

A randomized, controlled study by Santambrogio showed that VATS excision of solitary pulmonary nodule for diagnostic purposes resulted in less postoperative pain and a shorter hospital stay when compared with conventional surgery (Santambrogio *et al*., 1995a). However, it is not suitable for excising lesions larger than 3 cm in size as these have a high risk of being malignant. The same group also recently described how thoracoscopic ultrasound might aid the identification of deep, invisible pulmonary lesions to help biopsy a suspect lesion (Santambrogio *et al*., 1995b).

Staging Thoracoscopy

Mediastinoscopy is the gold standard for staging lung cancer, but it does have its limitations. For example, access to the subcarinal and subaortic window lymph nodes is poor. VATS, however, offers good visualization of both of these regions from the right and left side, respectively.

As with laparoscopy, VATS also allows the inspection of the pleural surfaces for detection of metastatic disease and to help in the management of suspected mesothelioma.

Thoracoscopic Resection for Malignancy

Thoracoscopic wedge resection and lobectomy for non-small cell lung cancer (NSCLC) have both been described. However, wedge resection for NSCLC gives inferior results in the management of early (T1) tumours in terms of both local recurrence rates and 5-year survival and should be reserved for patients not suitable for more radical surgery. Also, although some studies report advantages of thoracoscopic lobectomy for cancer in terms of reduced postoperative pain and hospital stay compared with thoracotomy, this is not borne out by the only randomized control trial that has been published comparing VATS and muscle-sparing thoracotomy (Kirby *et al*., 1995). As a consequence, open lobectomy should remain the treatment of choice in the management of malignancy outside of clinical trials.

SENTINEL LYMPH NODE MAPPING

Over the last two decades, breast surgery has also seen a move to less radical surgery with the use of wide local excision or quadrantectomy plus radiotherapy compared with the previous widespread use of radical mastectomy as advocated by Halstead.

The recent advance that has been seen in the surgical treatment of breast cancer is in the management of the axillary lymph nodes. The standard treatment for patients with invasive carcinoma is complete or level three axillary dissection, or lymph node sampling followed by radiotherapy to the axilla if histology confirms lymph node involvement. Both of these management strategies are associated with complications, e.g. pain, numbness, limited arm movement, upper limb lymphoedema, brachial plexus neuropathy and radiotherapy-induced osteonecrosis. Therefore, nearly all patients underwent some form of axillary dissection for both staging and therapy, and as a consequence were at risk of developing these complications. This led to the development of sentinel lymph node (SLN) mapping (**Figure 5; see colour plate section**), which has the potential to eliminate axillary dissection for the enlarging cohort of breast cancer patients who are node-negative and so avoid the possibility of the previously discussed complications.

The principle of SLN mapping is that it seeks to identify the first draining lymph node from the primary tumour. Using a radioisotope, patent blue dye or both methods, surgeons can successfully localize SLNs in more than 90% of cases. Generally it has been found that using a combination of radioisotope and blue dye is superior to using either technique alone. After identification of the SLN, the node is examined for evidence of metastatic disease by frozen section, imprint cytology, histology or immunohistochemistry. SLN mapping reliably predicts axillary node status in 98% of all patients and 95% of those who are node-positive (Cox *et al*., 1998).

SLN mapping has also found a role in the management of cutaneous malignant melanoma. Until the emergence of SLN-guided surgery, clinically node-negative patients with poor prognosis melanoma, i.e. melanoma thicker than 1.5 mm, underwent elective lymph node dissection (ELND). However there are no data that show any survival benefit over a selective policy, probably because a large proportion of these patients will have node-negative disease. This, together with the high morbidity of lymphadenectomy with seroma formation, wound breakdown and lymphoedema, means that ELND as routine practice has gone out of favour.

SLN mapping for malignant melanoma is performed, as in breast cancer, using blue dye, radioisotope or both. The marker is injected intradermally around the primary tumour or scar, and this method identifies the sentinel node in more than 90% of cases as seen with breast cancer (Schachter *et al*., 2000).

Therefore, SLN mapping can help identify those patients with microscopic lymph node metastases who may benefit from a block dissection, while sparing those with true node-negative disease an unnecessary major operation. It may also allow better staging of the disease to aid in the planning of adjuvant therapy. SLN could also benefit patients with truncal melanoma. These patients are likely to have multiple draining nodal basins and it has been shown that multiple diseased basins correlate with a poorer prognosis.

However, the use of SLN mapping in the management of breast cancer and malignant melanoma has still to be shown to be better than conventional therapy in terms of long-term outcome. Randomized clinical studies are to start shortly in breast surgery.

IMAGE-GUIDED SURGERY

Stereotactic Breast Surgery

Open surgical biopsies are commonly performed for non-palpable, mammographic abnormalities, but about 70% of these lesions will be found to be benign. These lesions are difficult to manage with regard to localization, amount of tissue to be excised and cosmesis.

Although fine needle aspiration cytology (FNAC) has a role in the management of breast lesions, accuracy is both operator and cytologist dependent. It is also difficult to diagnose carcinoma *in situ* with FNAC. Core biopsies may provide enough tissue sample for histological diagnosis, but it is not always sufficient for biochemical assessment of receptor status.

Other instruments have been developed to excise a greater amount of tissue, allowing a more accurate diagnosis to be made. These include the Mammotome and the Advanced Breast Biopsy Instrumentation (ABBI) system. The mammotome is used for percutaneous breast biopsy and is guided by either stereotaxis or ultrasound. It consists of a hollow outer sleeve through which a vacuum is applied, drawing the breast tissue into the probe prior to the advancement of a rotator cutter producing a larger specimen than conventional techniques.

The ABBI system is used under stereotactic guidance. Different-sized cannulas allow lesions up to 20 mm in diameter to be excised through a small skin incision. The lesion is initially centred using a compression paddle and two stereotactic views are taken. The three-dimensional position of the lesion is calculated from these views and the coordinates are programmed into the computer system. Under local anaesthetic the lesion is localized with a 14-gauge needle coaxial to the cannula and the position is again checked prior to the insertion of a steel wire. After a skin incision long enough to accommodate the cannula has been made, it is advanced into the breast while the blade is rotated. When the blade is approximately 15 mm beyond the radiological abnormality, the cannula is advanced and the blade retracted. The specimen is excised using cautery and withdrawn in the cannula. Radiographs are taken of the breast and specimen to ensure adequate excision.

Initial experience with the ABBI system shows that the advantages that it has over open biopsy includes improved cosmesis as it excises a smaller amount of tissue to produce an accurate diagnosis (Damascelli *et al.*, 1998). It is also less expensive than surgery as it can be performed in the radiology department on an outpatient basis. However, it does take over 1 h to perform this technique and the digital images used are not as good as conventional mammography.

Stereotactic Neurosurgery

Until recently, neurosurgeons used free-hand techniques based on the hand–eye coordination of the surgeon and information derived from a neurological examination and radiographs. Consequently, lesions were often difficult to localize despite the use of generous exposures.

Advances in both computer and radiological imaging technology have allowed the development of stereotactic surgery. The introduction of CT and, later, MRI allowed the surgeon to visualize the tumour and plan the operation in more detail and, with the development of interactive image-guided neurosurgical techniques, surgical access and the amount of tissue damage can be minimized as the lesion can be accurately located during surgery.

Stereotactic neurosurgery requires a head holder for fixing the skull rigidly, a mechanism for directing the instrument to the target, a stereotactic data acquisition unit and surgical instruments which can be directed to the lesion through a frame. Initially anatomical atlases were used for 'imaging' to direct the surgeon, but stereotactic surgery is now performed with the aid of three-dimensional real-time images derived from CT or MR scanning.

Neuronavigation systems have now been developed. These are intraoperative position sensing systems that display in real time the location of the surgical tool on preoperative images (**Figure 6; see colour plate section**). The first system, described in 1987 by Watanabe, consisted of a six-joint articulated positioning arm, an image scanner and a 16-bit microcomputer which tracked the position of the joints of the operating arm. Others have since been developed, including the Neuro-Sat and the Milano Arm.

The first US Food and Drug Administration (FDA)-approved system was the ISG Viewing Wand (ISG Technologies, Canada), the use of which has been well described for both primary and metastatic brain tumours and nononcological procedures. This equipment has in the main replaced stereotactic surgery by removing the need for a cumbersome frame and intraoperative scanning. The use of the Viewing Wand resulted in a shorter operating time and length of hospital stay, reduced the exposure required, aided localization of vital structures in anatomy distorted by the pathology and, as an accuracy of 1–2 mm can be achieved using this system, minimised trauma to normal tissue (Sandeman *et al.*, 1994). An obvious disadvantage of the Viewing Wand is the size and weight of the articulated arm. To overcome this problem, armless devices have now been developed which use sonic, electromagnetic or infrared motion analysis systems to track the instrument.

Radioimmunoguided Surgery

Despite patients undergoing what is perceived to be a 'curative' operation for cancer, many still go on to develop locoregional recurrence and metastatic disease. This is due to the presence of occult disease not detected at surgery or by standard diagnostic and staging techniques.

The Radioimmunoguided Surgery System (RIGS) has been developed to increase the detection and excision of cancer. It is used intraoperatively to detect microscopic disease and to assess the extent of the malignant process. The system uses a radioiodine-labelled antitumour-associated glycoprotein monoclonal antibody. Both overt and occult disease are identified at surgery by a gamma-detecting probe that identifies the antibody in the tissue.

Its use has been most widely reported in the staging and treatment of colorectal cancer but it has also been used in pancreatic cancer. In one series, 25% of patients had their surgical management altered by the RIGS system and 27% of patients with incurable disease were identified solely by RIGS. More recently, it has been demonstrated that it may help improve the outcome of patients undergoing surgery for recurrent colorectal cancer by identifying all residual disease (Martinez *et al*., 1997).

Interventional Magnetic Resonance (IMR) and Surgery

The development of dynamic real-time MRI has led to new possibilities in image-guided surgery. It combines the soft-tissue imaging capabilities of MR with real-time or dynamic imaging. The scanner at the author's institution is a Signa 0.5-T Spio IMR (**Figure 7a**), which is an open configuration system allowing the surgeon access to patients. In order to operate in the IMR unit, magnetic safe instruments, which the scanner also has the ability to track, have had to be developed.

Although the use of intraoperative MRI is in its infancy, its use in a number of different procedures has been reported. IMR has been used in the localization and biopsy of breast lesions, but there are still a number of problems that need to be overcome, including needle artefact, tissue shift during the procedure and fast equalization of contrast enhancement. We have performed a number of different procedures in the IMR including the excision of benign breast lumps (**Figure 7; see colour plate section**), colectomies and laparoscopic procedures such as cholecystectomy.

IMR-guided biopsies, craniotomies and microsurgical tumour resections have been performed by neurosurgeons. The accurate and immediate information that is provided by intraoperative MR imaging allows definitive localization and targeting of the lesion while monitoring the anatomical changes that occur during surgery.

The advantages that may be gained with IMR include visualization of structures deep to the two-dimensional image during endoscopic surgery and clarification of the surgical anatomy and may aid in defining the excision margins required for complete resection of cancers. This is particularly pertinent in neurosurgery, where tissue damage needs to be kept to a minimum.

Interstitial Thermotherapy (ITT)

Approximately 50% of patients with colorectal cancer develop metastatic disease including liver secondaries. Only a minority of patients are suitable for liver resection surgery, which would offer the best outcome. Thermal ablation techniques have been described under ultrasound, CT or conventional MR guidance, but these methods are limited by lack of real-time imaging and difficulties in the interpretation of grey-scale changes during thermal ablation.

The IMR overcomes these limitations with the use of a real-time colourization thermal monitoring system that detects changes in tissue temperature and hence can show the site and size of the evolving thermal lesions (**Figure 8; see colour plate section**). The liver tumours are punctured after the administration of intravenous Mangafodipir trisodium using dynamic MRI, usually under local anaesthetic. The tumour is then treated using a water-cooled interstitial fibre and Nd:YAG laser source while the treatment is monitored with real-time colour sequencing. Although complete tumour ablation may be achieved after only one procedure, repeated thermotherapy can be performed (de Jode *et al*., 1999). IMR has also been used with interstitial thermotherapy in the treatment of other cancers such as head and neck tumours, where clinical symptoms such as swallowing difficulties and nerve compression were improved, and in the management of pelvic soft tissue tumour recurrences.

Therefore, ITT may be used to treat metastatic disease and can obtain good local tumour control.

THE FUTURE

As further developments are made in computer hardware and software, there is likely to be more integration of imaging into surgical practice in the planning phase, to practice the operation and also during the procedure. The Virtual Human Project is a major initiative in medical education and research. It aims to make available both sectional and digitally reconstructed images of the human body – the virtual human. This project will build up a library of tomographic three-dimensional images that may be used in the future for simulation and the operation itself.

Robotic technology in the form of master–slave manipulators is already being used in the operating theatre, particularly in the field of minimal access surgery. Robotic camera assistants, such as EndoAssist™ and the Automated Endoscopic System for Optimal Positioning (AESOP) robot, have been used in a number of procedures such as laparoscopic adrenalectomy. The robot can be controlled with a foot pedal, via a voice-controlled interface, as with the AESOP, or by tracking the head movement of the surgeon. These robots have been shown to

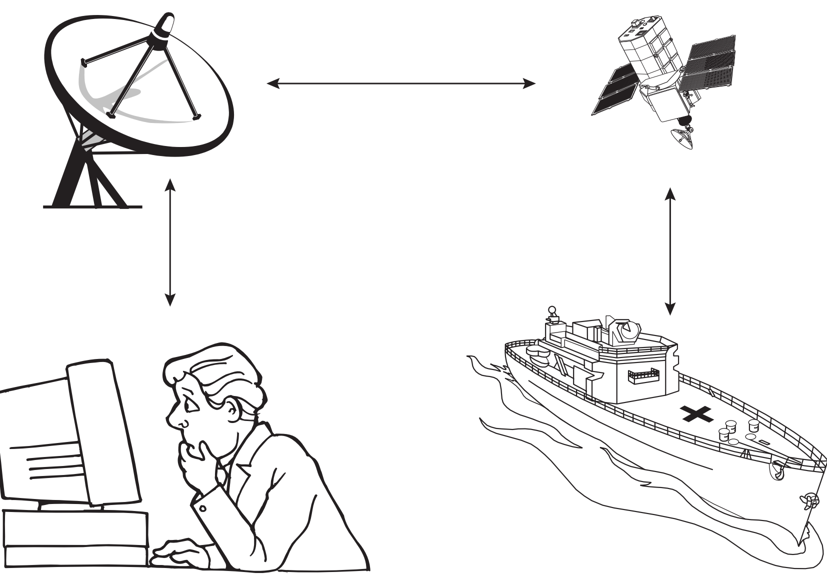


Figure 10 With the aid of surgical robotics, it can easily be envisaged that in the near future surgeons will be able to perform operations at a site remote from the patient.

be superior to the human assistant. The robot reduces the number of camera corrections that are made and the frequency of lens cleaning, both of which can disrupt the surgeon's performance. These tools may aid surgeons as endoscopic surgery for malignancy becomes more popular.

The Da Vinci^{TM} System (Intuitive Surgical, USA) is a computer-enhanced, minimal access surgical system that uses advanced electronics, mechanics and three-dimensional visualization to improve surgical techniques. It has been designed to make minimal access surgery more precise and easier to perform, which may lead to more complex procedures, that at present are not technically possible, being performed endoscopically.

This surgical system consists of two parts: the surgeon's viewing and control console, and the surgical arm unit. This latter unit controls detachable instruments, which access the operative site through 1-cm skin incisions, and these instruments have been designed with 6° of movement to enhance the dexterity of the surgeon's hand and wrist movements (**Figure 9; see colour plate section**).

The surgeon performs the operation from the console, which displays a high-resolution three-dimensional image of the surgical field. The surgeon holds an instrument below the monitor and the surgeon's exact hand movements are simultaneously transferred from the console to precise movements of the instruments at the operative site.

Over 250 operations have been performed with the Da Vinci system, including coronary bypass and mitral valve replacement. Nissen's fundoplication and benign colorectal procedures such as rectopexy have been performed in the author's unit. Although new, this technology may help the conversion of more complex open surgery into minimal access procedures including in the field of surgical oncology.

The development of these robots may eventually lead to remote surgery or telerobotics, enabling surgeons to operate on patients miles away from the medical centre (**Figure 10**).

CONCLUSION

Recent advances in oncological surgery have been in the development of techniques which result in the reduction of access related trauma. This is due to the advances that have been made in optics and computer-enhanced imaging, resulting in the production of images the quality of which aids the surgeon in the preparation and execution of a procedure. However, it is important that any new technique is fully assessed as to its oncological safety before it is offered as routine practice.

REFERENCES

Bernheim, B. (1911). Organoscopy: cystoscopy of the abdominal cavity. *Annals of Surgery*, **53**, 764–767.

Bouvy, N. D., *et al*. (1998). Effects of carbon dioxide pneumoperitoneum, air pneumoperitoneum, and gasless laparoscopy on body weight and tumor growth. *Archives of Surgery*, **133**, 652–656.

Buess, G., *et al*. (1985). Endoscopic surgery in the rectum. *Endoscopy*, **17**, 31–35.

Cox, C. E., *et al*. (1998). Guidelines for sentinel node biopsy and lymphatic mapping of patients with breast cancer. *Annals of Surgery*, **227**, 645–651; discussion, 651–653.

Damascelli, B., *et al*. (1998). Stereotactic excisional breast biopsy performed by interventional radiologists using the advanced breast biopsy instrumentation system. *British Journal of Radiology*, **71**, 1003–1011.

de Jode, M. G., *et al*. (1999). MRI guidance of infra-red laser liver tumour ablations, utilising an open MRI configuration system: technique and early progress. *Journal of Hepatology*, **31**, 347–353.

Jacobaeus, H. (1910). Ueber die Moglichkeit die Zystoskopie bei untersuchung seroser hohlungen anzuwenden. *Münchener Medizinische Wochenshrift*, **57**, 2090–2092.

Jimenez, R. E., *et al*. (2000). Impact of laparoscopic staging in the treatment of pancreatic cancer. *Archives of Surgery*, **135**, 409–414; discussion, 414–415.

Kirby, T. J., *et al*. (1995). Lobectomy – video-assisted thoracic surgery versus muscle-sparing thoracotomy. A randomized trial. *Journal of Thoracic and Cardiovascular Surgery*, **109**, 997–1001; discussion, 1001–1002.

Lee, S. W., *et al*. (1998). Abdominal wound tumor recurrence after open and laparoscopic-assisted splenectomy in a murine model. *Diseases of the Colon and Rectum*, **41**, 824–831.

Litwin, D. E., *et al*. (2000). Hand-assisted laparoscopic surgery (HALS) with the HandPort system: initial experience with 68 patients. *Annals of Surgery*, **231**, 715–723.

Lo, S. K. (1999). Metallic stenting for colorectal obstruction. *Gastrointestinal and Endoscopic Clinics of North America*, **9**, 459–477.

Martinez, D. A., *et al*. (1997). Radioimmunoguided surgery for gastrointestinal malignancies: an analysis of 14 years of clinical experience. *Cancer Control*, **4**, 505–516.

Monson, J. R., *et al*. (1992). Prospective evaluation of laparoscopic-assisted colectomy in an unselected group of patients. *Lancet*, **340**, 831–833.

Nakamura, K., *et al*. (1999). An early gastric carcinoma treatment strategy based on analysis of lymph node metastasis. *Cancer*, **85**, 1500–1505.

Sandeman, D. R., *et al*. (1994). Advances in image-directed neurosurgery: preliminary experience with the ISG Viewing Wand compared with the Leksell G frame. *British Journal of Neurosurgery*, **8**, 529–544.

Santambrogio, L., *et al*. (1995a). Videothoracoscopy versus thoracotomy for the diagnosis of the indeterminate solitary pulmonary nodule. *Annals of Thoracic Surgery*, **59**, 868–870; discussion, 870–871.

Santambrogio, R., *et al*. (1995b). Intraoperative ultrasound during thoracoscopic procedures for solitary pulmonary nodules. *Annals of Thoracic Surgery*, **68**, 218–222.

Schachter, J., *et al*. (2000). Standard and nonstandard applications of sentinel node-guided melanoma surgery. *World Journal of Surgery*, **24**, 491–494; discussion, 494–495.

Schwenk, W., *et al*. (2000). Inflammatory response after laparoscopic and conventional colorectal resections – results of a prospective randomized trial. *Langenbecks Archives of Surgery,* **385**, 2–9.

Stone, N. N., *et al*. (1997). Laparoscopic pelvic lymph node dissection for prostate cancer: comparison of the extended and modified techniques. *Journal of Urology*, **158**, 1891–1894.

van Dijkum, E. J., *et al*. (1999). Staging laparoscopy and laparoscopic ultrasonography in more than 400 patients with upper gastrointestinal carcinoma. *Journal of the American College of Surgeons*, **189**, 459–465.

Wexner, S. D. and Cohen, S. M. (1995). Port site metastases after laparoscopic colorectal surgery for cure of malignancy. *British Journal of Surgery*, **82**, 295–298.

FURTHER READING

Debatin, J. F. and Adam, G. (eds) (1998). *Interventional Magnetic Resonance Imaging* (Springer, Berlin).

Geraghty, J. G., *et al*. (eds) (1998). *Minimal Access Surgery in Oncology* (Greenwich Medical Media, London).

Monson, J. R. T. and Darzi, A. (eds) (1995). *Laparoscopic Colorectal Surgery* (Isis Medical Media, Oxford).

Walker, W. S. (ed.) (1999). *Video-assisted Thoracic Surgery* (Isis Medical Media, Oxford).Chap18

Chapter 102

Translational Research (Overview of Phase I, II and III Clinical Trials)

Nancy L. Lewis and Louis M. Weiner
Fox Chase Cancer Center, Philadelphia, PA, USA

CONTENTS

- Introduction
- Phase I Trials
- Phase II Trials
- Phase III Trials
- Meta-analysis
- The Protocol Document
- Informed Consent
- Institutional Review Boards
- Gemcitabine in Pancreatic Cancer
- The Story of Herceptin and Rituxan

INTRODUCTION

In order to improve the quality of care for cancer patients and to prevent, treat and potentially cure this group of devastating diseases, new therapies must continually be developed. Each year, the Cancer Therapeutic Evaluation Program (CTEP) at the National Cancer Institute (NCI) screens over 20 000 potential new anticancer agents using tumour cell assays. The pharmaceutical industry screens at least as many potential agents, if not more. The most promising agents are then further evaluated in animal studies. Only about 30 new agents are taken to human clinical trials per year. In the United States there are over 60 cancer centres encompassing 1700 institutions with over 8000 clinical investigators. The majority of these centres participate in one or more of 10 government-funded cooperative clinical trials groups (**Table 1**). These groups are consortia of academic institutions and clinical practitioners that take a lead role in defining the value of proposed new therapies for cancer. Together, over 20 000 cancer patients are treated on clinical trials per year using these government-supported mechanisms. The pharmaceutical industry, in cooperation with these groups or institutions, frequently funds additional studies to evaluate new drugs. Despite these efforts, only 3% of all patients with cancer are treated on clinical trials. In many cases, this occurs because of limitations imposed by patients' diseases or performance status. However, additional efforts are needed to enroll patients on clinical research studies to accelerate the pace of cancer treatment improvement. Clinical trials, which are expensive and time consuming, require careful planning and execution in order to answer clinically relevant

Table 1 Cooperative clinical trial groups

Group	Website
American College of Surgical Oncology Group	http://www.ascog.org
Cancer and Leukemia Group B	http://www.calgb.org
Gynecologic Oncology Group	http://www.gog.org
National Cancer Institute of Canada Clinical Trials Group	http://www.ctg.queensu.ca
North Central Cancer Treatment Group	http://www.napbc.ims.nci.nih.gov/database/con_mayo.html
Radiation Therapy Oncology Group	http://rtog.org
Southwest Oncology Group	http://www.swog.org
Children's Oncology Group	http://www.nccf.org
Eastern Cooperative Oncology Group	http://ecog.dfci.harvard.edu
National Surgical Adjuvant Breast and Bowel Project	http://www.nsabp.pitt.edu
European Organisation for Research and Treatment of Cancer	http://www.eortc.be

questions regarding the safety, feasibility and efficacy of new agents or combinations. This chapter provides an overview of the purpose, design, and implementation of phase I, II and III clinical trials and addresses some legal and ethical issues involved in human research.

The Food and Drug Administration (FDA) is a division of the Department of Health and Human Services. Its purpose is to ensure the safety and efficacy of new drugs before they are marketed in the United States. It takes, on average, 8.5 years for a drug to go from the laboratory through animal testing to the completion of clinical trials before it is approved. Prior to entering human clinical trials, a pharmaceutical or institutional application for an Investigational New Drug (INDA) must be submitted to the FDA's Center for Drug Evaluation and Research (CDER). Based on the studies carried out in animal models, the IND application may be approved, and only then can the drug be tested in human subjects.

PHASE I TRIALS

This is the first time a drug will be administered to humans. The goal of a phase I clinical trial is to determine the maximum tolerated dose (MTD) of a new drug or treatment regimen, and to make recommendations regarding the dose and schedule to be used in future studies. In addition, the route of drug administration (oral, intravenous, subcutaneous or intramuscular injection) and information regarding the clinical pharmacology of these new agents are determined. Pharmacokinetics, or the relationship between time and plasma concentration, encompasses absorption, distribution, metabolism and excretion of the drug. These studies are designed to learn about how the patient affects the drug. Pharmacodynamic studies determine how the drug affects the patient by examining the relationship of the plasma concentration and drug effect. Expected and unexpected toxicities are recorded. Phase I trials are neither powered nor designed to determine the efficacy of a drug (although patients are routinely evaluated for clinical response). As such, there is no need to utilize homogeneous populations of patients with regard to disease site or extent of prior treatment. Indeed, individuals who enroll in phase I studies are usually heavily pretreated patients who have failed conventional therapies for their underlying disease or for whom no standard therapy exists. For obvious safety reasons, patients with abnormal end-organ function or poor overall performance status are usually excluded.

In phase I studies, conservative safe starting doses and rules for dose escalation are preset. It has been traditionally assumed that the higher the dose, the more efficacious is the drug, and that the probability of toxicity increases with increasing dose. This presumption has been accurate for conventional cytotoxic agents, but may prove overly simplistic in the evaluation of biological agents that manipulate the host response or perturb cell signalling mechanisms. Animal studies most often utilize rats and mice, but frequently must include larger mammals and subhuman primates, particularly if relevant targets are present in these species. Historically, a safe starting dose has been one-tenth the murine equivalent LD 10 (0.1 MELD 10), expressed in milligrams per square metre of body surface area. The LD 10 is the dose at which 10% of study mice die when that dose is administered. If there is discordance between the mouse and rat data, then beagle dogs are often used as the animal model, given the similarities between human and canine biliary systems. In this case, one-third of the toxic dose low (TDL) in dogs is used as the starting dose. Dose limiting toxicities (DLTs) are clearly defined prior to initiating the study. Toxicities are traditionally graded on a scale of 1–4 using the NCI common toxicity criteria guidelines (National Cancer Institute–Cancer Chemotherapy Evaluation Program, 1999). Provided that no dose-limiting toxicity is encountered in the first cohort, the dose in the second cohort of patients is escalated. Using the modified Fibonacci dose escalation scheme, the second cohort receives a 67% increase of the first dose, the third cohort would receive a 50% increase over the dose given to the second group and the fourth cohort 40% of the dose given to the third group. Subsequent cohorts receive 33% dose increments until a maximum tolerated dose is determined (Collins *et al*., 1986). Usually, 2–3 patients per cohort are used (Simon *et al*., 1997). Typically, three patients are treated initially at each dose level. If none of the patients experiences a dose-limiting toxicity, then the dose is escalated to the next higher level for three additional patients. If one of the three patients experiences a DLT, three additional patients are accrued at the same dose. If two or more of the six total patients in a given dose level experience a DLT, then the MTD has been exceeded and additional patients are treated at the next lower dose level, to confirm that this dose is safe for future studies.

This approach allows a more aggressive initial dose escalation for presumably nontherapeutic doses of the agent. The Fibonacci dose escalation scheme has some disadvantages, such as requiring a relatively large number of patients, and frequently takes many months to complete, as the decision to go to the next level is based on first-cycle toxicities. As the majority of clinical responses occur in patients who receive 80–120% of the recommended phase II dose (Von Hoff *et al*., 1991), there are increasing ethical concerns about treating an excessive number of patients at doses which are retrospectively determined to be subtherapeutic. As a result, there have been a number of attempts to increase the starting dose in a safe manner, and to utilize quicker, yet safe, dose escalation schemes. Two novel approaches include escalation with overdose control (EWOC) (Babb *et al*., 1998) or a modified continual reassessment method (MCRM) (Goodman *et al*., 1995). Under EWOC, dose levels are selected so that the

probability that the dose exceeds the MTD is less than a specified value. The most likely curve describing the probability of a DLT versus dose, based on the experience in the previous patients is computed. This method is statistically designed to converge towards the MTD from doses below the MTD, and controls the probability that any given patient will receive an overdose. Similarly, the MCRM continually modifies the predicted function describing the dose toxicity curve based on the toxicity experienced by all of the patients entered onto the trial. Both of these methods decrease the number of patients required for each dose level, thereby decreasing the number of patients treated at subtherapeutic levels. However, patients still must be observed for defined intervals of time (typically 3–6 weeks) in order to escalate safely to the next highest dose level. The other disadvantage to these accelerated methods of dose escalation is that it limits opportunities to observe interpatient variability in pharmacokinetics and toxicity that might otherwise be seen. As endpoints other than toxicity become more important in identifying working doses for future studies (i.e. measured target effects), assigning fewer patients to each dose level may be detrimental to using potentially unvalidated measures of biological responses to define future working doses. Conceivably, this drawback might be overcome by modifying the designs of phase II studies.

In summary, phase I clinical trials are dose-finding studies to identify toxicity, the MTD and to determine the appropriate dose for further study. Relatively few patients with a variety of primary tumours are enrolled. Historically, only 4–6% of patients enrolled in phase I trials have clinically significant documented tumour responses, but virtually all drugs that are approved exhibit at least some anti-tumour activity in phase I trials. Approximately 70% of the drugs tested in a phase I clinical trial go on to further testing in phase II studies.

PHASE II TRIALS

Once a maximum tolerated dose has been determined in phase I studies, a recommended dose for further study in phase II trials is established. This is usually the dose level below which 33% of patients treated experienced dose-limiting toxicity. The purpose of a phase II trial is to determine the efficacy of the new agent (or combination) in a relatively homogeneous patient population. These trials are designed to eliminate those agents with insufficient activity before embarking on larger, more expensive phase III studies. Patients on phase II trials typically have the same primary tumour, similar prior treatment histories, or may be previously untreated. Typically, if the response rate in a phase II study is less than 20%, then the agent is unlikely to have any significant clinical benefit. If no clinical responses are observed in the first 14 patients treated in a phase II trial, there is less than a 5% likelihood that the employed agent(s) will exhibit a 20% response rate, and the trial can be terminated. However, the number of patients treated in a phase II study is typically greater than the number in any given phase I trial. Additional information regarding toxicity, clinical pharmacology and feasibility is also gleaned in phase II trials.

The endpoint of the phase II study is most often the clinical response rate, although this does not always translate into clinical benefit or increased survival. Therefore, clinical response parameters need to be defined. By convention, a complete response is defined as the disappearance of all clinical evidence of a tumour for a minimum of 4 weeks. A partial response (PR) is usually defined as a 50% or greater decrease in the sum of the products of two perpendicular diameters of all measured lesions for a minimum of 4 weeks, without simultaneous increase in size of any lesion or the appearance of any new lesions. Progression of disease (PD) is defined as an unequivocal increase of at least 25% in the size of any measurable lesion. Stable disease is defined as less than a partial response without evidence of progression, lasting a minimum of 8 weeks. Newer criteria for measuring response has recently been put forth by the NCI (Therasse *et al*., 2000). The RECIST criteria for disease evaluation utilize the sum of the longest diameters of up to 10 identified target lesions. By these criteria, a complete response is defined as the disappearance of all target lesions, a partial response is at least a 30% decrease in the sum of the longest diameter of target lesions and progression is defined as at least a 20% increase in the sum of the longest diameter of the target lesions. Stable disease is defined as neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD. For biological agents that target host response or perturb tumour cell signalling mechanisms, response rate may not be a reasonable study endpoint, and time to disease progression or quantifiable improvement in quality of life may be utilized instead.

There are two types of phase II trial designs: single-stage and two-stage. Single-stage trials employ a fixed sample size (Gehan, 1961). The null hypothesis (that is, that the drug is inactive) is rejected if there is a predetermined number of responders. This may vary according to the indication. For example, the proportion of responders that would justify continued development might be as low as 10% for drug-resistant malignancies such as pancreatic cancer, but might be much higher for low-grade lymphoma, where many drugs have significant antitumour activity as single agents. Potentially active drugs typically go on to phase III study. If the threshold number of responders is not achieved, the drug is deemed inactive and further development for that indication is terminated. In two-stage designs, on the other hand, a given number of patients are enrolled in the first stage of the study. If there is less than a predetermined number of responders, the trial is terminated. If there are more responders, then additional patients are enrolled. By

employing an early stopping rule, the trial is terminated if fewer than expected responses are noted, deeming the agent ineffective. This minimizes the number of patients treated with an agent not likely to provide benefit. Most phase II trials are single-arm studies. Both phase I and II studies are subject to serious selection bias, secondary to investigator bias, and random variation bias, given the small number of patients enrolled.

Randomized phase II trials are frequently employed when several experimental agents are available for testing in a given study population. Patients who meet specific inclusion and exclusion criteria can be randomized to receive one treatment or another. These trials are not designed to be comparative, but may help in selecting the most promising agents or schedule for further study.

In summary, while phase II studies further evaluate safety and toxicity, they are designed to determine efficacy of the drug against a given tumour type. Given the strict patient criteria regarding performance status, extent of prior therapy and anticipated life expectancy, these studies can suffer from selection bias. Frequently, promising phase II results do not hold up in larger generalized populations. Only about one-third of the drugs tested in phase II studies are promising enough to be selected for phase III evaluation.

PHASE III TRIALS

After successfully completing the second stage of trial testing, a drug or combination is deemed safe and reasonably active for a particular disease in a group of highly selected individuals. At this point the drug enters phase III testing. These much larger studies directly compare the agent with a known control, usually the standard treatment for that particular group of patients, in order to determine the relative merits of treatment. Once patients meet eligibility criteria, they are randomized to receive one of the study's treatment regimens. Phase III endpoints are usually response, overall survival or progression-free survival (the time from randomization to death or progression). These trials employ large numbers of patients, and are most often conducted at multiple institutions that include community-based practices. Eligibility criteria are also less restricted, to ensure that the results of the study can be generalized to an unselected population. These trials must be statistically powered to assure that a sufficient number of events are observed in a reasonable length of time.

There are two types of phase III trials: equivalency and factorial. Equivalency trials are designed to demonstrate that the new treatment has efficacy equal to, but not necessarily superior to, that of the standard treatment arm, but may be less toxic. Factorial designs answer two or more questions. In a 2×2 factorial design, patients are randomized to receive treatment A or B. Once randomized, patients undergo a second randomization for another variable, for example differing length of treatment or an additional agent C. There are therefore four treatment arms: patients receiving A only, those receiving B only, those receiving both A and C and those receiving B and C. Randomization decreases chance and bias (unconscious and conscious) and balances predictive factors (known and unknown). Ethical aspects regarding the process of randomization have been discussed at length. It is considered an ethical process in the setting of 'equipoise,' that is, a state of genuine uncertainty on the treating physician's part or that of the medical community regarding the relative merits of different therapies (Fried, 1974).

Statistical Considerations for Phase III Trials

The sample size for any given phase III trial is predetermined and formulated to ensure that adequate numbers of events will occur in a timely fashion, and that differences between the arms can be determined. Sample size is chosen based on four factors: probability of type I error, probability of type II error, power and variability. Type I error (α) is the probability of rejecting the null hypothesis when, in fact, the null hypothesis is true, that is, a false-positive result. A false-negative result or the probability of accepting the null hypothesis, when in fact the null hypothesis is false, is the type II error (β). Power ($1-\beta$) is an estimate of the smallest difference, or treatment effect, that the investigators wish to detect. For a fixed sample size, as the difference between the two treatment arms is increased, the power to detect it also increases. Power increases with increasing sample size. Lastly, variance describes the variability of the outcome measure across the patients in the study group. Greater variability among patients requires larger sample sizes in order to achieve a given power. Studies typically are designed with an α error of 5% and an 80–90% power to detect differences.

Analysis of Phase III Trials

The intention-to-treat principle states that once randomized, all patients are evaluated as members of that group, whether or not the intended therapy was completed. This limits biasing the reported results by excluding patients who withdrew from the study or deviated from the treatment plan for any reason, as those patients might have experienced greater toxicity. If such patients are excluded from the analysis, conclusions drawn are suspect.

Analysing the data prior to final accrual and completion of the trial is called an interim analysis. This has ethical and statistical value in that interim analyses may prevent investigators from further treating additional patients in the event that one arm is clearly more superior or significantly more toxic than the other. It would be unethical to continue treating patients on the inferior arm or withholding more

Table 2 Comparison of phase I, II and III trials

Phase	Purpose	No. of patients	Type of patients	Study design
I	Dose finding	20–30	All tumour types Usually pretreated	Dose escalation
II	Efficacy	30–60	One tumour type Usually no prior therapy	One or two stage
III	Compare with standard	Hundreds	One tumour type Identical cohorts	Randomized Frequently multicentre

effective treatment from the general population. However, as more interim analyses are performed throughout a given trial, the statistical likelihood of detecting a difference between the arms increases spuriously (i.e. increasing the type I error). This can be misleading, and frequently accrual to a randomized trial falls off prematurely owing to physician bias, once the results of an interim analysis have been made available. Therefore, data monitoring committees, independent of the investigators participating in the trial, have been established to perform such analyses.

Phase I, II and III trials are compared in **Table 2**.

META-ANALYSIS

Despite large sample sizes, randomized trials are not always powered to detect small differences that when applied to the general population might have a significant health benefit. Meta-analysis is a reasonable way to draw conclusions about treatments by combining the results of many small, well-designed randomized trials. Basically, all data from randomized trials are pooled and analysed. This includes all relevant randomized trials, both published and unpublished, to eliminate the hazards of publication bias. Most published studies report positive findings while those studies with negative results appear in less widely read journals or are not published at all. Some additional drawbacks to meta-analyses is that these studies are retrospective and not all the information about each trial is readily available. The treatment regimens and patient populations studied have to be similar enough to make conclusions from these studies clinically meaningful.

THE PROTOCOL DOCUMENT

The clinical protocol is a written document that outlines the plan of the study and encompasses a variety of details (**Table 3**). It includes background information which validates the purpose of the trial and defines the study endpoints. The population to be studied is identified through specific inclusion and exclusion eligibility criteria. Patient registration and randomized procedures as well as plans for data collection and data analysis are determined. Drug supplier, starting doses, dose escalations and modifications are predetermined. The biostatistical section of the protocol is perhaps the most essential, as the number of patients needed to support or refute the hypothesis is calculated. The appropriate statistical methods to be used to evaluate the data are defined. Finally, an approved consent form for patient participation is included in the protocol document.

Table 3 Contents of a clinical protocol document

1.0	Introduction
2.0	Objectives
3.0	Selection of patients
4.0	Registration procedures to be used
5.0	Study plan
6.0	Measurement of effect
7.0	Study parameters
8.0	Drug information and procurement
9.0	Statistical considerations
10.0	Adverse event reporting
11.0	Pathology review
12.0	Records to be kept
13.0	Patient informed consent
14.0	References
15.0	Appendix

INFORMED CONSENT

The term 'informed consent' is often felt to refer to the written document that patients sign, agreeing to participate in a clinical trial. However, studies have shown that despite reading such a document which clearly outlines the treatment plan, risks and potential benefits and toxicities that a patient may experience, almost half of patients cannot reiterate this type of pertinent information that the term 'informed consent' implies (Olver *et al.*, 1995). True informed consent has three elements: disclosure, comprehension and voluntary consent. Disclosure of information to patients regarding the nature of their disease, and the proposed treatment plan (with its side effects, risks and potential benefits) as well as alternatives to participating in such a study is carried out by treating physicians through an open discussion. Clinical researchers must ensure that the patient has demonstrated an understanding of this information. The agreement to participate in the study is

entirely voluntary, and patients need to know their rights to withdraw from a study at any time, without undesirable consequences. There are guidelines for writing an informed consent document in the institutional review board guidebook available through the Office of Human Research Protections. The consent form should include the purpose of the study, the treatment procedures and schedule, potential risk and benefits of participating in the study as well as alternatives to treatment. In addition, a consent form should include a person to contact in the event that the participant has any additional questions or concerns regarding the study.

INSTITUTIONAL REVIEW BOARDS

Each institution that conducts clinical research has an institutional review board (IRB). The purpose of the institutional review board is to protect human subjects involved in research. The IRB is made up of a diverse group of individuals, each with an area of expertise. IRB members include medical, administrative, and legal professionals as well as clergy, ethicists, patient advocates and other community members. By law, the board must have at least five members, include both genders and include one member not affiliated with that particular institution. The responsibilities of an IRB include reviewing and approving every clinical research protocol involving human subjects conducted at that institution, to ensure that medically important questions are being asked in a scientifically and ethically sound manner, and to protect the rights and welfare of patients participating in that clinical research. In addition to reviewing the clinical protocol, the IRB reviews the informed consent document to ensure that it is written in a clear and understandable way, and contains all of the necessary components. In addition, protocols are periodically reviewed on a regular basis to include adverse events, and amendments to the protocol.

Once a new drug has been evaluated in all three phases of clinical study, the drug sponsor can choose to submit a New Drug Application. The Oncologic Drugs Advisory Committee (ODAC), composed of research scientists and medical professionals, provides advice to the Food and Drug Administration (FDA) regarding all of the available evidence for safety, effectiveness and appropriate uses of the new agent. Usually this committee must approve the application before the FDA recommends the drug for marketing.

GEMCITABINE IN PANCREATIC CANCER

To illustrate the process of evaluating new drugs in clinical trials, the development of several agents will be reviewed. An example of a conventional cytotoxic antimetabolite, gemcitabine, has undergone phase I through phase III testing. Gemcitabine is a novel nucleotide deoxycytidine analogue with similarities to cytarabine. Intracellular gemcitabine is activated to gemcitabine monophosphate and then to the diphosphate and triphosphate forms. The mechanism of gemcitabine activity is inhibition of DNA synthesis through chain termination and substrate depletion. When gemcitabine is incorporated into DNA, only one deoxynucleotide can be inserted before the DNA polymerase fails to complete further elongation. The depletion of the DNA substrate is achieved by inhibition of the enzyme ribonucleotide reductase by gemcitabine diphosphate.

In preclinical studies, gemcitabine has a broad activity for solid organ tumours including colon, breast, lung and pancreatic tumour cell lines. The drug has undergone numerous phase I studies on several different dosing schedules. When administered daily $\times 5$, gemcitabine caused a high incidence of non-haematological toxicities with a maximum tolerated dose of only 10 mg/m^2. No antitumour responses were noted (O'Rourke *et al*., 1994). When given twice weekly as a 30-min infusion, myelosuppression was the dose-limiting toxicity. MTDs, depending upon duration of infusion, were 65 and 150 mg/m^2 (Popllin *et al*., 1992). Gemcitabine was then administered weekly for 3 weeks by a 24-h continuous infusion with an MTD of 180 mg/m^2. In this setting, neutropenia and lethargy were dose-limiting toxicities (Anderson *et al*., 1996). When given over 30 min every other week, a much higher MTD of 4560 mg/m^2 was identified. The most commonly used weekly regimen, administered 3 weeks on with 1 week of rest, has had the most favourable toxicity profile to date. Further evaluation of the weekly schedule was therefore proposed for phase II study in various solid organ malignancies including lung cancer, ovarian cancer, breast cancer and pancreatic cancer. A phase I/II study was carried out in pancreatic cancer with a starting dose of gemcitabine of 800 mg/m^2 and then in the absence of dose-limiting toxicity, doses were escalated to 1000, 1250 and 1500 mg/m^2. Of the 44 patients entered into this study that were evaluable for response, five had evidence of objective response with a median duration of 13 months. The median survival was 5.6 months with 23% alive at 1 year. Most interestingly, performance status either stabilized or improved in all of the responding patients and documented analgesic requirements were decreased in three of the five responders (Casper *et al*., 1994). A larger phase II trial was carried out in Europe with a regimen similar to that used in the US study for locally advanced or metastatic disease. Of the 32 evaluable patients, two (6.3%) achieved a partial response and six (18.8%) had stable disease. Median survival was 6.3 months and again a similar experience in symptomatic relief, improvement in performance status and decreased analgesic consumption was noted. Neither of these trials had included formal quality of life assessments, however. Therefore, two subsequent

studies – a randomized phase III trial in previously untreated and a phase II trial in 5-fluorouracil (5-FU)-refractory patients – were undertaken (Rothenberg *et al.*, 1996; Burris *et al.*, 1997). The phase III randomized trial compared gemcitabine versus 5-FU in previously untreated patients with locally advanced or metastatic pancreatic carcinoma. A total of 126 patients were assigned to receive either 5-FU 600 mg/m^2 over 30 min weekly or gemcitabine 1000 mg/m^2 for 7 weeks followed by 1 week rest then weekly $\times 3$ every 4 weeks. The primary endpoint of this study was a clinical benefit response as demonstrated by decreased pain intensity, decreased analgesic intake or improvement of Karnofsky performance status. A clinical benefit response was reported in 23.8% of the gemcitabine-treated patients and only 4.8% of the 5-FU recipients ($p = 0.0022$). The median duration of clinical benefit was 18 weeks with gemcitabine and 13 weeks with 5-FU. Tumour responses were dismal in both groups at 5.4% and 0%, respectively. Although the median survival of 5.7 months in the gemcitabine versus 4.41 months in the 5-FU arm were not considerably different, the clinical benefits of receiving gemcitabine were confirmed.

The phase II study that assessed clinical benefit response of gemcitabine enrolled only patients with pancreatic carcinoma who had progressed on 5-FU-based therapy. Similarly, the endpoints in this trial were palliative. Sixty-three patients received gemcitabine using the same regimen as employed in the phase III study with dose escalation up to 1250 mg/m^2 if no grade 2 toxicity had been observed. A clinical benefit response was achieved in 27% of the patients with a median time to response of 3 weeks and a duration of response of 14 weeks. There was no association between clinical benefit and survival. Although the improvement in overall survival is modest, symptom palliation appears to be markedly improved with the use of gemcitabine. This drug received FDA approval and is now considered the standard of care for metastatic pancreatic cancer. The development of gemcitabine illustrates the need to evaluate multiple dosing schedules in a variety of cancers to identify a useful drug and indication, and that objective response rates may not always be the best indicator of clinical benefit. However, the modest efficacy of gemcitabine in pancreatic cancer indicates the urgent need to develop new agents and therapeutic strategies.

THE STORY OF HERCEPTIN AND RITUXAN

As we begin to understand cancer on a molecular level, new therapeutic targets are discovered. Two novel antitumour agents that have undergone extensive clinical trial testing in all three phases will be reviewed. Herceptin (trastuzumab) and Rituxan (rituximab) are monoclonal antibodies that have been rationally designed to target different cellular membrane proteins. Her2/*neu* is a cellular receptor that is frequently overexpressed on the surface of malignant cells, particularly breast cancer. It belongs to a family of receptor tyrosine kinases that play a critical role in the control of cell cycle progression. As such, it is a very attractive target for new agents. In the mid-1980s, monoclonal antibodies were developed against this receptor. A recombinant humanized version of the anti-Her2/*neu* monoclonal antibody, which binds to the extracellular domain of this receptor, entered clinical trials. Three phase I trials were undertaken in patients with tumours that overexpressed Her2/*neu* to evaluate the safety and pharmacokinetics of trastuzumab as a single agent and in combination with cisplatin. The drug was found to be well tolerated when given on a weekly schedule with doses ranging from 10 to 500 mg.

Pharmacokinetic studies revealed an increased serum half-life of trastuzumab with increasing dose. No pharmacokinetic interactions were noted when this new agent was administered with cisplatin. Trastuzumab was further evaluated in two phase II clinical trials to determine its anticancer activity in patients with refractory metastatic breast cancer when given as a single agent as well as with cisplatin. Based on pharmacokinetic parameters, a loading dose was used to achieve therapeutic levels rapidly with continued weekly administration. Objective tumour responses were the efficacy endpoints of these trials. In one study objective responses were seen in five of 43 patients (11.6%) to include one CR and four PRs (Baselga *et al.*, 1996). In the other phase II trial, a 24% partial response rate was demonstrated in 37 evaluable patients (Pegram *et al.*, 1998). After demonstrating efficacy in this group of previously treated refractory patients, this novel agent went on to phase III study in combination with chemotherapy for patients with metastatic breast cancer whose tumours overexpressed Her2/*neu*. In one phase I trial, patients who had not received prior anthracycline therapy were randomized to receive adriamycin and cyclophosphamide alone versus adriamycin, cyclophosphamide and trastuzumab. Those patients who had previously received prior anthracycline therapy were randomized to paclitaxel alone versus paclitaxel plus trastuzumab. Primary endpoints in this trial included time to disease progression with secondary endpoints being response rate, response duration, time to treatment failure and 1-year overall survival. This international trial enrolled 469 patients at 150 clinical sites in 14 countries. The overall response rate of chemotherapy alone was 36.2%. When given with trastuzumab, the response rate increased to 62%. In both drug regimens, the addition of trastuzumab significantly increased the response rate and decreased the time to tumour progression. A significant increase in cardiac toxicity was seen in the adriamycin arm, which illustrates the need for vigilant monitoring as these targeted therapies may not only demonstrate new side effects, but may potentiate known toxicities. As a result, this agent has now

received FDA approval for use in women with metastatic breast cancer as a single agent and in combination with paclitaxel.

Rituximab is another novel chimaeric monoclonal antibody which targets CD20. CD20 is a 35 000-Da phosphoprotein that is found on B lymphocytes including B cell lymphomas. By binding to the CD20 antigen, the monoclonal antibody activates complement, induces antibody-dependent cellular cytotoxicity, inhibits proliferation and directly induces apoptosis. Rituximab was the first monoclonal antibody approved for cancer treatment and perhaps illustrates the minimum amount of time a drug can take from early development to FDA approval. Rituximab entered phase I clinical trials in 1993 in a single intravenous dose trial. Initial doses started at $10\,mg/m^2$ and were escalated to $500\,mg/m^2$ with no significant dose-limiting toxicities noted (Maloney *et al*., 1994). B cell depletion lasting 1–2 months was observed at doses as low as $50–100\,mg/m^2$. A second phase I trial with multiple dosing was undertaken. Rituximab was given weekly $\times4$ at doses ranging from 125 to $375\,mg/m^2$. Tumour responses were observed at all dose levels. Again, no dose-limiting toxicities were noted, illustrating the 'new paradigm' of noncytotoxic drug development. Given the lack of cumulative toxicity, the upper dose level of $375\,mg/m^2$ was chosen for phase II study. In a multicentre clinical phase II trial, 37 patients with relapsed low-grade or follicular non-Hodgkin lymphoma were treated. Seventeen of the 34 evaluable patients demonstrated a partial or complete response for an overall response rate of 50%. Time to tumour progression for these responding patients was approximately 1 year. In a pivotal phase II study conducted at 31 centres, 166 patients with relapsed indolent lymphoma were treated with four weekly doses of rituximab, $375\,mg/m^2$ (McLaughlin *et al*., 1998). The response rate of 48% was comparable to that seen with single-agent cytotoxic chemotherapy. This led to the approval of rituximab in this setting, and illustrates that drug approval need not require a phase III study if it is clearly effective. Commonly observed toxicities in the phase I and phase II trials included fever, chills, pruritus, fatigue and transient hypotension. The more serious toxicities of bronchospasm and significant hypotension were seen in patients with greater tumour burden and it has been postulated that complement mediated rapid tumour lysis may have occurred. As such, it was recommended that infusions of rituximab start at $50\,mg\,h^{-1}$ with premedication with a slow increase in the rate in the absence of symptoms. Most patients can be treated subsequently with minimal side effects.

Rituximab has been studied extensively combined with CHOP-based chemotherapy. In 40 patients with low-grade/follicular non-Hodgkin lymphoma, CHOP was administered at standard doses every 3 weeks for six cycles along with six infusions of rituximab at the recommended phase II dose. Two doses of the antibody were given at the beginning and the end of therapy and single doses before the third and fifth cycles. Thirty-eight patients were treated with an overall response rate of 100% (58% complete response, 42% partial response) with a median duration of response of 39.1 + months. This phase II study (Czuczman *et al*., 1999) demonstrated that rituximab could be safely combined with CHOP and produced a high response rate with a prolonged time to disease progression. Interim results of a large randomized trial evaluating CHOP versus CHOP plus rituximab in elderly patients with diffuse large cell lymphoma were recently presented in preliminary form at the American Society of Hematology Meeting, 2000. The addition of rituximab to CHOP increased the CR (76% versus 60%, $p = 0.004$) and also significantly improved event-free and overall survival (69% versus 49% and 83% versus 68%, respectively) (Coiffer *et al*., 2000). Rituxan has been approved by the FDA for relapsed or refractory CD20 (+) low-grade or follicular non-Hodgkin lymphoma. A phase III trial of front-line CHOP versus CHOP plus rituximab in older patients with high-risk non-Hodgkin lymphoma is currently ongoing through the Eastern Cooperative Oncology Group.

The process of taking a new therapeutic anticancer agent from the laboratory to widespread use is long and arduous. The pitfalls encountered in clinical trial design, implementation and interpretation are numerous. As we enter an era of molecularly targeted new agents, measurement of biological effects will become increasingly vital. Carefully planned and executed clinical and laboratory correlative studies are of paramount importance in advancing the treatment and prevention of malignant disease.

REFERENCES

Anderson, H., *et al*. (1996). A phase I study of a 24 hour infusion of gemcitabine in previously untreated patients with inoperable non-small cell lung cancer. *British Journal of Cancer*, **74**, 460–462.

Babb, J., *et al*. (1998). Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine*, **17**, 1103–1120.

Baselga, J., *et al*. (1996). Phase II study of weekly intravenous recombinant humanized anti-p185 Her-2 monoclonal antibody in patients with HER2/neu overexpressing metastatic breast cancer. *Journal of Clinical Oncology*, **14**, 737–744.

Burris, H. A., *et al*. (1997). Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *Journal of Clinical Oncology*, **15**, 2403–2413.

Casper, E. S., *et al*. (1994). Phase II trial of gemcitabine (2′,2″-difluorodeoxycytidine) in patients with adenocarcinoma of the pancreas. *Investigational New Drugs*, **12**, 29–34.

Coiffier, B., *et al*. (2000). Mabthera (Rituximab) plus CHOP is superior to CHOP alone in elderly patients with diffuse large B-cell lymphoma (DLCL): interim results of a randomized GELA trial. *Blood,* **96**, 223a, abstract 950.

Collins, J. M., *et al*. (1986). Potential roles for preclinical pharmacology in phase I clinical trials. *Cancer Treatment Reports*, **70**, 73–80.

Czuczman, M. S., *et al*. (1999). Treatment of patients with low grade B-cell lymphoma with the combination of chimeric anti-CD20 monoclonal antibody and CHOP chemotherapy. *Journal of Clinical Oncology*, **17**, 268–276.

Fried, C. (1974). *Medical Experimentation: Personal Integrity in Social Policy* (North-Holland, Amsterdam).

Gehan, E. A. (1961). The determination of the number of patients required in a follow-up trial of a new chemotherapeutic agent. *Journal of Chronic Diseases*, **13**, 346–353.

Goodman, S. N., *et al*. (1995). Some practical improvements in the continual reassessment method for phase I studies. *Statistics in Medicine*, **14**, 1149–1161.

Maloney, D. G., *et al*. (1994). Phase I clinical trial using escalating single dose infusion of chimeric anti-CK20 monoclonal antibody (IDEC-C2B8) in patients with recurrent B-cell lymphoma. *Blood*, **84**, 2457–2466.

McLaughlin, P., *et al*. (1998). Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. *Journal of Clinical Oncology*, **18**, 2825–2833.

National Cancer Institute–Cancer Therapy Evaluation Program. (1999). *Common Toxicity Criteria Manual, Version 2.0.*

Olver, I. N., *et al*. (1995). The adequacy of consent forms for informing patients entering oncological clinical trials. *Annals of Oncology*, **6**, 867–870.

O'Rourke, T. J., *et al*. (1994). Phase I clinical trial of gemcitabine given as an intravenous bolus on 5 consecutive days. *European Journal of Cancer*, **30A**, 417–418.

Pegram, M. D., *et al*. (1998). Phase II study of receptor-enhanced chemosensitivity using recombinant humanized anti-p185 HER2/neu monoclonal antibody plus cisplatin in patients with HER2/neu-overexpressing metastatic breast cancer refractory to chemotherapy treatment. *Journal of Clinical Oncology*, **8**, 2659–2671.

Popllin, E. A. D., *et al*. (1992). Difluorodeoxycytidine (dFdC, gemcitabine): a phase I study. *Investigational New Drugs*, **10**, 165–170.

Rothenberg, M. L., *et al*. (1996). A phase II trial of gemcitabine in patients with 5-FU refractory pancreas cancer. *Annals of Oncology*, **7**, 347–353.

Simon, R., *et al*. (1997). Accelerated titration designs for phase I clinical trials and oncology. *Journal of the National Cancer Institute*, **89**, 1138–1147.

Therasse, P., *et al*. (2000). New guidelines to evaluate the response to treatment in solid tumors. *Journal of the National Cancer Institute*, **92**, 205–216.

Von Hoff, *et al*. (1991). Response rates, duration of response, and dose response effects in phase I studies of anti-neoplastics. *Investigational New Drugs*, **9**, 115–122.

FURTHER READING

Dawson, D. and Trapp, R. G. (2000). *Basic and Clinical Biostatistics* (Appleton and Lange, Norwalk, CT).

Piantadosi, S. (1997). *Clinical Trials: A Methodologic Approach* (Wiley, New York).

Websites

http://www.cancertrials.nci.nih.gov
http://trialpages.com
http://ctep.info.nih.gov
http://www.clinicaltrials.gov
http://cnetdb.nci.nih.gov/pdq.html

Index

A
A20 131
ABC (ATP binding cassette) transporters 1348
  anthracyclines resistance 1338
  multidrug resistance 1045–6, 1351
  structure 1348
ABCP (BCRP/MXR) transporter 1338–9, 1045
abdominal plain radiography 1147, 1149, 1153
Abelson murine leukaemia virus
  cultured cell transformation 962
  v-*abl* 964
aberrant crypt foci 252
abl
  chromosomal translocation 27–8, 174, 1106
    chronic myelogenous leukaemia 33, 174, 1106
  v-*abl* 963–4
  see also bcr-abl
Abl tyrosine kinase 99–100, 174, 1363
abscess, computed tomography (CT) 1180
accelerator mass spectrometry (MAS) 298
accutin 1468–9, 1477
acetaldehyde 414
acetaminophen 649–50
acetic acid, percutaneous injection 1261, 1264
2-acetylaminofluorene 277, 282–3, 925–7
5-*N*-acetylardeemin (NAA) 1048–51
N-acetylcysteine (NAC) 427–8, 438–9, 1431
N-acetyltransferase polymorphism 380–1
  bladder cancer susceptibility 380–2
  colorectal cancer susceptibility 549
achalasia 528
acidic fibroblast growth factor (aFGF; FGF-1) 1487
  angiogenesis 237, 240–1, 939
acinar cell carcinoma 583
acini 580
acoustic neuroma 38, 846
acromegaly 609
actin 984, 985
  angiomatoid fibrous histiocytoma 748–9
actin-binding proteins 983–4
actinic cheilitis 493
actinic (solar) keratosis 253, 480–1, 826–7, 1425, 1435
  malignant potential 481
  pathology 480
Actinomyces 1340
actinomycin A 1333
actinomycin D 756, 1333, 1340
  clinical activity 1340
  mechanism of action 1340
  pharmacology 1340
  resistance 1048, 1076, 1341
  retinoids combined treatment 1411
  structure 1340
  toxicity 1340
actinomycins 1340–1
active specific immunotherapy (ASI)
  clinical studies 1061–2
  micrometastases model studies 1061
adamalysin-related membrane proteinases 231
adamantinoma 728–9
ADAMTS 1482
adaptor molecules, tyrosine kinase signalling 166, 168
adenine phosphoribosyltransferase (APRT) 1272
adeno-association viral vectors 1442–3
adenocarcinoma 465
  ceruminal gland 844
  cervix 617–20
  clear cell 661
  colon/rectum 547, 554–5
  endometrioid 629
  gastrointestinal tract 547
  larynx 508
  lung 517
  middle ear 848–9
  oesophagus 525–6, 528, 532–3, 535–7, 540–2
  ovary 637
  pancreatic duct 580–2, 587
  prostate 666–72
  urethra 661–2
adenofibroma 637
adenoid cystic carcinoma 836
adenoids 503
adenoma 465
  adrenal 608–9, 1207
  ceruminal gland 844
  gastrointestinal tract 547
    see also colorectal adenoma
  hepatocellular 576
  middle ear 844–5
  neuroendocrine cell 599–600
  parathyroid 250–1, 600
  pituitary 606–7
  sebaceous 550, 828
  thyroid 600, 604, 606–7
adenosine deaminase 1280
adenosine deaminase deficiency 1280
adenosquamous carcinoma
  lung 518
  penis 676
S adenosylmethionine (*SAM*) 318–9, 441
  dietary precursor deficiency 329
adenovirus
  antisense oligonucleotide/ribozyme delivery vectors 1442–3
  DNA vaccine vectors 1395
  E1A 1071
  E1B 1071
  experimental brain tumour induction 1128–9
  gene therapy vectors 1522, 1527–8
    brain tumours 1135
    endostatin angiogenesis inhibition 1075
    p53 496, 1070–1
    thymidine kinase–ganciclovir system 1536
  haematological malignancies 1104–5
  oncogenic activity 71
  oncolytic activity 1071
    tumour-selective therapeutic strategies 1071
  ONYX-015 mutant 1071, 1427–31
  soft tissue tumours 737

adenylate kinase 1272–3
adherens junctions 215–16, 218–19, 983, 1081–2
basement membrane-induced mammary epithelium differentiation 216, 217
integrins 215
adhesion molecules *see* cell adhesion molecules
adipocytes 1025
differentiation 964
mammary gland development 1026–7
adoptive immunotherapy 1062–3, 1550–1
adrenal adenoma 1207, 608–9
adrenal cortical carcinoma 609
epidemiology 600
adrenal cortical nodules 603
adrenal gland
multidrug resistance proteins (MRPs) 1352
Pgp expression 1349
adrenal tumours
computed tomography (CT) 1165
cortex 603, 608
magnetic resonance imaging (MRI) 1207
medulla 599–600, 602
neuroblastic 794
percutaneous biopsy 1247–8
steroid hormone secretion 600
adriamycin *see* doxorubicin
adult T cell leukaemia/lymphoma 80–4, 707–8
clinical features 80, 707–8
histology 708
leukaemia cells 80
human T cell leukaemia virus type 1 (HTLV-1) 76–7, 80, 341–2, 707–8, 1104–5, 1108
clonal selection of infected T cells 84
pathogenic mechanisms 83–4, 1108
sero-epidemiology 80
Tax protein actions *see* Tax protein
trans-activation of transcription 80–1, 83
immunophenotype 708
Advanced Breast Biopsy Instrumentation (ABBI) system 1564
adverse reactions reporting 1042
aerodigestive tract tumours 505–9
clonality analysis 250–1
tobacco association 402
aetiology
causal inference 259
see also cancer causation inference
epidemiological evidence 265–8
evaluation of evidence 259
hypothesis formulation 260, 265
concomitant variation 265
inference by analogy 265
rule of agreement 265
rule of difference 265
aflatoxin B_1 273–5, 284, 389–90, 438, 1090
DNA damage 284, 299, 1508
adduct formation 296–7, 438
hepatocellular carcinoma 273–5, 379, 389–90
mechanism of carcinogenesis 284, 390
metabolic activation 390
oltipraz protective effect 423
p53 mutations induction 288–9, 299, 379, 390, 1090
permitted dietary contamination levels 390
structure 390
aflatoxin B_2 284
aflatoxin G_1 284, 390
aflatoxin G_2 284, 390
aflatoxin M_1 390
aflatoxins 273–5, 390, 414–5, 1081
chemoprevention target 1433
DNA adducts formation 279
detection in urine 301
DNA damage 299, 568
hepatocellular carcinoma 338–9, 565–6, 568
permitted dietary contamination levels 390
rodent inbred strain responses 367
AFT-1/FUS fusion 748–9
AG-3340 (prinopastat) 1479
Agaricus bisporus 391
agaritine 391
ageing
cancer risk 262
DNA methylation 322, 329
energy restriction response 395
mitochondrial DNA damage 94
aggressive fibromatosis
APC mutations 205–6
nasopharynx 503
aggressive NK cell leukaemia/lymphoma 707
AGM 1470 1471
AgNORs 426
Ah receptor 394
AIB1
pancreatic cancer 592
see also SRC-3
air pollution 1081
polycyclic aromatic hydrocarbons (PAH) exposure 379
white blood cell–DNA adduct measurements 379
AJCC/UICC sarcoma staging system 742–3
Akt signalling
apoptosis regulation 130–1
antimicrotubule agent resistance 1328
invasion/metastasis 225–8
AKT2 592
Alagille's syndrome 568
albinism 481
albolabrin 1477
albumin 570–1
alcohol consumption 260, 262–3, 342, 589, 794
hepatocellular carcinoma 566–8
laryngeal squamous cell carcinoma 508
oesophageal cancer 528
oral squamous cell carcinoma 492–3
rectal cancer 549
alcohol, percutaneous injection 572, 1176, 1261, 1264
hepatocellular carcinoma/liver tumours 572, 1176, 1264
mode of action 1264
ultrasound guidance 1264
aldehydes, tobacco smoke 301, 509
aldosterone 1048
ALK/NPK fusion 691
ALK/NPM fusion 711
ALK/TMP3/TMP4 rearrangements 747–8
alkyl sulfonates 1300
alkylating agents 414–5, 737, 783, 1295–1302
combined drug therapy 1334–5
mechanism of action 1295–6, 1308
DNA damage/adduct formation 279, 295, 298, 1295–1300, 1302
mismatch-repair-defective cancer cell resistance 95
neurocarcinogenesis 1130–1
see also platinum-based drugs
alkylbenzenes, dietary 391
all-*trans*-retinoic acid 966–7, 1031
acute promyelocytic leukaemia treatment 473, 1112, 1414

differentiation induction mechanism 1414–15
resistance 1112, 1414
arsenic trioxide combined treatment 1412
histone deacetylase inhibitor combined treatment 1411–12
in vitro responses 1409–10
retinoic acid syndrome 1414
Allium 424, 438
allopurinol 1287
allylic sulphides 424, 438
α cells 580
α-1-antitrypsin 570–3
α-1-antitrypsin deficiency 568
α-actinin 173, 214, 985
α-catenin 215
α-difluoromethylornithine 429
α-fetoprotein 966
fetal liver production 563
germ cell tumours 473
hepatoblastoma 573
hepatocellular carcinoma 563, 570–2
testicular germ cell tumours 683
tumour marker function 472
yolk sac tumours 682, 813
α-heavy chain disease (immunoproliferative small intestinal disease) 548
α-lactalbumin 966
α-particles 307
α-tubulin 1323
aluminium 300, 512
alveolar carcinoma see bronchoalveolar carcinoma
alveolar rhabdomyosarcoma 739, 752
chromosomal translocations 753
minimal residual disease detection 475
alveolar soft part sarcoma 759
chromosomal translocations 759
alveoli 504
aly 154
American Type Culture Collection (ATCC) 883, 1044
Ames test 352, 425
2-amino-1-methyl-6-phenylimidazole[4, 5-b]pyridine (PhiP) 281, 298
aminoglutethimide 1450, 1454, 1456
aminolaevulinic acid hydrochloride 1425, 1435
aminopterin 1289–90
AML-1 184
AML-1-ETO fusion protein 1407
AML-2 184
AML-3 184
amphiregulin 1486–7
amphotericin B 882
amsacrine 1315–16
AN-9 (pivaloyloxymethyl butyrate) 1406
anabolic–androgenic steroids 783
hepatocellular carcinoma risk 568
anal cancer
epidemiology 265
human papillomavirus (HPV) 336–7
squamous cell carcinoma 547
anaplasia 959
^{18}FDG imaging principle 1228
anaplastic astrocytoma 801, 819, 821, 1123–5
experimental induction 1129, 1131
anaplastic large cell lymphoma 710–11
antigen expression 710
clinical features 710
genetic features 711
NPM/ALK fusion 711
primary cutaneous 711
anaplastic rhabdomyosarcoma 752
anaplastic thyroid carcinoma 600–2, 605–6
anastellin 1476
anastomotic leakage, computed tomography (CT) 1180
anastrazole 1456–7
anchorage-independent growth 961–2
cell transformation assays 964
androgen insensitivity syndromes 684
androgen receptor 135, 140, 673
prostate cancer 135, 1449
androgens 135
promoting activity 287
terminal differentiation induction 966–7
aneugens 361
aneuploidy 88–9, 358–9
in vitro micronucleus assay 359–60
aneurysmal bone cyst 1148
angiofibroma 507
angiogenesis 26, 229–30, 235–45, 472–3, 971, 1028–9, 1072–3, 1524
angiopoietins 1072–3
antiangiogenic gene therapy 1072–4
antibody targets 1384
antisense oligonucleotide targeting 1443
basement membrane degradation 237, 241–2, 937, 939
cell migration 237, 242–3, 939
chemokines 239
cytokines 239, 1028–9
definition 235
Doppler imaging 1189
ovarian cancer 1194–5
endogenous inhibitors 241–2, 1469
endothelial cell proliferation 243, 939
fibroblast growth factors (FGFs) 240–1, 243, 1028, 1072–3
growth factors 229, 239, 971
hypoxia induction/oxygen regulation 237, 977, 980, 1469–70
retinopathy of prematurity 977
invasion 229–30
macrophages 239–40
mathematical modelling 945–7, 950
matrix metalloproteinases (MMPs)/tissue inhibitors of metalloproteinases (TIMPs) 229, 241
mechanisms 237
metastasis 225, 971–2, 999–1000, 1086
models
advantages/disadvantages 971–2
angiogenesis windows 973–4
applications 972–3
chick chorioallantoic membrane assay 974–5
co-culture 979
exteriorized tissue 974–5
GFAP-v-*src* transgenic mouse 1133
immortalized endothelial lines 979–80
in vitro 972–3, 977–80
in vivo 972–7
induction on normally avascular tissues 975
intradermal assays 974
matrix implants 975–7
mouse cornea 1028
physiopathological induction 977
signalling event studies 980
spontaneous tubulogenesis 977–9
oncogene stimulation 1469
platelet-derived growth factor (PDGF) 1072–3
prognostic significance 244–5
systemic markers 244–5
therapeutic strategies 235–6, 442, 1465
transforming growth factor-β (TGF-β) 180, 191, 1072–3
tumour necrosis factor α (TNF-α) 239–40

angiogenesis – *continued*
tumour progression 191, 235–6, 938–40, 971, 1083
'angiogeneic switch' 235
vascular endothelial cell growth factor (VEGF) 238–9, 243–4, 939, 1028, 1072–3, 1088, 1470
therapeutic approaches 21, 1073, 1384
vessel formation/maturation 243–4
angiogenesis windows 973–4
cranial window 973
dorsal skinfold chamber 973
intravital microscopy 973–4
angiogenic factors 235, 237–41, 243–4, 938–9
endothelial cell interactions 939
mathematical modelling 945–7
tumour-associated macrophages production 239
angiogenin 229, 939, 971
angioimmunoblastic T cell lymphoma 708–9
angiolipoma 469–70
angiomatoid fibrous histiocytoma 748–9
paraneoplastic syndromes 743
angiopoietin-1 244, 1469
angiopoietin-2 244, 1469
angiopoietins 1072–3
knockout mouse studies 244
angiosarcoma 465, 563–4, 575–6, 737, 755
bone 729
central nervous system 810–11
heart 777–8, 781, 783
immunohistochemistry 575–6, 745, 755
liver 413
penis 676
postirradiation 486, 755
prognosis 564–5
seminal vesicle 678
skin 486, 755
soft tissue 755
angiostatin 241–2, 971, 1469
angiogenesis inhibition by gene therapy 1073, 1470
mechanism of action 1469
angiotropic (intravascular) B cell lymphoma 704
angiozyme 1445
anilines 33
bladder cancer association 655
animal fats 33
animal models 864–5, 913–21
animal welfare issues 913, 919–20
brain tumours 1126–33
in vivo experimental techniques 1133–5
breast cancer *see* breast cancer models
drug pharmacokinetics 920–1
endocrinological investigations 921
limitations 920–1
metastasis 864–5, 1002–3
mouse models 903
selection principles 915
xenografts *see* xenograft studies
see also knockout mouse studies; rodent carcinogenicity studies; transgenic models
animal welfare issues 913, 919–20
anklyosing spondylitis 265, 311–12
annexin II 1016
annexins 985
anogenital tumours 336–7
anogenital warts 54
antagonistic drug combinations 1051
clinical protocol design 1052–3
combination index (CI) 1052
median-effect plots 1051–2
quantitation 1051–2
selectivity 1052
therapeutic goals 1051
anthracenedione antibiotics 1339
structure 1339
anthracycline analogues 1338–9
P-glycoprotein modulation 1350
anthracyclines 1333–9
clinical activity 1334–5
combined drug therapy 1334–5
mechanism of action 1334
metabolism 1335–6
pharmacology 1335–6
resistance 1076, 1337–8, 1353
structure 1333–4
toxicity 1336–7
cardiac dysfunction 1336
transport 1337–8
antiangiogenic therapy 971, 1465–73
adhesion molecule targeting 1468–9
angiogenesis window assay 973
angiogenic receptor/post-receptor signalling targets 1468
angiogenic stimulator inhibition 1466–8
antisense/ribozyme targeting 1445, 1470
assessment of response 1472
cytotoxic therapy combined treatment 1472–3
endogenous angiogenesis inhibitors 1469
epidermal growth factor receptor (EGFR) blockade 1490
extracellular matrix targeting 1468–9
gene therapy 1072–4, 1470
hypoxia response pathway 1470
immune mechanisms 1472
integrin antibodies 1384, 1477
principles 1465
problems 1472
proteolysis inhibitors 1470–1
see also vascular targeting
antibiotics
antitumour 1333–46
tissue culture media 882, 1043
mycoplasma contamination control 883
antibodies 1379
anticancer effects 1379
chimaeric 1380
fully human 1380
humanized 1380
nomenclature/suffixes 1380
antibody response
DNA vaccines 1397
human papillomavirus (HPV) infection 54–5
antibody-based therapy 1371, 1379–87
angiogenesis targets 1384, 1477
anti-idiotype vaccines 1384
antibody types 1380
antigen targets 1380
biological basis 1379–80
bispecific antibodies 1387
epidermal growth factor receptor (EGFR) blockade 1492
vascular targeting 1471
growth factor receptor blockade 1488–91
immune stimulants 1384
pre-targeting approaches 1386–7
therapeutic targeting agents 1384–7
immunotoxins 1386
radioisotope-labelled antibodies 1385–6
toxicities/infusional syndromes 1382

unlabelled antibodies 1382–4
antibody-dependent cellular cytotoxicity (ADCC) 1379
antibody-directed prodrug enzyme therapy (ADEPT) 1386, 1533
antiemetics 1255–6
antifolate compounds 1287–92
combined drug therapy 1334–5
nonclassical 1291
antigen presentation/processing
following DNA vaccination 1397
MHC molecules 1056–7
specific immune response 1389–90
T cells 172–3
antigen recognition 1056
antigen–antibody complexes, immunohistochemical detection 467–9
heat-induced epitope retrieval (HIER) 468–9
antigen-presenting cells 1056
cellular vaccines development 1391
dendritic cells 1389–92
monocyte-derived APCs 1392
lymph nodes 690
antigenotoxins 421–32
definition 421
endogenous 422
exogenous 423–4
mechanism of action 425–7
see also chemoprevention
antihormones 439
anti-idiotype antibody therapy 1382
anti-idiotype vaccines 1384
antimetabolites 1271–92
antifolate compounds 1287–92
definition 1271
intracellular metabolism 1272
nucleoside analogues 1272
purine analogues 1279–87
pyrimidine analogues 1273–9
threshold concentration 1271
transport systems 1272
antimicrotubule agents 1323–31
combined drug treatment 1327–8
depolymerizing agents 1323, 1329
high concentration effects 1327
historical aspects 1329
mitotic arrest 1324–5
apoptosis 1325–6
mitotic slippage 1325–7
polymerizing (stabilizing) agents 1323, 1329–31
resistance 1328
side effects 1329
stabilizing/destabilizing agent combinations 1331
antimitotic agents *see* antimicrotubule agents
antimycotics, tissue culture media 882
antioestrogens 1426
historical aspects 1451
mechanism of action 137, 1453–4
pure antagonists 1459–60
treatment strategies 1454–5
antioxidants
chemopreventive activity 437–8, 442–3
endogenous antigenotoxins 422, 428
antisense oligodeoxynucleotides (ODNs) 1439–47
biological activity by nonantisense mechanisms 1441–2
clinical trials 1445–6
delivery 1442–3
design 1441–2
formulation 1442
mechanism of action 1439
preclinical studies 1443
'second generation' 1442
signal transduction inhibition 1498
target genes/mRNA 1441
therapeutic potential 1439
anus 545
AP-1 137, 1364–5, 1409, 1451, 1490
AP12 700–1
APC
axin binding 200–1
scaffold protein function 201
Wnt signalling pathway (β-catenin destruction complex formation) 196, 199, 201, 550, 556
APC mutations 29, 195–6, 436, 746, 1091
aggressive fibromatosis 205–6
β-catenin interaction alterations 25–6, 205–6
breast cancer 205–6
cancer-related 204–7, 550, 1090
colorectal adenoma
clonality analysis 252
sporadic 205–6
colorectal cancer 201
progression 32–3, 556
sporadic 205–6
familial adenomatous polyposis (FAP) 37, 205–6, 325, 550
frame-shift 94
Gardner's syndrome 737
gastric cancer 250–1, 542
I1307K 550
oesophageal adenocarcinoma 542
promoter hypermethylation 325
thyroid carcinoma 607
APC/C (anaphase-promoting complex/cyclosome) 1513
APC/C-dependent checkpoint 1515
spindle assembly 1517
aplasia 5
aplastic anaemia, stem cell transplantation 1545, 1546–8
Apo2 (DR4; trail receptor-1) 123–5
Apo2L see Trail protein
apocrine sweat glands 477
apoptosis 4, 13, 17–18, 119–33, 953, 1081–2, 1525
anthracyclines mode of action 1334
antimetabolites mode of action 1281–3
thymidylate synthase inhibition 1275
avoidance by cancer cells 11, 25, 119
adult T cell leukaemia pathogenesis 84
invasion/metastasis 228
mismatch repair-defects 95
β catenin regulation 206–7
caspases *see* caspases
conserved pathways/protein domains 119–20
death domains 153
DNA damage-related 280, 1507–9
cell cycle checkpoints 19
DNA mismatch repair system induction 95–6
extracellular matrix regulation 220
hypoxia-related 1360–1
mitotic arrest 1325–6
multidrug resistance (MDR) 1046–8
mismatch repair defects 95
p53 regulation 29, 891, 1516
platinum-based drug mechanism of action 1305–8
radiotherapy response 1361, 1363–5, 1367
signal transduction pathway interactions 131–2
chemokine receptor signalling 157
transforming growth factor-β receptor signalling 183
tumour necrosis factor receptor signalling 153–4

apoptosis – *continued*
small intestinal crypt cells 7–8
terminally differentiated cells 1403–4
therapeutic approaches 132–3, 430–2, 442
tissue homeostasis 119
topoisomerase inhibition 1313, 1317
apoptosomes 120
apoptotic bodies 119
apoptotic protease activating factor-1 (Apaf-1) 126–30, 132
knockout mouse studies 127–8
appendix 545
APUD (amine precursor uptake and decarboxylation) system 599
aquilide A 390
arachidonic acid 395
metabolism 278, 442–3
chemopreventive intervention targets 432, 436, 439–40, 442–3, 552
arcitumomab (CEA-scan) 1223
ardeemin 1048
areca nut 492
arginine butyrate 1413
Argos 171
Arias Stella reaction 633
aromatase 1455–6
aromatase inhbitors 438
breast cancer 1455–8
classification 1456
competitive 1456–7
suicide 1456–8
aromatic amines
dietary carcinogens 89
DNA damage 281
adduct formation 279, 298
DNA interaction sites 296–7
tobacco smoke 301
aromatic fatty acids, differentiation induction 1406–7
clinical studies 1413
arresten 1476
arsenic 260, 275, 407, 478, 481, 512, 737, 783, 1086–8
arsenic trioxide 1414
all-*trans*-retinoic acid combined treatment 1412
artificial skin reconstructs 1116–17
artificial sweeteners 655
aryl hydrocarbon hydroxylase (AHH) 408–9
arylalkyl isothiocyanates 437–8
arylamines 377–8, 380–1
bladder cancer association 655
N-acetyltransferase polymorphism-related susceptibility 380–2
exposure assessment 377–8
DNA adducts 378
haemoglobin adducts 377–8
asbestos 275–6, 414–5, 418, 498, 512
chromosomal damage 768, 771
fibre dimensions 275–6
mesothelioma association 413, 417, 767, 769–70, 783–4
'bystander'risk 767
preventive measures 769
properties 768
workplace surveillance 419
ascites fluid 467
ascorbic acid see vitamin C
Aspergillus
aflatoxins 275, 284, 390, 568
sterigmatocystin 390
Aspergillus flavus 273–5, 284, 389–90
Aspergillus fumigatus 568
aspirin 423, 428, 432, 549, 554, 1091, 1417
ASPL/TFE3 fusion 759
astrocytes 789–90
astrocytoma 792, 819, 1123
anaplastic 801, 819, 821, 1123, 1125, 1128–9, 1131
animal models 1127
chromosomal changes 821, 1123–4
desmoplastic infantile 802–3, 807
magnetic resonance imaging (MRI) 800–1
NF1/NF2 mutations 30
pathology 800–3
pilocytic 802–3, 1123–4
pleomorphic xanthoastrocytoma 802–3
progression 821, 1123–4
retina 830–1
subependymal giant cell 818
ataxia telangiectasia (Louis–Barr syndrome) 19, 87–8, 104, 117, 382, 568, 636, 836–7, 1363
ataxia telangiectasia mutated see ATM; ATM (Atm)
ataxia telangiectasia-like disorder 104
ATCR see SRC-3
ATF-1/EWS fusion 757
ATM (ataxia telangiectasia mutated) kinase 104, 1515–18
ATM (*Atm*) 1131
mutations
B type chronic lymphocytic leukaemia/small lymphocytic lymphoma 696–7
cell cycle regulation defects 18–19
familial breast cancer 42–3
mantle cell lymphoma 698
radiation-induced DNA damage response 1363, 1366–7
ATM-related (ATR) kinase 1515–18
atomic bombs 311–12, 531–2
atrioventricular benign nodal tumours 777
atypical adenomatous hyperplasia of prostate (adenosis) 668–71
atypical alveolar hyperplasia 509, 513–14
geographical distribution 513
immunohistochemistry 513
molecular genetics 513–14
atypical dyct hyperplasia 250
atypical fibroxanthoma 748
atypical polymorphous lymphoproliferative disorders 718
atypical teratoid/rhabdoid tumour 800
chromosomal abnormalities 820
auricle 841
autoclavable materials 877
autoclave sterilization 876–7
bottles 877
liquids 877, 879
wrapped/boxed items 877
autoclave tape 877
autocrine signalling 4, 953, 1084–5
endothelium 1012
haematopoiesis 6
receptor tyrosine kinases 169, 176
tumour cell growth regulation 25, 1084–5
autoimmune disease, stem cell transplantation 1553
autoimmune gastritis 531
autoimmune lymphoproliferative syndrome 123
automated cell counting 878, 884
avian erythroblastosis virus 962
avian myelocytomatosis virus 962
avian sarcoma virus 1128–9
avidin–biotin complex 745
avidin–biotin peroxidase complex 468
awd 868–9
axillary lymph nodes 612
sampling 615

Axin
cancer-related mutations 200–1, 204–7, 556–7
Dishevelled binding 200–1, 203
scaffold protein function 200–1
tumour suppressor function 206
Wnt signalling pathway (β-catenin destruction complex formation) 196, 199, 201–2
Axin1 206
axons 789
azaarenes, tobacco smoke 301, 405
5-azacytidine (5-azaC) 330–1, 966–7
5′-aza-2′-deoxycytidine 1411
azathioprin 716
chronic graft versus host disease 1548
aziridines (ethylenimines) 1298–9
azo dyes 277
azoxymethane 288
rat colon tumour model 1426

B
B6C3F1 mouse 365
B16 melanoma cells 918, 1116
brain tumour models 1128
gene-modified dendritic cell antitumour response 1075–6
metastasis models 1000–1, 1004, 1116, 1535
B72.3 467
B cell lymphoma 695–705
anti-idiotype antibody-based therapy 1382
Bcl-2 translocation 128, 131
central nervous system 812
diffuse large B cell 702–4
marginal zone 700–2
peripheral 696–705
precursor B cell lymphoblastic/lymphoblastic leukaemia 695–6
WHO classification 694
see also MALT lymphoma (MALToma)
B cells 1056, 1379
bone marrow precursor cells 6
co-stimulation 1056
isolation from blood 1108–9
lymph node follicles 689–90
germinal centres 689–90
Bacillus Calmette-Guerin (BCG) 1421–3, 1434
bacterial artificial chromsomes (BACs) 892–3
bacterial infection-related cancers 344
bacterial mutation assays 352–4
analysis of results 353–4
bacterial strains 352–3
procedure 353
Bad 129–31
Bag1 129–30
Bag3 129–30
BAG4 see SODD (silencer of death domains)
Bak 129–30, 1307–8
expression in cancers 131
Balkan nephropathy 653–4
BALT (bronchus-associated lymphoid tissue) 504
Bannayan–Zonana syndrome 907
Bap31 124–6, 129–30
BAR 124–6, 129–30
barium contrast imaging 1149
Barrett oesophagus
growth factors overexpression 542
oesophageal adenocarcinoma 525, 528, 1432
pathology 536–7
screening/surveillance 532–3
Barrett's model 25
basal cell carcinoma 313, 461, 477–8
clinical features 482–3
clinical management 488
eyelids 827
pathology 482–5
penis 676
prognosis 487–8
ultraviolet radiation association 478, 482
basal cell naevus syndrome see Gorlin–Goltz syndrome
basal epidermal layer 8, 9
Basal Medium Eagle 880
base excision repair 90, 94, 280, 310, 1361, 1509–10
base-free (abasic) sites 90–1, 280
mitochondrial DNA 94
base-free (abasic) site repair 90, 280
basement membrane 5, 210, 987, 1010, 1081
breast tissue 9
cell adhesion 230, 985
degradation 220, 987, 993
angiogenesis 237, 240–2, 939
chemopreventive approaches 432, 442
differentiation induction 216–17
mammary epithelium *in vitro* model 216–17
endothelium 236, 1010
epidermis 9
integrins binding 213
proteins 210
conserved structure 210
differentiation promotion 213
basic fibroblast growth factor (bFGF; FGF-2) 1488
angiogenesis 229, 237, 239–41, 243, 939, 974, 1028, 1072–3, 1490
antiangiogenic therapy 1466–8
breast tissue regulation 9–10
invasion/metastasis 1034–6
ribozyme targeting 1445
basil 391
basiloid squamous carcinoma 494
basophils 1056
batimastat (BB-94) 1479
BAX (Bax) 896, 905–7, 1131, 1307, 1327
apoptosis regulation 19, 129–32, 1307–8, 1364
expression in cancers 131
bronchopulmonary neuroendocrine tumours 511
frameshift mutations 94
colon cancer 96, 556–7
gene therapy 132
Bax, topoisomerase II activity modulation 1317
Bay 12–9566 1470, 1479
BB-94 (batimastat) 1479
BB-25160 (marimastat) 1470, 1479
BCL-1 translocation
mantle cell lymphoma 698
multiple myeloma 730
Bcl-2
antiapoptotic activity 129–30, 1047–8, 1108, 1307–8
bronchopulmonary neuroendocrine tumours 511
dimerization regulation 130–1
follicular lymphoma 699–700
gastroenteropancreatic endocrine tumours 607
necrosis (caspase-independent death) 130
phosphorylation in mitotic arrest 1325–6
synovial sarcoma 758
see also Bcl-2 protein family
BCL-2 (bcl-2) 11, 495, 558, 905–7
antisense oligonucleotide targeting 132, 1443, 1446
antitumour compound molecular targeting 1044

BCL-2 (bcl-2) – continued
translocation in B-cell lymphomas 128, 131
Bcl-2 homology (BH) domains 119–20
Bcl-2 homology-3 (BH3) domain 130, 132
therapeutic approaches 132
Bcl-2 protein family 128–1
antiapoptotic 128–30, 1047
mechanisms 129–30
channel/pore function 129–30
conformational states 130
cytochrome *c* release regulation 128, 130
dimerization regulation 130–31
expression in cancers 131
homologues in lower organisms 129
necrosis (caspase-independent death) regulation 130
proapoptotic 19, 129–30
cisplatin mode of action 1307–8
BCL-6
Burkitt's lymphoma 704
diffuse large B cell lymphoma 703–4
follicular lymphoma 699
Bcl-10 (hE10; CIPER) 127, 556–7a
Bcl-W 129, 1307–8
Bcl-X_L 131
antiapoptotic activity 129–30, 1047–8, 1307–8
dimerization regulation 130–1
expression in cancers 131
BCNU (carmustine) 1300–1
resistance 1076–7
bcr-abl fusion
antisense oligonucleotide targeting 1443
receptor tyrosine kinase inhibition 1500–1
ribozyme targeting 1445
see also Philadelphia chromosome
BCRP (ABCP/MXR) transporter 1045, 1337–9
Beckwith–Wiedemann syndrome 40–1, 573, 647, 737
becquerel (Bq) 308
beer 391–2
beige/nude (*bg/nu*) mouse studies 917
beige/nude/X-linked immunodeficiency (*bg/nu/xid*) mouse studies 917
benign tumours 461
benz[*a*]anthracene 415
benzene 414
benzidine 273, 413, 419
mechanism of carcinogenesis 283–4
benzo[a]pyrene 273, 284–5
cancer chemopreventive strategies 410
dietary 392
mechanism of carcinogenesis 284–5
DNA adduct formation 297, 299
metabolic activation 294, 379, 408
neurocarcinogenesis 1129
occupational exposure 415
rodent breast cancer models 925–6, 933–4
rodent lung tumour models 1426
tobacco smoke 405, 408–9, 509
O^6-benzylguanine 1512
Ber-EP4
endolymphatic sac papillary tumour 848
immunohistochemistry 467
beryllium 275, 512
β cells 580
β-carotene 1432
cancer chemoprevention 411, 430
lung cancer prevention trials 1431–2
β-catenin 218–9, 746
adherens junctions 215
apoptosis regulation 206–7
axin binding 200–1
cancer-related mutations 195–6, 204–7
cell adhesion function 203
cell cycle regulation 1082
destruction box sequence 199–200
oncogenic mutations 199–200, 201
destruction complex 196, 199–201
antagonism 201–3
defects 204–7
E-cadherin interaction 1082
LEF/TCF transcription factor binding 203–4
transcription activation mechanism 204
Wnt signalling 185, 196, 199
intracellular level regulation 199–203
nuclear gene expression regulation 203–4
Wnt-independent regulation 204
β-lapachone 22
β-tubulin 1323
mutations in drug resistance 1328
betacellulin 1486–7
betel 399, 478, 481, 492
oral cancer association 403
Bfl-1 129, 131
bHLH 184
bias 377
case control studies 267
DNA repair–cancer susceptibility studies 383
epidemiological studies 262, 269
mutational spectra studies 378–80
Bid 129–30, 1307–8
bidis 511
Bik 129–30
bile acids 549
bile canaliculi 563–4
bile duct epithelium 10
bile duct stones 574
bilharzia *see* schistosomiasis
biliary cystadenocarcinoma 575
biliary cystadenoma 575
biliary stents 595
biliary system
congenital abnormalities 574
development 563
biliary tumours 577
magnetic resonance imaging (MRI) 1206
telomerase activity 113
ultrasound 1197
bilirubin 422
Bim 129–30, 132, 1307–8
biological safety cabinet 873–6
air flow 875
maintenance/disinfection 876
vacuum source 876
biomarkers 1418–23
chemopreventive drug development 449
DNA adducts 378–9, 438
identification 1420–1
lung cancer screening 419
response markers 1420–1
risk markers 1420
surrogate endpoint biomarkers (SEBs) 1420–3
tissue acquisition standardization 1420
tobacco smoke uptake 408
biopsy 461, 464
CT-guided 1166–9
mutational spectra studies 378

ultrasound guidance 1190
breast cancer 1190
cervical lymph node metastases 1192
thyroid cancer 1192
urothelial carcinoma screening 656
see also percutaneous biopsy
BIR domains 119–20
caspase inhibition 122–3
Birbeck granules 730
biscuits 392
bispecific antibodies 1387
epidermal growth factor receptor (EGFR) blockade 1492
vascular targeting 1471
bisphosphonates 1417, 1480
Bittner factor 924
BK mole *see* dysplastic naevus
BK virus 71, 794
BL-22 1386
bladder
anatomy 645–6
development 645
bladder cancer 654–60
N-acetyltransferase (NAT) polymorphism-related susceptibility 380–2
adenocarcinomas 654, 660
chemical carcinogens
2-acetylaminofluorene 282–3
anilines 33
tobacco smoke 402, 407, 655
chemoprevention 1434–5
dietary factors 379
DNA adduct levels 384
drug resistance 1351
epidermal growth factor receptor (EGFR) overexpression 1488
Fas (CD95) somatic mutations 123
human papillomavirus (HPV) 337
magnetic resonance imaging (MRI) 1209–10
mesenchymal tumours 654
multidrug resistance proteins (MRPs) expression 1353–4
mutiple tumours, clonality analysis 250–1
occupational exposure-related 273, 417
workplace surveillance 419
Rb mutations 908
schistosomiasis association 335, 345, 660
squamous cell carcinoma 654, 660
see also urothelial carcinoma
bladder exstrophy 660
bladder tumour antigen 656
bladder tumours, benign 654
blastocyst microinjection 904–5
bleomycin 1333, 1344–6
clinical activity 1345
mechanism of action 96–7, 1345
pharmacology 1345
resistance 1345–6
structure 1344
toxicity 1345
BLM 100, 104
defect in Bloom syndrome 100–1
S phase nuclear foci 101
BLM mutations 94
blocking agents 421, 427–8, 436–9
blood flow
Doppler imaging 1189
ovarian cancer 1194–5
positron emission tomography (PET) studies 1231
tumours 236–7
blood oxygen level dependent (BOLD) functional MRI 1202
blood transfusion-related infections 78, 567
blood vessels
formation/maturation during angiogenesis 243–4
morphology 236
blood–brain barrier 1123, 1349
disruption for brain tumour metastasis 1125
experimental considerations 1125, 1134
blood–testis barrier 1349
Bloom syndrome 19, 87–8, 117, 381
RecQ-family helicase defects 100–1
see also BLM; *BLM* mutations
blue naevus 835
BMP receptors 183
BMS-214662 1498
BMS-275291 1479
BNP1350 1319
Bok 129–30
bone
cortical 722
development 721–2
epiphysoid 721–2
flat 721–2
growth plates 721–2
metabolism, oestrogen receptorα knockout mouse 136
ossification 722
structure 722
trabecular 722
tubular 721–2
bone infarcts 722–3
bone marrow
cell isolation for culture 1109
haematopoiesis 1102–3
malignant infiltration 1102
metastases 474–5
progenitor cells 6
stem cells 6, 10, 1103
transplantation *see* stem cell transplantation
bone metastases 1001, 1216
magnetic resonance imaging (MRI) 1212
scintigraphic scanning see bone scanning
bone morphogentic protein 1 type metalloproteinases 231
bone scanning 1216–18
^{18}F positron emission tomography (PET) 1231
bone metastases detection 1216–17
false-negative studies 1218
gallium-67 citrate 1222
indications 1216
interpretation 1217
procedure 1216
sensitivity 1217–18
specificity 1218
treatment-related flare phenomenon 1217
bone tumours 721–30
aetiology 722–3
benign 723–4
bone scanning 1216–18
thallium-201 1221
classification 721
epidemiology 722
giant cell 724
magnetic resonance imaging (MRI) 1211–12
occupational exposure-related 311–12
plain radiography 1148–9
syndromic associations 722
borderline ovarian tumours (ovarian tumours of low malignant potential) 461
botryoid rhabdomyosarcoma 752

botryoid rhabdomyosarcoma – *continued*
  middle ear 849
Bournveille disease *see* tuberous sclerosis
Boveri, T. 27, 1523
bovine leukaemia virus (BLV) 77
bovine papillomavirus 4 (BPV4) 53
bovine papillomaviruses (BPV) 50, 53
bowel obstruction
  colonic stenting 1562
  computed tomography (CT) 1180
Bowenoid papulosis 676
Bowen's disease 676, 251
Bowman–Birk soybean trypsin inhibitor (BBI) 438–9
Boyden chamber assays 992
bracken fern (*Pteridium aquilinum*) 390
brain tumours
  aetiology 794
  animal models 1127–33
    external radiation effects 1134–5
    in vivo experimental techniques 1133–5
    induced tumours 1128–31
    intracranial tumour implantation 1134
    spontaneous tumours 1127–8
    transgenic/knockout studies 1131–3
    treatment safety evaluation 1126
  antigenic expression patterns 1123–4
  blood–brain barrier 1125, 1134
  clinical symptoms 794–5
    intracranial expansion 1124–5
    tumour location 1125
  computed tomography (CT) 792–4
  early detection 794–5
  epidemiology 792–4
  experimental design considerations 1123
  extraneural metastasis 1125
  gene therapy 1135
  genetic prodrug activation therapy 1535–6
  grading 1202
  hereditary syndromes 795, 817–20
  human tumour cell models 1126–7
    intracranial 1127
    xenografts 1127
  immunohistochemistry 795
  injection/infusion treatments, experimental approaches 1135
  magnetic resonance imaging (MRI) 1202–3, 792–4
  metastatic 814, 1123, 1125
    experimental model 1134
  microvessel permeability/transport changes 1011
  molecular genetics 1123–4
  pathology 795–814
    germ cell tumours 812–3
    haematopoietic system tumours 812
    mature neuroepithelial tumours 800–8
    meningeal tumours 808–10
    mesenchymal tumours 808, 810–11
    primitive neuroepithelial tissue tumours 795–800
    secondary tumours 814
  percutaneous ablative therapy 1261
  positron emission tomography (PET)
    ^{18}FDG 1216
    protein synthesis markers 1231
  predisposing conditions 795
  thallium-201 scanning 1220–1
  see also central nervous system tumours
branchial arches 841
branchial pouches 841
BRCA1 462, 1363
  breast cancer progession models 866
  multiprotein complexes 104
  oestrogen receptor corepressor function 140
  p53 binding sites 99–100
  Rad51 binding 99–100
  tumour radiosensitivity influence 1362
BRCA1 29–30, 33, 44, 325, 380, 612, 955, 1092–3
  breast cancer association 19–20, 29–30, 33, 42–3, 612, 897
  frame-shift mutations 94
  ovarian cancer association 44, 636, 641
  promoter methylation in tumours 31, 325
Brca1 knockout mouse studies 897–8, 1094–5
BRCA1-associated genome surveillance complex (BASC) 104
BRCA2
  breast cancer progession models 866
  Rad51 binding 99–100
  tumour radiosensitivity influence 1362
BRCA2 29–30, 612, 1092–3
  breast cancer association 19–20, 29–30, 42–3, 612
  frame-shift mutations 94
  ovarian cancer association 636
  pancreatic cancer association 587–9, 592–3
    pancreatic intraepithelial neoplasia (PanIN) 581–2
breast
  benign lesions 613
  developmental phases 9, 611–12, 1026–8
    fetus 9
    pregnancy 9
    puberty 9, 611
  duct system/end buds 9, 611
    development 9
    rat tumour induction models 927–33
    terminal duct lobular unit 611
  ductal cell/alveolar cell differentiation 9
  extracellular matrix–cell interactions 216–7
  lymphatic drainage 612
  mammary epithelium 611
    basement membrane-induced differentiation 216–17
    cell culture 886
    pregnancy-related basement membrane deposition 216
    stromal interactions 1025
    tissue organization loss-related tumorigenesis 218
  menstrual cycle changes 9
  myoepithelial cells 611
    differentiation 9
  oestrogen receptor expression 142–3
  stromal heterogeneity 1029
  tissue organization 9–10, 611–12
    stem cells 9
breast cancer 262–3, 611–15, 1092–3, 1449–51
  advanced disease 1449–50
  aetiology 311, 347–9, 612
  angiogenesis 1028
  animal models *see* breast cancer models
  APC mutations 205–6
  associated benign lesions 613
  Bcl-2 overexpression 131
    therapeutic approaches 132
  BRCA1 association 19–20, 29–30, 33, 42–3, 612, 897
  BRCA2 association 19–20, 29–30, 42–3, 612
  cellular autocrine signalling 25
  cumulative lifetime oestrogen exposure association 33
  diagnosis 613
    mammography 1236–41
  dietary factors 394–5
  drug resistance 1351
  enzyme genetic polymorphism-related risk 304

epidemiology 276–77, 612, 1449
epithelial–stromal interactions 612, 1028–9
Epstein–Barr virus 60–1
familial 19, 29–30, 35, 42–3, 612, 897, 1425
^{18}FDG positron emission tomography (PET) 1217
fine needle aspiration cytology 1564
galactography (ductography) 1241–3
genetic prodrug activation therapy 1537
glutathione-*S*-transferase (GST) isoform expression 1355
grading, Bloom–Richardson criteria 471
HER2/neu overexpression 467, 472–3, 612, 615, 966
  topoisomerase II coexpression 1314–15
integrins expression 1026
interleukin 2 (IL-2) therapy 1375
intraductal carcinoma imaging 1242
magnetic resonance imaging (MRI) 1209
metastases
  bone 1001, 1216
  brain 1125
  liver 575, 1196–7
  middle ear 849
  ovary 640
  trastuzumab (Herceptin) treatment 1383
metastatic spread 228–9, 612
  gene expression changes 869
  micrometastases detection 1058
molecular genetics 612, 1092–3
  metastasis suppressor genes 868–9
  oncogene expression 867–9
multidrug resistance proteins (MRPs) expression 1353
mutifocal, clonality analysis 251
oestrogen dependence 276–7, 1449
oestrogen receptors/receptor status 467, 472, 1449–51
  endocrine manipulation response prediction 1450–1
  expression 142–3
  gene mutations 135, 612
  hormone resistant cancer 144
oxidative DNA damage 89
p53 mutations 29, 473, 612, 615
pathology
  atypical lobular hyperplasia 612, 1092, 1238
  cystosarcoma phyllodes 614–5
  ductal carcinoma *in situ* 1238, 1242, 1092, 1449–50
  invasive ductal carcinoma, not otherwise specified 614
  invasive lobular carcinoma 614
  lobular carcinoma *in situ* 1092
percutaneous ablative therapy 1261
percutaneous core biopsy 1238, 1247, 1250, 1564
  ultrasound guidance 1190
preneoplastic lesions 612
  clonality analysis 250
  see also breast carcinoma *in situ*
preventive treatment in high-risk women 1421–3, 1433–4, 1450, 1458, 1460
prgnostic factors 615
PTEN mutations 30
pulmonary lymphangitic carcinomatosis 1178
Rb mutations 908
risk factors 612, 1029–30
scintimammography 1225
screening 462, 612
  elderly women 1236
  mammography 1235–6
  ultrasound 1236
sentinel lymph node biopsy 472, 1232–3, 1563
src mutations 27–8
staging 1241
  stem cell transplantation 1546–7
  surrogate endpoint biomarkers (SEBs) 1421–3
  telomerase activity 113–16
  tenascin-C expression 1026
  TGF-β receptors
    expression 187–8
    TβRI mutations 186–7
  treatment strategies 615, 1450
    hormone therapy 1449, 1454–5
    oestrogen actions blockade 1458–60
    oestrogen synthesis inhibition 1455–8
    selective oestrogen receptor modulation 1460–3
    surgery 1557, 1563–4
    surgical/radiation-induced ovarian ablation 1450, 1455
  TSG101 mutations 30
  tumour growth rates 254
  ultrasound 1190
breast cancer models 863–5, 868, 915, 923–34, 1092–5
  chemically induced tumours 925–6, 1093–4
    age-dependent responses 929–30
    cell kinetics 932–3
    cell of origin 931–2
    mammary gland differentiation influence 929–31
    mouse 925
    rat 925–6, 927–9
  gene therapy 1070
  genetic background effects 927
  human cell lines 1094
  knockout mouse studies 1094–5
  limitations 933–4
  Min mouse 1091
  radiation-induced tumorigenesis 926–7
  rat/mouse inbred strains 1093–4
  spontaneously developed tumours 924–5
    mouse 924
    rat 924
  transgenic studies 895–6, 1094–5
  tumour progression 865–7
    cell phenotypic diversification 866–7
    cell signalling 866
    genetic changes 865–66
  usefulness of rodent models 923
  xenografts 1094
breast cancer resistance protein *see* BCRP
breast carcinoma *in situ* 612
  ductal (DCIS) 613–4
  lobular (LCIS) 614
  Paget disease of nipple 614
  telomerase activity 113
breast cyst 1190, 1238–41
  terminal duct 1242
breast fibroadenoma 613
breast fibrocystic change 613
breast imaging
  diagnostic mammography 1236–41
  galactography (ductography) 1241–3
  magnetic resonance imaging (MRI) 1238, 1241
  ultrasound 1236, 1238–41
breast papillomatosis 613
  intraductal 1242
breast sclerosing adenosis 613
breast-feeding
  Bittner virus transmission 335
  human T cell leukaemia virus type 1 (HTLV-1) transmission 78, 85, 342
Brenner tumours (ovarian transitional cell tumours) 637, 639–40
bromodeoxyuridine labelling 426, 957–9, 255

bronchi 504
bronchial papilloma 510
bronchial papillomatosis 510
bronchial squamous dysplasia 509
bronchioles 504
bronchoalveolar carcinoma (alveolar carcinoma) 517
  atypical alveolar hyperplasia differentiation 513, 517
  geographical distribution 513
  treatment 517
bronchogenic carcinoma
  asbestos fibre dimensions 275–6
  see also lung cancer
bronchopulmonary carcinoid tumours 510–11
  atypical 510–11
  large-cell neuroendocrine carcinoma 510–11
  molecular genetics 511
  small-cell carcinoma *see* small-cell carcinoma
  typical 510–11
bronchus-associated lymphoid tissue (BALT) 504
bryostatin 1502
budesonide 449
Burkitt-like non-Hodgkin's lymphoma 694, 704–5
Burkitt's lymphoma 704–5
  atypical variants 705
  c-myc activation/translocation 339–40, 704
  classification 694
  clinical features 704
  endemic form 339–40, 704
  Epstein–Barr virus 55, 59–62, 335, 339–40, 1104–5
  geographical distribution 59–60
    malaria relationship 264
  historical aspects 693
  HIV-related 704
  immunophenotype 704
  intestinal tract 548
  p16/*p15* promoter hypermethylation 323–4
  sporadic form 339–40, 704
  t(8;14)(q24.1;q32) 704
burn scars, skin cancer association 478, 481
buserelin 1455
busulfan 1300, 1552
buta-1,3-diene 407, 414
N-butyl-*N*-(4-hydroxybutyl)nitrosamine 1427
butylated hydroxyanisole 410
butylnitrosourea 925–6
butyramide 1406
butyrate 1404–6
  clinical studies 1413
  retinoid combined treatment 1411
butyric acid 549
butyroids 1406
bystander effect
  antibody-mediated antitumour effects 1380
  gene therapy
    brain tumours 1135
    viral transduction efficiency 1525–7
  genetic prodrug activation therapy 1069, 1533, 1534–7, 1540–1
  intercellular communication via gap junctions 1069–70, 1535
  suicide gene therapy 1069
bZIP 183

C

C6 gliomas 1130–1
4C8 gliomas 1128
C68 748–9
C225 antibody therapy 496, 1383, 1488–9, 1500
  antiangiogenic activity 1490
  combined treatment approaches 1491
  pancreatic cancer 1489–90
C1300 neuroblastoma 1128
C/EBPβ 137
CA2.2 590
CA15-3 636
CA19-9 590–1
CA-50 590
CA-125
  endometrial cancer screening 632
  ovarian cancer 636–7, 641–2
  pancreatic cancer 590
  seminal vesicle carcinoma 678
CA-242 590
cadherins 5, 1034
  actin cytosleleton association 1082
  blocking antibody studies 988
  catenin associations 984–5, 1082
  cell adhesion 230, 984
  functional modification studies 988
  malignant invasion 230, 984–5, 999–1000
  negative cell growth regulation 25–6
cadmium 275, 407, 509, 512, 672
caecum 545
Caenorhabditis elegans 964
  basement membrane proteins 210
  programmed cell death studies 1047–8
caffeic acid 424, 427
calcifying epithelioma of Malherbe (pilomatrixoma) 827
calcineurin 129–30
calcium
  chemopreventive actions 437
    clinical trials 1432–3
    terminal differentiation induction 429, 441
  dietary 424, 549
calcium channel blockers 1048
caldesmon 745
Call–Exner bodies 640
CALLA *see* CD10
calmodulin antagonists 1048
calorific intake 394–6
calponin 745
calretinin 242
  mesothelioma 770
Cambridge filter 404–5
CAMPATH-1H (anti-CD52) 1383
Camptotheca acuminata 1318
camptothecin 1318, 1334
  mechanism of action 1317
cancer causation inference
  criteria 269–70
  epidemiological evidence 269–71
  individual exposure 270–1
  tobacco use 401
cancer models, problems/utility 864
  cellular instability 863–4
cancer registry data 264
cancer-prone mutations 19, 35–44
  see also susceptibility markers
canine venereal transmissible sarcoma 335
cannulated sponge model 976
canstatin 1476
capecitabine 1277
capillary haemangioma
  bone 723–4
  orbit 826, 834, 838

retina 831, 838
CAPL/Mts 1 1445
capromab pendetide (Prostascint)
accuracy 1223–4
prostate cancer 1223
carbamates 437–8
carbon black 414
carboplatin 1303, 1501
chemistry 1303–4
resistance 1303
structure 1303
carboxyaminotriazole 1468, 1477–8
carboxylesterase 1319
irinotecan prodrug activation system 1540
carboxypeptidase G2 1386–7
CMDA prodrug activation system 1539
carcinoembryonic antigen (CEA)
cellular vaccine target 1392–3
embryonal carcinoma 813
endometrial adenocarcinoma 633
hepatocellular carcinoma 570–1
immunohistochemistry 467, 570–1, 770
medullary thyroid carcinoma 608
ovarian cancer 636
seminal vesicle carcinoma 678
urothelial carcinoma 657
carcinogenesis 10–11, 937
biomarkers *see* biomarkers
cell division relationship 113
cell immortalization 959
cell proliferation 247–55
chemopreventive intervention strategies 435, 1417
clinical aspects 33
DNA promoter hypermethylation 322
drug development targets 449, 1042
genomic instability effects 87
hyperplasia 10–11
in vitro model limitations 967
in vivo assay 886
interpretation of *in vivo* experiments 899–901
survival analysis 899
molecular mechanisms 31–2, 293, 229–30
initiation events 32, 1085–6
tumour-promotion events 32, 1085–6
multistage model 10, 31–2, 1085–6
two-hit hypothesis 31
organization field theory 250, 967
oxidative DNA damage 89
promotor insertion via provirus integration 76
radiation exposure 311–14
ionizing radiation 311–12, 314
ultraviolet 311, 313–14
somatic mutation theory 967–8
stem cell involvement 11
tumour-suppressor genes 30
carcinogens 259
aetiological hypothesis formulation 260
chemical *see* chemical carcinogens
identification 259
occupational 273, 413–15
physical 307–14
preventive strategies
deactivation/detoxification enhancement 438
inhibition of formation/activation 421, 427, 437–8
inhibition of uptake 437
carcinoid syndrome 510, 599–600
carcinoid tumours 599–600, 609
larynx 508
liver 1259
lung 510–11
middle ear 845
carcinoma 465
aetiological factors 1081
see also epithelial cancers
carcinoma *in situ* 225, 1085
carcinosarcoma 519
CARD domains 119–20, 122–4, 126–7
CARD4 (Nod) 127
cardiac fibroma 777
cardiac malformations 573
cardiac myxoma 778
associated conditions 783
cell of origin 777–8
clinical management 799
clinical symptoms 778–9
electron microscopy 784
embolism 778–9
epidemiology 783
familial 779, 784
immunohistochemistry 784
molecular genetics 785
pathology 778–80
microscopic features 780
prognostic factors 785
cardiac rhabdomyoma 777
cardiac sarcoma 778, 781–2
aetiology 783
cell of origin 777–8
clinical management 786
epidemiology 783
immunohistochemistry 784
molecular genetics 785
prognostic factors 785
see also angiosarcoma; myofibroblastic (intimal) sarcoma
cardiac teratoma 777
CARET 1431–2
CAREX database 418
carmustine *see* BCNU
Carney complex 827–8
Caroli disease 574
carotenoids 430
cartilage 722
endochondral ossification 722
case control studies 261–2, 267
advantages/disadvantages 267–8
bias 267
Casein Kinase Iϵ (CKIϵ) 196, 199, 201, 203
Cash 125–6
CASK 214–5
caspase-1 124
caspase-2 122
caspase-3 121, 153
caspase-4 124
caspase-5 124, 556–7
caspase-7 121
caspase-8 153
caspase-9 122, 131
knockout mouse studies 127–8
caspase-inducible DNA endonucleases *see* CIDE family
caspase-recruitment domains *see* CARD domains
caspases 153
activation 120–1
exogenous protease introduction 121
induced-proximity model 120, 123

caspases
  activation – *continued*
    intrinsic/extrinsic pathways 126
    mitochondrial pathways 123–6
    receptor-mediated 123–6
    safety-catch model 121
  apoptosis 120–2, 1325–6
  cleavage amplification 120
  inactivation
    cancers 122
    IAP (inhibitor of apoptosis) family proteins 122–3
  proinflammatory cytokines processing 124
  protease domains 120
  target proteins 120–1
Casper 124–6
Castleman's disease
  hyaline vascular type 690–1
  Kaposi sarcoma-associated herpes virus association 65, 341
  multicentric (multicentric angiofollicular lymphoid hyperplasia) 691
Cat B 495
Cat D 495
Cat L 495
catenins 230
  cadherin associations 984–5, 1082
  negative cell growth regulation 25–6
  see also α-catenin; β-catenin
cathepsin D 1034–6
cathepsin L 558
cathepsins
  malignant invasion/metastasis 231–2, 985, 999–1000, 1034–6
  therapeutic inhibition 1480
cationic trypsinogen mutations 588
causal inference 259
  see also cancer causation inference
caveolin-1 212
cavernous haemangioma
  bone 723–4
  orbit 826, 834
  retina 831
CB1954 1537–9
CBFA1/AML-3 184
CBP/p300 183–4, 204
  cointegrator function 140, 1453
  SRC coactivator complex 140
CC chemokines 239
CCI-779 1502
CCNU (lomustine) 1300–1
CD1 730
CD2
  adult T cell leukaemia/lymphoma 708
  T cell large granular lymphocytic leukaemia 707
  T cell prolymphocytic leukaemia (T-PLL) 706
CD3 172–3
  adult T cell leukaemia/lymphoma 708
  precursor T cell lymphoblastic/acute lymphoblastic leukaemia
      (precursor T cell lymphoblastic lymphoma) 706
  T cell large granular lymphocytic leukaemia 707
  T cell prolymphocytic leukaemia (T-PLL) 706
CD4 172–3
  T cell prolymphocytic leukaemia (T-PLL) 706
CD4 T cells 1056–7
  cytokine activation 1371–2
  gene-modified dendritic cell antitumour response 1075–6
  mycosis fungoides 486
CD5
  adult T cell leukaemia/lymphoma 708
  B type chronic lymphocytic leukaemia/small lymphocytic lymphoma 696
  diffuse large B cell lymphoma 703
  T cell prolymphocytic leukaemia (T-PLL) 706
CD7
  T cell large granular lymphocytic leukaemia 707
  T cell prolymphocytic leukaemia (T-PLL) 706
CD8 172–3
CD8 T cells 1056–7
  cytokine activation 1371
CD10 (CALLA)
  Burkitt's lymphoma 704
  diffuse large B cell lymphoma 703
  follicular lymphoma 699
  multiple myeloma 729–30
  precursor B cell lymphoblastic leukaemia/lymphoma 696
CD11b/CD18 (Mac-1) 1441–2
CD15 467
  Hodgkin's disease 715
CD16 1387
CD19
  B type chronic lymphocytic leukaemia/small lymphocytic
      lymphoma 696
  precursor B cell lymphoblastic leukaemia/lymphoma 696
CD20 469, 1382
  antibody *see* rituximab
  B type chronic lymphocytic leukaemia/small lymphocytic
      lymphoma 696
  diffuse large B cell lymphoma 703
  lymphoma expression 467, 472–3
  nodular lymphocyte predominance Hodgkin's lymphoma 714
  precursor B cell lymphoblastic leukaemia/lymphoma 696
  radioisotope-labelled antibodies 1385
CD21 (CR2) 55
CD22, precursor B cell lymphoblastic leukaemia/lymphoma 696
CD23
  B type chronic lymphocytic leukaemia/small lymphocytic
      lymphoma 696
  diffuse large B cell lymphoma 703
  follicular lymphoma 699
  lymphoplasmacytic lymphoma 702
CD25
  adult T cell leukaemia/lymphoma 708
  anaplastic large cell lymphoma 710
  immunotoxin target 1386
CD28/CTLA4 1398–9
CD30 152
  anaplastic large cell lymphoma 710
  cutaneous CD30-positive lymphoproliferative disorders 711
  diffuse large B cell lymphoma 703
  Hodgkin's disease 715
  testicular germ cell tumours 683
CD31 810–11, 1020–1
  angiosarcoma 575–6, 729, 755
  Kaposi sarcoma 754
  soft tissue tumours 745
CD33 1386
CD34 6, 1546
  angiosarcoma 575–6, 729, 755
  Kaposi sarcoma 754
  precursor B cell lymphoblastic leukaemia/lymphoma 696
  soft tissue tumours 745
CD38, multiple myeloma 729–30
CD40 152, 1390
CD40L 1076
CD43 467
  B type chronic lymphocytic leukaemia/small lymphocytic
      lymphoma 696
CD44 171, 558

haematopoietic cell interactions 1104
CD45 173–4, 466–7, 469
lymphoma immunophenotyping 467
CD52 1383
CD56 (neural cell adhesion molecule; NCAM)
malignant invasion 230
multiple myeloma 729–30
nasal NK/T cell lymphomas 712
neuroendocrine tumours 602
small-cell lung carcinoma 518
CD57 (Leu-7)
endolymphatic sac papillary tumour 848
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
neuroendocrine tumours 602
T cell large granular lymphocytic leukaemia 707
CD58, multiple myeloma 729–30
CD64 1387
CD68 745
CD79a
B type chronic lymphocytic leukaemia/small lymphocytic lymphoma 696
diffuse large B cell lymphoma 703
multiple myeloma 729–30
nodular lymphocyte predominance Hodgkin's lymphoma 714
CD80 1390, 1398–9
CD86 1390, 1398–9
CD95 *see* Fas
CD99 469, 745
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7
CD138 *see* syndecan
CD rat 365
CD-1 mouse 365
cdc2 17, 1324, 1507
cyclin B complex, mitosis initiation/exit 1324–5
cdc25 17
cdc25A
G_1 checkpoint 1515–17
TGF-β-induced growth arrest 181
cdc25C, G_2 checkpoint 1515–17
cdk1 1512
cdk2 15, 1512
TGF-β-induced growth arrest 180–1
cdk2–cyclin A 141–2
cdk2–cyclin E 1408, 1513–14, 1516–17
cdk4 15, 740, 1507, 1521
Tax protein activation 82
TGF-β-induced growth arrest 180–1
CDK4 (*cdk4*) 750
mutations 1514–15
cdk6 15, 1512
Tax protein activation 82
TGF-β-induced growth arrest 180–1
CDKN2A 772–3
CDKN2B 772–3
cdx 10
CDX-2 556–7
CEA *see* carcinoembryonic antigen
CEA-scan *see* arcitumomab
celecoxib 443, 1424–5, 1427, 1432
cell adhesion
cell culture assays 886
invasion 230
metastasis 229, 984–5
models 985–95
adhesion assays 988–9
applications 993
cell detachment 988
cell spreading 991–2
flow/dynamic adhesion 989–91
limitations 994–5
tumour cells
cell functional changes 987
cell shape alterations 985–6
dynamic conditions of fluid flow 986
endothelial cell interactions 984–5
extracellular matrix interactions 985–6
cell adhesion assays 988–9
hydrodynamic 989–90
limitations 995
cell adhesion molecules 983–4, 1081–2
antiangiogenic therapy targeting 1468–9
blocking antibody studies 988
calcium requirements 988
cell adhesion 230
cell migration processes 8
endothelial cells 1014–17
adhesion assays 1016–17
identification techniques 1015
organ distribution 1016
extracellular matrix 5, 8, 1026
mathematical modelling 945–7
functional domains therapeutic potential 1475
haematopoiesis 1104
immune response 1390
dendritic cells 1391
malignant invasion 230, 999–1000
mutation in carcinogenetic processes 11
negative cell growth regulation 25–6
signalling receptor function 983–4, 1034
tumour cell line expression 1056–7
cell counter 878
cell culture 873
antitumour compounds screening 1043–4
hollow-fibre assay 1044
applications 960–1
cell adhesion studies see cell adhesion, models
cell line establishment 960
chemopreventive agent testing 425, 445
DNA transfection technique 962–3
drug resistant cells generation 1044
genotoxicity testing 354–5
chromosomal damage tests 355–6
haematological cells 1109–10
safety precautions 1109
human brain tumours 1126
limitations 967–8, 994–5
cell cycle 1512–18
antibody markers 958
cell ageing 17
checkpoints 4, 17–20, 25, 1324, 1508, 1512, 1515–17
cancer cells 19–20
DNA damage-related activation 18, 1515–16
heritable defects 19
radiotherapy responses 1363
S phase 18, 102
therapeutic manipulation 21–3, 1518
see also G1 checkpoint; G2 checkpoint; mitotic (M-phase) checkpoint
chromatin structural changes 317
G_0 phase 13, 18, 953–5
G_1 phase 13–17, 953–5, 1324
G_2 phase 13–14, 17, 953–5, 1324

cell cycle – *continued*
historical aspects 13
M phase 13–14, 953–5, 1324
restriction point *see* restriction point (R-point)
rodent chemically-induced mammary carcinogenesis 932–3
S phase 13–14, 17, 953–5, 1324
fraction labelling 957–8
signalling molecules 14
synchronization in cell culture 888
TGF-β-induced growth arrest 180
therapeutic approaches 20–3
cell cycle arrest
butyrate actions 1405
p53 activity 905–7
radiotherapy response 1363–4
cell cycle regulation 13–23, 953–5, 1512, 1526
cancer cell defects 19–20, 1507
cyclin-dependent kinases 1324, 1512–14
extracellular matrix interactions 213
G1 phase 14–17
G2 phase 17
growth factors 14–15, 953–5, 1526
regulatory genes 14
restriction point control 1514–15
S phase 17
therapeutic approaches 1517–18
transforming growth factor-β (TGF-β) 180
cell division 13
base pair loss from chromosome ends 108
regulation 4
stem cells 5
telomere control (divisional clock) 17, 107–9
tumorigenesis relationship 113
see also mitosis
cell freezing 884
solutions 882–3
cell growth regulation
apoptosis *see* apoptosis
extracellular matrix interactions 26
growth factors 25
intercellular signalling 25–6
regulatory gene mutations 25
signalling events/receptors 25
cell lines 883, 964
cross-cellular contamination 883, 1126–7
DNA transfection 962–3
genotoxicity testing 354
metabolically competent 363
human brain tumours 1126
immortalization 961
mycoplasma contamination 883, 1126–7
neoplastic transformation 961–2
osteosarcoma 1117–18
cell migration 214
angiogenesis 237, 242, 938–9
chemotaxis assay 992, 995
extracellular matrix function 8–9
integrins 242–3
in vitro assay 886
motility 987
assays 992
tumour cells 26, 987
invasion/metastasis 231, 999–1000, 1082
cell morphology
changes during adhesive interactions 985–6
cell spreading assay 991–2
under flow conditions 986
extracellular matrix response 213–14
cell proliferation 4, 13, 225, 247–55
anchorage-independent 961–2
angiogenesis 242
antibody cell cycle markers 958
chemopreventive intervention targets 436
differentiation inverse correlation 959
establishment of clones 252–3
extracellular matrix–integrin-mediated 211–12
flow cytometry 255, 958–9
gastric epithelium 524
insulin-like growth factor I receptor (IGF-IR) 965–6
metaphase arrest technique 958
metastases 999–1000
population kinetics 957
prognostic significance 255
rate in tumours 254–5, 957
measurement 255, 957–9
regulation 4, 953
defects in cancer development 4
rodent chemically-induced mammary carcinogenesis 932–3
S phase fraction labelling 957–8
Tax protein in adult T cell leukaemia pathogenesis 83
transforming growth factor-β (TGF-β) inhibition 180, 186
cell proliferative index 957
antibody markers 958
antigenotoxin assessment 426
S phase fraction labelling 957–8
cell regeneration (replacement) capacity 3–4
conditionally renewing tissues 4
non-renewing tissues 4
rapidly self-renewing tissues 4
cell-permeable peptides 965
cellular hierarchies 4–5
cellular immune response
early tumour growth 938
human papillomavirus (HPV) infection 55
cellular vaccines 1389, 1391–4
adjuvants 1391–3
autologous antigen-presenting cells 1391
clinical trials 1393–4
dendritic cells 1391–2
immune response markers 1392
immunostimulatory cytokines production 1391
tumour cell lines 1391
tumour-specific peptides 1392
central nervous system development 789
central nervous system tumours 789, 1123–5
biological categories 791–2
clinical management 821–2
chemoembolization 1256
stereotactic surgery 1564
cytogenetics 820
cytokine secretion 1124
DNA mismatch repair defects 550
epidemiology 792–4
grading 821
growth patterns 1123–4
magnetic resonance imaging (MRI) 1202–3
models 1123–36
animal 1127–33
applications 1126–33
human tumour cells 1126–7
in vivo experimental techniques 1133–5
interpretation of results 1135
molecular genetics 817–21
pathology 790–2

anaplasia 1123–4
primary 1123
prognostic factors 821
secondary (metastatic) 792
stem cell transplantation 1546–7
telomerase activity 113
treatment-related inflammatory response 1125–6
see also brain tumours
central neurocytoma 807
centrifuge 878
centroblasts 689–90
centrocytes 689–90
centrosome abnormalities 88–9
ceramide 1364–5
Cerberus 199, 204
cerebellar dysplastic gangliocytoma (Lhermitte–Duclos disease) 806, 820, 907
ceruminal gland adenocarcinoma 844
ceruminal gland adenoma 844
pleomorphic 844
ceruminal gland tumours 843–4
classification 843
ceruminal glands 841
cervical cancer 617–27
adenocarcinoma 617–20
adenoma malignum variant 617–20
aetiology 620
tobacco use 402
chemoprevention 1434–5
clear cell variant 617–20
clinical management 627
epidemiology 263, 265, 620
human papillomavirus (HPV) 54, 336–7, 617–20, 626, 1435
HPV 16 335, 337
hybrid capture technique 623–5
immune response (E2/E4 antibodies) 54–5
in situ hybridization 625–6
screening 337, 349
magnetic resonance imaging (MRI) 1208
molecular genetics 626–7, 1314–15
mouse model 1426
multidrug resistance proteins (MRPs) expression 1353
pathology 617–20, 626
gross findings 626
premalignant lesions 617
prognostic factors 627
recurrence 1178
screening 462, 621–5
HPV testing 621, 623–5
Pap smear 621–3
small-cell carcinoma 617
squamous cell carcinoma 617
squamous intraepithelial lesions (SILs) 617, 617–20
high-grade 623, 626
low-grade 623, 626
cervical intraepithelial neoplasia (CIN)
CIN2 251–2
CIN3 251–2
clonality analysis 251–2
human papillomavirus (HPV) 336–7
progression 336
squamous intraepithelial lesions (SILs) 617–20, 623, 626
cervical lymph node metastases
magnetic resonance imaging (MRI) 1204
ultrasound 1190–2
cervix 617
CGP 41251 (Midostaurin) 1501–2
CGP 57148B (imitanib; Glivec; ST1571) 1497, 1500–1
CGS 27023A 1479
checkpoint kinases 1515–18
therapeutic inhibition (UCN-01) 1518
chemical ablation therapy 1261
chemical carcinogens 33
active metabolites 296–7
brain tumour models 1129–31
breast cancer models *see* breast cancer models
colorectal cancer models 1090–1
critical gene targets 288–9
direct-acting 275
DNA damage 89, 278–80, 293–5, 925, 933
adduct formation 293, 295–8
DNA reaction sites 296–7
epigenetic 278
exposure assessment 377–9
extracellular matrix effects 1088–9
free radicals generation 278–9
hormones 276–7
inert 275–6
inorganic 275
lung cancer models 1092
mechanism of action 273–90
metabolic activation 277–8, 294–5, 407–8, 425–7
in vitro system 352
multistep carcinogenesis 286–8
mutation induction 280–1
mutational fingerprints 281, 379, 379–80
organic 273–5
promoting agents 287–8
screening
expert computer systems 362–3
high-throughput 362
soft tissue sarcomas 737
thresholds for exposure 361–2
tobacco smoke 301, 405
see also genotoxicity testing; occupational carcinogens
chemoattractants, invasion/metastasis 231–2
chemoembolization 1255–60
ancillary medication 1255–6
applications 1256
contraindications 1259
liver tumours *see* liver tumours
metastatic tumours 1259
principle 1255–6
procedure 1255–6
radiofrequency ablation combined treatment 1257–9, 1263
chemokine receptors
endothelial cells 239
HIV-1 coreceptor activity 155
signal transduction 154–7
apoptosis 157
focal adhesion kinases 156–7
heterotrimeric G-protein interactions 155
phosphoinositide 3-kinase γ 156
phospholipases 155–6
tyrosine phosphatases 157
structural aspects 155
chemokines 154
angiogenesis 239
functional aspects 154
G-protein-coupled receptors 155
metastasis 229
chemoprevention 435–51, 1417–36
agent therapeutic index enhancement 1424–5
bladder cancer 1434

chemoprevention – *continued*
  blocking agents 421, 427–8, 436–8
  breast cancer 1421–3, 1433–4, 1450, 1458, 1460
  carcinogen formation inhibition 421, 427
  cervical cancer 1434–5
  colorectal cancer 552, 1427, 1432–3
  efficacy evaluation 425, 444–5
    animal models 425, 445, 449
    cell-based assays 425, 445
    mechanistic assays 444–5, 449
    transgenic/knockout mice 445
  future development 1435–6
  gastric cancer 1432
  goals 1427
  head and neck cancer 496, 1431
  hepatocellular carcinoma 1433
  lung cancer 1092, 1431–2
  molecular targets 435–6, 443–4
  oesophageal cancer 1432
  prostate cancer 1434
  regional delivery systems 1425
  skin cancer 1435
  suppressing agents 421, 429–32
  target sites 432
  test systems 425
  tobacco smoking 410–11, 1431
  toxicity/safety standards 1424–5
  see also antigenotoxins
chemopreventive drug development 445–9, 1425–35
  agent combinations 449, 1425
  carcinogenesis biomarkers 449
  clinical trials 425, 448–9, 451, 1426–7, 1427–35
    phase I studies 450–1, 1426–7
    within 'therapeutic' trials 1425
  dietary components 449–50
  genotoxicity tests 450
  lead compound identification 1425
  pharmacokinetic testing 450
  preclinical efficacy development 449–50, 1426
    carcinogen-induced tumour models 1426
  preclinical toxicity testing 450
chemotherapy 20–1, 473
  bladder cancer 660
  breast cancer 615
  clash hypothesis 18
  DNA adduct detection in tissues 301
  drug resistance *see* drug resistance
  interleukin 2 (IL-2) combined therapy 1376
  islet cell tumours 595
  mesothelioma 774
  myelosuppression 1076
  nephroblastoma (Wilms' tumour) 648
  neuroendocrine tumours 609
  oral cancer 496
  pancreatic cancer 594
  selective protection of normal cells 23
  soft tissue sarcomas 761
    metastatic disease 762–3
  trans-catheter therapy *see* chemoembolization
chest radiography 1147
  Kerley B lines 504–5
  lung cancer screening 1150
chfr 1517
chick chorioallantoic membrane (CAM)
  angiogenesis models 242, 974–5, 980
  metastasis models 1003
    videomicroscopy 1000–1
chief cells 524, 528–9
chimaeric antibodies 1380
chimaeric mice 904–5
Chinese hamster cell genotoxicity assay 354
Chk1 checkpoint kinase 1515–7
  therapeutic inhibition (UCN-01) 1518
CHK2 42, 1363
  mutations 737
Chk2 checkpoint kinase 1515–8
  cellular irradiation response 1363
chlorambucil 1296
  resistance 1356
  structure 1296
2-chloroethylnitrosoureas 1300–2
  carbamoylation activity 1302
  DNA adduct formation 1301–2
  resistance 1510
chloroform 414
chlorophenoxy herbicides 414
chlorozotocin 1300–2
chlorpromazine 1048
CHN fusions 738–9, 757
CHO 354
cholangiocarcinoma 564, 573–5
  aetiology 573–4
  diagnosis 575
  epidemiology 564, 573
  immunocytochemistry 575
  liver fluke infection association 347, 573–4
  mixed hepatocellular carcinoma 575
  molecular genetics 575
  pathology 574–5
    macroscopic 574
    microscopic 574–5
  precancerous changes 574
  presentation 575
  prognosis 564–5, 575
  treatment 575
  ultrasound 1197
choledochal cysts 574
cholelithasis 577
cholesteatoma (keratoma) 842–3
cholesterol 599
chondroblastoma 723
chondrocytes 722
chondroid chordoma 728
chondroitin 210
chondroma 503
chondromyxoid fibroma 723
chondrosarcoma 465, 725–6
  clear-cell 726
  clinical management 726
  cytogenetic abnormalities 726
  dedifferentiated 726
  genetic predisposition 722
  grading 725–6
  larynx 508
  mesenchymal 726
  nasopharynx 503
  pathology 725
  plain radiography 1149
  prognostic factors 726
CHOP/EWS fusion 738–9, 750
CHOP/TLS fusion 750
chordoma 727–8, 811–12
  chondroid 728
  dedifferentiated 728

choriocarcinoma 465–6
central nervous system 813
immunohistochemistry 683
ovary 640
testis 1192–3, 682
choristoma 827–8
choroid plexus 789–90
multidrug resistance proteins (MRPs) 1352
choroid plexus carcinoma 808
choroid plexus papilloma 808
experimental induction 1128–9
choroid plexus tumours 792
pathology 807–8
choroid tumours 833
choroidal melanoma 833, 838
choroidal naevus 833
chromatin
condensation 320–2
structure 317–18
chromium 275, 407, 505, 512
chromogranin 467
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7
islet cell tumours 586
jugulotympanic paraganglioma 846
neuroblastic tumours 815
neuroendocrine tumours 602
small-cell carcinoma 671
cervix 617
lung 518
chromogranin A 602
chromogranin B 602
chromophobe renal cell carcinoma 651–3
chromosomal abnormalities
asbestos-induced 768, 771
central nervous system tumours 820
chondrosarcoma 726
early response markers 377
fatty tumours 469–70
genotoxicity testing
aneuploidy tests 358–60
exposure thresholds assessment 361–2
in vitro cell culture 355
rodent models 355
head and neck cancers 495
hepatocellular carcinoma 571
lung cancer 512
lymphomas 473, 692–3
oncogenesis 27
osteosarcoma 724
pulmonary hamartoma 509–10
salivary gland carcinoma 498–9
chromosomal translocations
abl activation 27–8
acute promyelocytic leukaemia (APL) 473, 909
alveolar soft part sarcoma 759
anaplastic large cell lymphoma 711
angiomatoid fibrous histiocytoma 748–9
Burkitt's lymphoma 704, 1105–6
c-*myc* activation 1105–6
clear cell sarcoma (malignant melanoma of soft parts) 757
congenital fibrosarcoma 747
dermatofibrosarcoma protuberans 748
desmoplastic small round tumour 759
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 722–3, 756
follicular lymphoma 699–700
leukaemias 1111
liposarcoma 469–70
lymphoma 692
MALT type marginal zone B cell lymphoma 700–1
molecular diagnostic methods 475
multiple myeloma 730
myxoid chondrosarcoma 757
myxoid/round cell liposarcoma 750
oncogene expression changes 867–8
rhabdomyosarcoma 753
sarcomas 470
diagnosis 473–4
soft tissue sarcoma 737–9
synovial sarcoma 470, 758
tyrosine kinase displacement 176
chromosome 11 deletions, mouse models 898–9
chromosome condensation 1324–6
chromosome engineering 898–9
chromosome instability
molecular origins 88
tumours 88
chrysene 273
chutta (reverse smoking) 492
CI-1033 1500
cIAP-1 131
cIAP-2 (cIAP-2) 122–3, 131
CIDE (caspase-inducible DNA endonuclease) family 119–21
cigarette smoke condensate (CSC) 404
cigarette smoke/smoking *see* tobacco smoke; tobacco smoking
ciliary body 825–6
ciliary body medulloepithelioma 831
ciliary body tumours 833
CIPER *see* Bcl-10
circumcision 676–7
cirrhosis, hepatocellular carcinoma 568–9, 572
screening 569
cisplatin 660, 685, 774, 1071, 1302–8
analogues 1303
apoptosis induction 1307–8
chemistry 1303–4
clinical utility 1303
combined p53 gene therapy 1528
DNA interactions 91, 1304–7
historical aspects 1303
irinotecan sulphate (CPT-11) combination 1320
liver tumour chemoembolization 1256
mechanism of action 1303, 1305
neuroendocrine tumours 609
p53 activation 1307
resistance 1303, 1353, 1356
DNA mismatch repair system defects 95, 1314–15
ribozyme reversal 1445
side effects 1303
structure 1303
citrus fruits 410, 424
cladribine 1284
Clara cells 504
classification, tumour 465–71
electron microscopy 466
histogenesis 465
immunohistochemistry 466–9
molecular studies 469–71
clastogens 361–2
clear cell adenocarcinoma 661
clear cell carcinoma
endometrium 629
ovary 637, 639

clear cell chondrosarcoma 726
clear cell renal cell carcinoma 649–53
clear cell sarcoma (malignant melanoma of soft parts) 756–7
  chromosomal translocations 757
  EWS fusions 738–9
  immunohistochemistry 745
cleft palate 573
clinical drug trials 1569–76
  chemopreventive agents 425, 448–9, 450–1
  good clinical practice (GCP) 1041
  IND (investigational new drug) applications 1041–2
  informed consent 1573–4
  institutional review boards 1574
  meta-analysis 1573
  phase I 448–51, 1042–3, 1570–1, 1573
  phase II 448–9, 1042–3, 1571–3
  phase III 448–9, 1042–3, 1572–3
  protocol document 1573
  starting doses 921, 1570–1
clofibrate 1408
clonal nature of cancers 247, 1085
  clonality evaluation *see* clonality analysis
  establishment of clones 252–3
  second primary tumours, oral cavity 493–4
clonality analysis 247–8, 606
  atheromatous plaque 248–9
  hyperplastic lesions 251–2
  intestinal crypts/gastric glands 252–3
  intestinal metaplasia 252
  lymphoma/lymphocytic proliferations 692
  patch size influence 248–9
  post-transplantation lymphoproliferative disorders 717–18
  preinvasive lesions 250–2
  preneoplastic lesions 249–50, 252
  synchronous/second primary/multifocal tumours 250–1
  X chromosome inactivation 247–8, 606
cloning, tissue culture technique 885–6
clonogenic stem cells 7–8
Clonorchis sinensis 347, 573–4
clustering 213–4
CM101 1471
CNS-1 glioma 1131
CNT1 1272, 1278–9
CNT2 1272
CO_2 buffer system, cell culture 881
coactivators
  definition 138
  oestrogen receptor function 138–40, 1452–3
    CBP/p300 140, 1453
    nuclear receptor (NR) box (receptor-interacting domain) 139
    SRC family 138–40, 1453
coal tars 273, 413, 415, 512, 1086–7
  skin cancer 478, 481
cobalt 275
cochlea 841
Cockayne syndrome 87–8
  cellular characteristics 92–3
  clinical features 92
  CS-A/CS-B complementation groups 92
    gene products 92–3
  transcription-coupled repair (TCR) defect 92–94
Cockayne syndrome group B (CSB) protein 91
co-culture angiogenesis models 979
coffee 392, 589, 636, 655
cohort studies 261–2, 266–7
  advantages/disadvantages 267–8
  closed 267
  open 267
  prospective/retrospective 266–7
COL1A1/PDGFβ fusion 748
COL-3 (metastat) 1480
colcemide 1331
colchicine 1329, 1331, 1435
  P-glycoprotein modulation 1048
Coley's toxins 152
collagen
  angiogenesis induction 971
  antineoplastic/antiangiogenic fragments 1476
  basement membrane 210, 236, 987, 1010–12
  extracellular matrix 5, 8, 210, 1026
    degradation during malignant invasion 231
    TGF-β regulation of expression 181
  integrins binding 210, 242, 1026
collagenase 1 1034–6
collecting duct carcinoma 651–2
COLO320 1091
colon
  development 545
  gross appearance 545
colonic stenting 1562
colorectal adenoma 547, 552–4, 1421
  adenocarcinoma differentiation 552–3
  cell proliferation rate 252–3
  clinical management 559
    transanal endoscopic microsurgery 1562
  clonality analysis 252
  dysplasia 552
  early morphogenesis 553–4
    aberrant crypt foci 554
  endoscopic screening 551, 1421
  malignant change 547, 552
  pathology
    macroscopic features 552
    microscopic features 552–3
  serrated 554
  telomerase activity 113
  tubular 552
  tubulovillous 552
  villous 552
  see also colorectal polyps
colorectal cancer 19–20
  adenocarcinoma 547, 554
    grading 555
  adenoma–carcinoma sequence 552, 556, 955
  aetiology 548–50
  APC mutations 201, 205–6
  Axin mutations 200–1, 206
  Bcl-2 overexpression 131
  cellular vaccines 1391
  chemoprevention 552, 1427, 1432–3
  chronic inflammation association 550
  clinical management 558–60
    adjuvant therapy 559–60
    arcitumomab 1223
    early cancer 559
    hand-assist laparoscopic surgery 1561
    interstitial thermotherapy 1565
    laparoscopic surgery 1559–60
    radioimmunoguided surgery 1565
    surgery 559
    transanal endoscopic microsurgery 1562
  dietary factors 265, 394–5, 549
    preventive approaches 552
  DNA mismatch repair defects 96

Dukes classification 557
epidemiology 265, 548
migrant studies 264, 548
erbB-2/neu expression abnormalities 966
familial adenomatous polyposis (FAP) 36–7, 43
^{18}FDG positron emission tomography (PET) 1216
genetic high-prevalence polymorphisms 549
genetic instability 556–7
glutathione-*S*-transferase (GST) isoform expression 1355
hepatic metastases 575, 1196–7
hereditary non-polyposis colorectal cancer (HNPCC) 43–4
immune response 556
inherited syndromes 549–50
genetic/predictive screening 551–2
invasion/metastasis 228–9, 555
Jass prognostic classification 558
magnetic resonance imaging (MRI) 1207–8
metastatic (secondary) tumours
chemoembolization 1259
interleukin 2 (IL-2) therapy 1375
microsatellite instability 96, 556–8
mutator phenotype 556–7
TGFβ-RII frameshift mutations 96, 556–7
models
chemically induced tumours 1090–1
Min mouse 1091
spontaneously arising animal tumours 1090–1
xenografts 1091
molecular genetics 36–7, 556–7, 1090
oncogene expression 867–8
multidrug resistance proteins (MRPs) expression 1353–4
mutiple tumours, clonality analysis 250–1
ovarian metastases 640
p53 mutations 29
P-glycoprotein expression 1350–1
pathology
gross appearances 554
histology 554–6, 558
percutaneous biopsy 1250
preinvasive lesions 552–4
preoperative staging 559, 1558
prevention 550–2
prognostic factors 557–8
progression, Vogelstein's model (Vogelgram) 32–3
satumomab pendetide 1223
schistosomiasis association 345
screening 550–1
computed tomography (CT) 1158–9
helical endoluminal CT colonography 1158–9
signet ring cell carcinoma 554
src mutations 27–8
staging
computed tomography (CT) 1165–6
TNM classification 557
surrogate endpoint biomarkers (SEBs) 1421
telomerase activity 113–15
tobacco smoking association 549
transforming growth factor-β (TGF-β) 96
receptor expression 187–8
TβRII mutations 96, 186–7, 556–7
tumour-suppressor genes 29
Wnt signalling pathway 195–6, 205–6
colorectal hyperplastic polyposis 547
colorectal polyps 29, 547
clonality analysis 252
computed tomography (CT) 1158
hamartomas 547
helical endoluminal CT colonography 1158–9
hyperplastic 547
malignant potential 547
see also colorectal adenoma
coltsfoot 390
combination index (CI) 1052
Combretastin A4 1471
comet assay 358
comfrey (*Symphytum officinale*) 390
common bile duct 563–4
tumours 577
comparative genomic hybridization (CGH) 899–900
lymphoma 693
tumour-suppressor gene identification 31
'Compare Program' 1043
complement-dependent cytotoxicity 1379–80
computed tomography (CT) 1147, 1149, 1155–83, 1558
brain tumours 792–4
complications investigation 1180–2
data acquisition 1155–6
diagnostic applications 1160–1
hepatic metastases 1165–6
image reconstruction 1156
lung cancer screening 1150
metastatic tumours 1149
multidetector-row helical scanning 1158
pancreatic cancer 589
percutaneous ablative therapy guidance 1262–3
radiotherapy guidance 1176
renal tumours 1197
scanner components 1155
scanner evolution 1156, 1158, 1182–3
screening applications 1158–9
soft tissue tumours 743–4
software advances 1183
spiral/helical scanning 1156
three-dimensional 1183
staging applications 1162–6
stent placement guidance 1176
surgical planning 1170–6
surveillance applications 1177–8
tissue sampling guidance 1166–9
percutaneous biopsy 1248
ultrasound comparison 1188
virtual imaging techniques 1183
window settings 1156
see also helical endoluminal CT colonography
concentrative nucleoside transport (CNT) system 1272
conditional gene knockouts 894–5, 897–8
conditionally renewing tissues 4, 9–10
cancer origins 10–11
cone biopsy 627
confluent cultured cells 888
confounding factors 260, 262, 268–9, 377
DNA repaircancer susceptibility studies 383
mutational spectra studies 378, 380
congenital fibrosarcoma 747
congenital hypertrophy of retinal pigment epithelium 37, 831–2
congenital melanosis 826–7
conjugation reactions 277, 294, 380–1, 427
conjunctiva 826
conjunctival dysplasia 826–7
conjunctival intraepithelial neoplasia (carcinoma-*in-situ*) 827
conjunctival melanoma 827
predisposing lesions 826–7
conjunctival naevi
progression to melanoma 826–7

conjunctival naevi – *continued*
  screening 827
conjunctival papilloma 826
conjunctival squamous cell carcinoma
  human immunodeficiency virus (HIV) 343
  human papillomavirus (HPV) 337, 343
conjunctival tumours 826–8
  clinical management 837–8
connective tissue 3
  development 733
  eye/ocular adnexa 825
connexins 9, 1031
  cytotoxic ‘bystander effect’ 1069–70, 1535
connexons 1031–2
Conn’s syndrome 603, 608
contact inhibition 1031
  loss in tumour cells 25–6
cooked food-related carcinogens 377–8
  benzo[*a*]pyrene 392
  heterocyclic amines 392
  PhIP 281, 298
cord blood *ex vivo* expansion 1549–50
cord blood transplantation 1549–50
  advantages 1549
  disadvantages 1549–50
  graft versus host disease 1549
corepressors
  oestrogen receptor interactions 140–1
    antiandrogens mechanism of action 1454
  steroid receptor interactions 138, 140
cornea 825–6
corneal micropocket assay 242, 975, 980
corticosteroids
  chemoembolization ancillary medication 1255–6
  chemoprevention 1426
  side effects 1151–2
corticotroph adenoma 600
Cortinellus shiitake (shiitake) 391
cortisol 1048
cosmic radiation 309
co-stimulatory molecules 1060, 1389–90
  dendritic cell expression 1391
  DNA vaccine response 1398–9
Coulter Counter 884
Cowden syndrome 636, 806, 827
  pathology 820
  PTEN knockout mouse studies 907–8
CP358774 (OSI-774) 1500
CPT-11 *see* irinotecan sulphate
CR2 (CD21) 55
CRADD 123, 154
craniopharyngioma 792–4, 813–4
Cre/loxP gene targeting 894–5, 905, 910
Crk 166–8, 174, 211, 1487
Crohn’s disease 550
croton oil 287, 1086–7, 1116
cruciferous vegetables 424, 438
cryoprotectants 882–3
cryptorchidism 683–4
Cryptothethya crypta 1278
crystals of Reinke 682–3
CSA (*CSA*) 92–3
CSB (*CSB*) 92–3
Csk 173
CtBP 204
CTLA4/CD28 1398–9
CTLA-4 blocking antibodies 1384
culture dishes 878
curie (Ci) 308
Cushing’s disease 600
Cushing’s syndrome 515, 608, 648
cutaneous CD30-positive lymphoproliferative disorders 711
cutaneous horn 480
cutaneous T cell lymphoma *see* mycosis fungoides
CXC chemokines 239
CXCR-1 239
CXCR-2 239, 1088–9
CXCR-3 239
CXCR-4 157, 239
cycad nuts 273–5, 390–1
cycasin 273–5, 390–1
cyclin A 15, 17, 19, 1512–13
  oestrogen receptor phosphorylation 141–2
  TGF-β-induced growth arrest 180–1
cyclin A kinase 17
cyclin B 15, 17, 1324, 1502–3
cyclin B1 17, 1324–5
  mitosis initiation/exit 1324–5
cyclin D 15, 17
  breast cancer 612
  TGF-β-induced growth arrest 180–1
cyclin D 1513
cyclin D1 19, 215, 395, 1512, 1526
  antisense oligonucleotide targeting 1443
  degradation for G_1 arrest 1508
  laryngeal squamous cell carcinoma 508
  mantle cell lymphoma 698
  see also p16cyclinD1pRb pathway
cyclin D1
  head and neck cancer 495
  laryngeal cancer 542
  oesophageal cancer 541–2
  regulation 1514–15
  Wnt target genes 203–4, 206–7
cyclin D2 1512
cyclin D3 1512
cyclin D kinase inhibitors 1513
cyclin E 15, 17, 19, 1408, 1512–14
  restriction point control 18
  TGF-β-induced growth arrest 180–1
cyclin E kinase regulation 1513, 1516–17
cyclin G 896
cyclin-activating kinase (CAK) 15, 1512–13
cyclin-dependent kinase inhibitors (CKIs) 15, 1364, 1513
  TGF-β-induced growth arrest 180–1
  therapeutic approaches 21, 1499–500, 1517
cyclin-dependent kinases (cdks) 14, 953–5, 1363, 1512, 1526
  cell cycle regulation 15, 1324, 1512–14
    restriction point control 1514
  genes 1513
  hepatocellular carcinoma 571
  HPV E7 protein interaction 53
  structure 1512
  subunits 1512
    phase-specific expression 1513
    regulation by phosphorylation 1512–13
  TGF-β-induced growth arrest 180–1
cyclins 14, 15, 953–5, 1363, 1526
  HPV E7 protein interaction 53
  phase-specific expression 1513
  phase-specific proteolysis 1513
  phosphorylation/dephosphorylation 1512–13
  TGF-β-induced growth arrest 180–1
  ubiquitination 1513

cyclobutane pyrimidine dimer formation 89, 310
  repair 280
cyclooxygenase inhibitors 432, 440, 443, 1426
cyclooxygenase 2 (COX-2) 443
cyclooxygenase 2 (COX-2) inhibitors 443, 552, 591, 1366, 1424–5, 1432
cyclophosphamide 301, 414–15, 756, 1296
  combined drug therapy 1489
    etoposide 1052
    interleukin 2 (IL-2) 1376
  in vivo metastasis models 1005–6
  metabolic activation 1296
  resistance 1355
  side effects
    bladder cancer 655
    osteoporosis 1151–2
  structure 1296
cyclosporin 414–15, 716, 971, 1552
  chronic graft versus host disease 1548
  P-glycoprotein interaction 1048, 1338
CYP polymorphisms, DNA adduct levels 304
CYP1A1
  polymorphism 304
  tobacco smoke induction 408–9
CYP1A2 polymorphism 549
CYP2D6 polymorphism 382
CYP3A4 673
cyproterone acetate 568
cyrosurgical ablation 1265
cystadenofibroma 637
cystectomy 660
cysteine 422
cysteinylglycine 422
cystic duct 563
cystic duct tumours 577
cystography 656
cystoprostatectomy 660
cystosarcoma phyllodes 614–15
cystoscopy, bladder cancer
  follow-up 660
  screening 656
cysts
  computed tomography (CT) 1161
  ultrasound imaging 1187
cytarabine 1552–3
  mechanism of action 1278–9
  metabolism 1278
  resistane mechanism 1279
  structure 1277
  transport 1278
cytidine analogues 1277–9
cytochrome *c* 128–31
  Apaf-1 activation 126–7
  mitochondrial caspase activation 126–7
  necrosis (caspase-independent death)-related release 130
cytochrome P450 422, 1410
  1A2–acetaminophen (paracetamol) prodrug activation system 1539–40
  4B1–2-aminoanthracene prodrug activation system 1539
  4B1–4-ipomeanol prodrug activation system 1539
  chemopreventive approaches 427
  inducers 428
  inhibitors 427–8, 437–8
  polymorphism 380–1
    colorectal cancer susceptibility 549
  tobacco smoke carcinogens activation 408
cytofectin 1443
cytogenetic alterations *see* chromosomal abnormalities
cytokeratins 467, 1081–2
  cholangiocarcinoma 575
  chordoma 811–12
  desmoplastic small round tumour 759
  embryonal carcinoma 813
  endolymphatic sac papillary tumour 848
  ependymal tumours 805
  epithelioid sarcoma 758–9
  hepatocellular carcinoma 570–1
  leiomyosarcoma 751–2
  micrometastases detection 474–5
  middle ear adenoma 845
  primitive neuroectodermal tumours 798–9
  renal cell carcinoma 652
  seminal vesicle carcinoma 678
  soft tissue tumours 745
  synovial sarcoma 758
  thyroid papillary carcinoma 605
  urothelium/urothelial neoplasms 657
cytokine therapy 1371
  biological basis 1371–2
  clinical aspects 1372
  immunotherapy 1551–2
cytokines
  angiogenesis 239, 971, 1028–9
  antimicrotubule agent high concentration effects 1327
  cell growth 25
  cell migration 231
  cellular vaccine production 1391
  central nervous system tumour secretion 1124
  early tumour growth 938
  gene therapy 1074–5
    delivery method 1074–5
    dendritic cell targets 1075
  haematopoiesis 6–7, 1102
  radiotherapy responses 1365
  signalling 147–57
  squamous cell carcinoma progression 1088–9
  TGF-β regulation of expression 181
cytopathology 461–2
cytosine deaminase 1386–7
  suicide gene activity 1069, 1536–7, 1541
cytoskeleton 1025–6
  adherens junctions communication 215
  alterations
    invasion/metastasis 231
    transformed cells 963–4
  extracellular matrix connections 214–15
  focal adhesions communication 215
  integrin connections 214, 243, 985
cytotoxic agents
  DNA cross-link formation 293–4
  drug development targets 1042
  occupational exposure 414–15
cytotoxic T cells 1389–90
  cytokine gene therapy targets 1074
  DNA vaccine response 1397–8
  gene-modified dendritic cell antitumour response 1075–6
  tumour cell responses 124, 1390
CYVADIC chemotherapy 762–3

D

DAB389–IL-2 fusion toxin 1386
dacarbazine 301
dactinomycin *see* actinomycin D
danazol 568
DAPH2 1491

DARC 239
daunomycin 1316
daunorubicin 1333–4
  cardiotoxicity 1336
  clinical activity 1334–5
  metabolism 1335–6
  pharmacology 1335
  resistance 1353–4
  structure 1333–4
  transport 1337–8
DCC (deleted in colon cancer) 29, 462, 955, 1090, 1525
  central nervous system tumours 1123–4
  endocrine tumours 607
  loss during colon cancer progression 32–33
  oesophageal adenocarcinoma 542
DD3 673
DDT 414
death domains (DDs) 119–20, 123–4, 153
death effector domains (DEDs) 119–20, 123–5
dedifferentiated chondrosarcoma 726
dedifferentiated chordoma 728
dedifferentiated liposarcoma 750–1
deep fibromatoses (desmoid tumours) 734
delayed type hypersensitivity, HPV lesions (warts) 55
deleted in colon cancer *see* DCC
deleted in pancreas cancer *see* DPC4
δ cells 580
dendritic cells
  antigen presentation 1389–91
  cancer-related defective function 1391–2
  cellular vaccines 1378, 1391–2
    tumour-specific peptides 1392
  follicular 689–90
  gene therapy targets 1075–6
    CD40L gene transfer 1076
    cytokine genes 1075
    cytotoxic T cell activation 1075–6
  in vitro production for immunotherapy 1391–2
  lymph nodes 690
  skin 477
Denys–Drash syndrome 647
5-deoxy-5-fluorouridine (doxifluridine) 1277
depreotide 1225
dermatofibrosarcoma protuberans 748
  immunohistochemistry 745
dermis 477
dermoid cyst
  orbit 826, 835
  ovary 635, 640
desmin 467, 966
  angiomatoid fibrous histiocytoma 748–9
  desmoplastic small round tumour 759
  leiomyosarcoma 751–2
  primitive neuroectodermal tumour 798–9
  rhabdomyosarcoma 752
  soft tissue tumours 745
  squamous cell carcinoma 481
desmoid tumours (deep fibromatoses) 734, 746
desmoplakins 9
desmoplastic cerebral astrocytoma *see* desmoplastic infantile ganglioglioma/astrocytoma
desmoplastic infantile ganglioglioma/astrocytoma 802–3, 807
desmoplastic medulloblastoma 798
desmoplastic small round tumour 759
  chromosomal translocations 759
  EWS fusions 738–9, 759
  immunohistochemistry 745, 759
desmosomes 9, 983, 1010, 1081–2
development 964
  biliary system 563
  bladder 645
  bone 721–2
  breast 9, 611–12, 1026–8
  central nervous system 789
  colon 545
  connective tissue 733
  ear 841
  eye 825–6
  female reproductive tract 617
  gall bladder 563
  gastrointestinal tract 545
  heart 777
  kidney 645
  liver 563–4
  mesenchymal–epithelial interactions 1026–8
  nervous system 789–90
  ocular adnexa 825–6
  oesophagus 523
  oral cavity 491
  pancreas 579
  penis 675
  peripheral nerve tissue 733
  prostate 665
  salivary glands 491
  seminal vesicle 677
  skin 477
  small intestine 545
  stomach 524
  testis 678–9
  ureter 645
  urethra 645
  urinary tract 645
  vascular tissue 733
dexamethasone 1048
DFMO *see* difluoromethylornithine
DHFR (dihydrofolate reductase)
  alterations in methotrexate resistance 1290–1
  gene transfer into haematopoietic stem cells 1076–7
diabetes mellitus 589, 632
diagnosis of cancer 461–75
  histological examination 463–5
  molecular techniques 473–4
  pathology reports 465
  prognostic markers 472–3
  routine procedures 463–73
  tumour classification 465–71
  tumour grading 471
  tumour markers 472–3
  tumour staging 471–2
  see also screening
diallyl sulfide 410, 427–8
dianilinophthalimide (DAPH) tyrosine kinase inhibitors 1491
diaphysis 721–2
dibenzacridine 405
dibenzanthracene 273
  breast cancer models 925, 933–4
  occupational exposure 415, 512
7H-dibenzo[*c,g*]cabazole 405
dichloromethane 414
Dickkopf (*Dkk*) 199
Dictyostelium discoideum 964
diesel engine exhaust 300, 414–15, 418
dietary factors 33, 389–96, 421, 794, 1081, 1417, 1482
  additives/preservatives 389

anticarcinogenic compounds 389, 423–4
cancer prevention 421, 423–4, 449–50
cholangiocarcinoma 573–4
colorectal cancer 265, 549
cooking products 389
DNA damage/adduct formation 89, 379, 392, 395
edible plant components 389
epidemiological studies 263–4
fat intake 265, 395–6
gastric cancer 264–5, 395–6, 526–31, 1432
genome hypomethylation 329
macrocomponents 394–6
microcomponents/contaminants 389–4
mycotoxins 389–90
nasopharyngeal carcinoma 340
occupational exposure 414–15
oesophageal cancer 525–6, 528
pancreatic cancer 589
prostate cancer 33, 395, 672–3
recommendations 396
sodium chloride 395–6, 529–1
storage/fermentation compounds 389
total calories 394–6
xenograft studies 920
dietary fibre 549, 1433
diethylstilboestrol 1140–1
cervical clear cell adenocarcinoma association 617–20
differential display techniques 900
differentiation 4–5, 225–8, 984, 1084, 1403
cell cycle G0 stage 953–5
diagnostic tumour markers 966
enhancement by retinoids 1409
genetic models 964
haematological malignancy pathogenesis 1101, 1106–7
in vitro models 1110
haematopoietic cells 1102–4
homeobox gene regulation 10
in vitro models 964
inducing agents 1403–12
malignancy relationship 959, 1101
oncogene activity 1106–7
pregnancy-related mammary gland changes 216–17
proliferation inverse correlation 959
regulation 5, 959
at cell surface 959–60
basement membrane proteins 213
gene expression 960, 1403
soft tissue neoplasms 733–4
tumour 953–68
see also terminal differentiation
differentiation therapy 22–3, 966–7, 1403–15
clinical aspects 1412–15
rationale 1403–4
differentiation-inducing agents 1042, 1403–12
combinations of agents 1411–12
cytokines 1410–11
histone deacetylase inhibitors 1404–7
peroxisome proliferator–activator receptor ligands 1407–8
planar–polar solvents 1408
retinoids 1409–10
vitamin D_3 1408–9
diffuse idiopathic pulmonary neuroendocrine hyperplasia (DIPNECH) 510, 514
diffuse large B cell lymphoma 702–4
anaplastic variant 703
centroblastic variant 702
clinical features 702
extranodal 704
genetic features 703–4
t(14;18)(q32;q21) translocation 692
immunoblastic variant 703
immunophenotype 703
intravascular (angiotropic) 704
molecular classification 470–1
primary effusion 704
primary mediastinal (thymic) 704
subtypes 704
T cell/histiocyte-rich variant 703
difluoromethylornithine (DFMO) 440, 448–50, 1425, 1432–3, 1435
digital rectal examination 666, 673
dihydrofolate reductase 1290
see also DHFR
diltiazem 1048
dimethyl sulphoxide (DMSO)
cell freezing solutions 882–3
retinoid combined treatment 1411
4-dimethylaminoazobenzene 273
dimethylbenz[*a*]anthracene (DMBA) 273
mouse melanoma models 1116–17
neurocarcinogenesis 1129
rodent breast cancer models 925-7, 929–31, 933–4, 1426
dimethylformamide 1411
dimethylnitrosamine 285–6
dioxins 737, 783
dietary 393–4
diphtheria toxin–VEGF construct 1466
disc angiogenesis system 976
Disheveled
Axin binding 200–1, 203
cytoskeleton association 201–2
DEP domain 201
DIX domain 201
localization within cell 201–2
PDZ domain 201
Wnt signalling (destruction complex antagonism) 196, 199, 201–3
disintegrins 1468–9, 1477
dissemination processes 225, 230
diterpene esters 1086–7
dithiolthiones 424
divisional clock 107–9
DLG 215–16
DLK1 327–8
dll (discs large lethal) 215–16
DNA adducts 89, 279, 293–304
adenine 296–7, 378, 407–8
alkylating agent mode of action 279, 295, 298, 1295–300, 1302, 1341
base 293–4
biological activity 299–300
position/nature influences 279, 297
carcinogenesis/cancer risk 279, 293, 299, 384
exposure biomarker 378–9, 438
internal dose markers 377
risk assessment 303
chemopreventive strategies 438
conversion to mutations 279, 281, 297–9
cytosine 90, 295–7
detection in human tissues 300–1
genotype correlations 304
detection methods 298–9
^{32}P postlabelling 298, 378–9
immunoassay 298–9
mass spectrometry 299
radiolabelled compounds 298

DNA adducts – *continued*
detection in urine 301
dietary factors 379
heterocyclic amines 392
polycyclic aromatic hydrocarbons 392
polyunsaturated fatty acid (PUFA) peroxidation products 395
dosimetry 300
endogenous 302–3
formation 293–4
chemical carcinogens 279, 295–8
DNA reaction sites 296–7
guanine 90, 279, 295–7, 302, 378, 407–8, 1295–6
nucleoside 293–4
nucleotide 293–4
platinum-based drug mechanism of action 1304–7
repair processes 280
DNA alkyltransferase 90
nucleotide excision repair 91
thymine 295
tobacco exposure relationship 299–302, 408–9
DNA alkyltransferase 90, 280
DNA bases nomenclature 1271–2
DNA damage 87, 1497–8
apoptotic response 19, 280, 1507–9
carcinogenesis initiation 286–7
cell cycle checkpoints activation 17–18, 1507, 1515–17
cancer cell defects 1507
spindle assembly checkpoint 1517
therapeutic applications 1518
chemical carcinogens 278–80, 288, 293–5, 925, 933
base hydroxylation 278–9
direct DNA binding 279
chemotherapeutic approaches 21
detection techniques 1510
dietary polyunsaturated fatty acid (PUFA) peroxidation products 395
dosimetry for cellular profiling 1511–12
double-stranded breaks 96–7, 293–4
induced 1507–8
insertion/deletion loops 94
ionizing radiation/radiotherapy 310, 1360–2, 1367, 1508
cellular radiosensitivity relationship 1362–4
microsatellite instability 94
oxidative stress 89
reactive nitrogen oxide species 395
spontaneous 422, 1507–8
sunlight exposure photoproducts 91, 280, 481–2
cyclobutane–pyrimidine dimers 91, 280
pyrimidine–6, 4–pyrimidine dimers 91
ultraviolet radiation 89, 310, 1508
see also DNA adducts; genomic instability
DNA ligase I 91
DNA ligase III 94
DNA ligases 1522–3
DNA low-fidelity polymerase 1509
DNA manipulation techniques 1522–3
DNA methylation (cytosine methylation) 248, 318–20
carcinogenetic mechanisms 328
chemopreventive approaches 432, 442
dietary deficiency of methyl donors 329
gene silencing 96, 319–20
hMLH1 in colorectal cancer 96, 324, 556–7
tumour-suppressor genes 322
genome mapping 329
genomic imprinting 325–6
heritability during cell replication 319
molecular modulators 320–2
mutagenic susceptibility targets 322, 379
pancreatic cancer screening 591
tumour cells 322
promoter region hypermethylation 322–5
tumour-suppressor gene identification 31
DNA methyltransferases (DNMTs) 318–19
therapeutic strategies 330–1
DNA packaging 317
DNA polymerase δ 1509–10
nucleotide excision repair 91
DNA polymerase ϵ 1509–10
nucleotide excision repair 91
DNA polymerases 17, 103, 1522–23
DNA repair 89–104, 280, 1507–10
base excision repair 280, 310, 1361, 1509–10
base-free (abasic) sites 90–1, 280
long-patch repair 91
cancer cell defects 1507
drug resistance 1046
cancer preventive strategies 428, 438–9
cancer therapy 1510–12
cellular profiling 1511–12
radiotherapy 1361–2
specific cotherapeutic inhibitors 1512
transgenic enhancement 1512
DNA alkyltransferase 90
double-stranded breaks 96–100, 310
homologous recombination 97–100, 1362
non-homologous end joining 97
radiotherapy-induced 1361–2
single-strand annealing 100
enzyme targets of antimetabolites 1277
functional tests 381–3
gene defects 11, 87–8, 1508
cancer susceptibility markers 382–3
frame-shift mutations 94
skin cancer susceptibility 478
genotoxicity test system in rat liver 355–7
Hus1:Rad1:Rad9 sliding clamp model 103
mechanisms 89–90, 280, 1508–10
major pathways 1509
mismatch repair *see* mismatch repair
mitochondrial DNA 94
multimolecular assemblies/nuclear foci 99–101, 103–4
nucleotide excision repair *see* nucleotide excision repair (NER)
one-step repair (OSR) 1509
p53 mutation-related defect 29
photodimers 91, 280
polymerases 103
Rad17 clamp-loader model 103
RecQ-family helicases 89–103
single-base damage 90
transcription-coupled *see* transcription-coupled repair (TCR)
WRN protein 102–3
DNA synthesis 13–14
checkpoint 18, 102
chemotherapeutic approaches 20–1
enzyme targets of antimetabolites 1277
DNA methylation 319
genotoxicity test system in rat liver 355–7
licensing 17
PET markers 1231
DNA topoisomerases *see* topoisomerases
DNA transfection 962–3, 965
DNA tumour viruses 49–71, 83, 965
haematological malignancies 1104–5
oncogenesis 27
DNA vaccines 1059–60, 1389, 1394–9

adenovirus vectors 1395
advantages 1398–9
antigen processing pathways 1397
current clinical trials 1399
delivery 1397
gene gunning into skin 1397
genes for whole proteins 1395
immune response induction 1397
antibody response 1397
co-stimulatory molecule response 1398–9
cytotoxic T cell response 1397–8
helper T cell response 1398
intramuscular/intradermal injection 1397
minigenes for single epitopes 1395–7
plasmid vectors 1394
regulatory elements 1394–5
retrovirus vectors 1395
safety 1398–9
DNA-dependent protein kinase (DNA-PK) 1361–2, 1515–16
DNMT1 318–19, 321
DNMT3a 318–19
DNMT3b 318–19
docahexaenoic acid 395
docetaxel 1329
combined drug treatment 1327–8
vinorelbine 1331
mechanism of action 1323, 1330–1
docking proteins
receptor tyrosine kinase signalling 166–8
T cell receptor activation 172–3
Dolabella auricularia 1329
dolastatin 10 1329
dolastatin 15 1329
dolastatins 1329
mechanism of action 1323, 1329
donor lymphocytes infusion 1550–1
Doppler ultrasound 1189
hepatocellular carcinoma 1196
ovarian tumour angiogenesis assessment 1194–5
renal tumours 1197
dormancy (quiescence) 13, 235, 938, 1036–7
double contrast radiographs 1149
Down syndrome 573
doxifluridine (5-deoxy-5-fluorouridine) 1277
doxorubicin 660, 756, 762–4, 1333–4
clinical activity 1334–5
combination therapy 1489, 1491
ethacrynic acid potentiation 1355–6
integrin conjugates 1469, 1481
liver tumour chemoembolization 1256
mechanism of action 1315–17
antitumour compound screening 1043
metabolism 1335–6
neuroendocrine tumours 609
pharmacology 1335
resistance 1353–6
HEp2 cell line 1351
side effects 1151–2
cardiotoxicity 1336
structure 1333–4
transport 1337–8
doxycycline 1480
DP1 99–100
DPC4 (deleted in pancreas cancer) 30, 188, 547
pancreatic cancer 30, 592–3
pancreatic intraepithelial neoplasia (PanIN) 581–2
see also Smad4
DR3 (weasle; tweak) 123–6
DR4 (trail receptor-1; Apo2) 123–6
DR5 (trail receptor-2) 123–6, 131–2
DR6 123, 124–6
Drosophila
basement membrane proteins 210
tissue polarization regulation 215–16
Wnt target genes 203–4
Drosophila melanogaster 964
Wnt signalling 196
drug combinations
specific DNA repair inhibitors 1512
see also antagonistic drug combinations; synergistic drug combinations
drug delivery, tumour blood flow 237
drug development 21, 1041–6
costs 1043
genetic approaches 1045–6
expression profiling technologies 1045–6
positional cloning 1045–6
lead candidate compounds 414, 1041, 1043
molecular strategies 1042
role of Food and Drug Administration (FDA) 1041–2, 1570
screening *see* screening antitumour compounds
stages of evaluation 1041–2
success rate 1043
see also chemopreventive agent development; clinical drug trials
drug efficacy 1041
cell culture assays 886
clinical trials 1042
drug combinations 1051
NCI Cancer Drug Screen procedures 1043
drug resistance 1046–53, 1347–56
antimicrotubule agents 1328
antisense oligonucleotide targeting 1443
apoptosis defects 119, 132
DNA mismatch repair systems 95, 1510
gene therapy 1076–7
mechanisms 1046
ribozyme targeting 1445
tumour cells 1044
see also multidrug resistance (MDR)
drug selectivity 1041
drug combinations 1051–2
drug toxicity
clinical trials 1042
phase I 1570
DNA damage 1508
drug combinations 1051
IND (investigational new drug) applications 1041–2
toxological profile 1041
drug transport systems
anthracyclines 1337–8
antimetabolites 1272
cytarabine 1278
daunorubicin 1337–8
drug resistance 1046
5-fluoro-2′-deoxyuridine (floxuridine) 1273, 1275
5-fluorouracil 1273
gemcitabine 1279
melphalan 1298
methotrexate 1288–9
mitoxantrone 1045
taxanes 1328
Vinca alkaloids 1328
DT diaphorase 1386–7
antitumour compound molecular targeting 1044

Dubin–Johnson syndrome 1352
duct ectasia 1242
duct of Santorini 579
duct of Wirsung 579
ductal carcinoma in situ (DCIS) 613–14, 1449–50
  comedo DCIS 613–14
  non-comedo DCIS 614
ductography *see* galactography
Dulbecco modified Minimal Essential Medium (DMEM) 880–1
Dulbecco's phosphate-buffered saline 880
Dulbecco's phosphate-buffered saline A 880
Duncan's syndrome (X-linked lymphoproliferative disorder) 59–60
duodenum 545
Dupuytren disease (palmar fibromatosis) 746
DX8951f 1319
dysembryoplastic neuroepithelial tumour 806–7
dysgerminoma 640
dysplasia 249–50, 1085, 1418–19
  clinical management approaches 1418–19
  colorectal adenoma 552
  colorectal lesions of inflammatory bowel disease 554
  larynx 508
  oesophageal cancer progression 523, 533
  surrogate endpoint biomarker (SEB) 1421–23
dysplastic naevus (BK mole) 479–80, 487
  clinical management 1435
  malignant change 479–80, 485
  pathology 479
dystroglycan 214–15
dystrophin 214–15

E

E1AF/EWS fusion 738
E2A 1107
E2F
  apoptosis induction 1516
  Rb protein complexes/cell cycle regulation *see* Rb–E2F pathway
  therapeutic approaches 21, 1408
E2F1 1516
E12 1107
E47 1107
E7070 1499
E-cadherin 8, 495, 554
  adherens junctions 215
  catenin complex signal transduction 1034
  functional inactivation
    cancer cells 25–6, 1082
    CpG island hypermethylation 324
  gastric cancer 542
  malignant invasion 230
  tumour progression 1082
E-selectin 244–5, 1016–17
  malignant invasion 230
EA hy926 cell line 979
ear
  development 841
  non-neoplastic lesions 842
  structure 841–2
ear tumours 841–9
  classification 842
Earle's balanced salt solution 880
early detection 462–3, 955–6
  brain tumours 794–5
  pancreatic cancer 582, 589
Early Lung Cancer Action Project (ELCAP) 1159
early response markers 377, 379–80
EB 1089 (Seocalcitol) 1414
Eb cells 1057
  tumour vaccine studies 1060
EB1 1517
eccrine sweat glands 477
ectoderm 1081
ectomesodermal syndromes (phakomatoses) 817–20, 836–7
ECV304 cell line 979–80
eicosapentaenoic acid (EPA) 395
EKB 1869 1500
elastofibroma 743–4
elderly patients 19
  breast cancer screening 1236
electric fields 307, 308
  activity units 308–9
  biological effects 311
  carcinogenesis studies 313–14
    animal brain tumour models 1128
  exposure limits 314
  sources of exposure 309
electron microscopy
  acinar cell carcinoma 583
  cardiac myxoma 784
  Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 799
  hepatocellular carcinoma 570
  jugulotympanic paraganglioma 846
  meningioma 810
  mesothelioma 771
  neuroendocrine tumours 602
  renal cell carcinoma 652
  soft tissue tumours 746
  tumour classification 466
electrophile scavengers 428
elF4E 495
ellagic acid 424, 427–8
embolic agents 1256
embryogenesis 1081
  DNA methylation patterns 318–19
  mesenchymal–epithelial interactions 1026–8, 1083–4
  Wnt signalling 195–6
  see also development
embryonal carcinoma 465–6
  central nervous system 791, 813
  developmental field effects 1403
  immunohistochemistry 683
  ovary 640
  testis 681–2, 1192–3
embryonal rhabdomyosarcoma 752
  genetic features 753
  middle ear 849
  penis 676
  prostate 671–2
embryonal sarcoma 576
embryonic stem cells 4–5, 891–2, 1081
  chromosome engineering 898–9
  experimental manipulation/culture 903–4
  injection into blastocysts 891, 894
  knock-in mouse generation 905
  knockout mouse generation 903–5
  transgene introduction by homologous recombination 892–5
EMD121974 1477
EMS-1 495
encephalo-trigeminal angiomatosis (Sturge–Weber syndrome) 836–7
enchondroma 722–3
endobronchial stents 1176
endocrine ablation therapy 1449

  breast cancer 1450
  surgical/radiation-induced ovarian ablation 1450, 1455
endocrine cells
  large intestine 547
  stomach 524–5
endocrine organs 599
endocrine regulation 953
  breast tissue differentiation 9
  receptor tyrosine kinases 169
endocrine system 4
endocrine tumours
  aetiology 601
  epidemiology 600–1
  molecular genetics 606–8
  pathology 599–600, 602–4
  prognostic factors 608
  screening 601–2
endoderm 1081
endoglin 1471
  transforming growth factor-β (TGF-β) binding 182
endoluminal stents 1176
endolymphatic sac papillary tumour 847–8
endometrial biopsy 632–3, 1193
endometrial cancer
  aetiology/risk factors 630–2, 1139–40
    oestrogen exposure relationship 630, 632, 1139
  clear cell carcinoma 629
  clinical management 634
  DNA mismatch repair defects 96, 550
  endometrioid adenocarcinoma 629
  epidemiology 630–2, 1139
  FIGO staging 633–4
  genetic factors 1139
  human papillomavirus (HPV) 633
  in vivo animal model 1140–2
    oestrogen induction 1140–2
    problems 1141
  magnetic resonance imaging (MRI) 1208
  molecular genetics 633
  papillary serous carcinoma 629
  pathology 629, 633
  prognostic factors 633–4
  PTEN mutations 907
  screening 632–3
  type I carcinomas 629, 1139
  type II carcinomas 629, 1139
  ultrasound 1193
    saline infusion sonohystography 1193–4
  vimentin expression 469
endometrial hyperplasia 628–9
endometrial polyp 627–8
endometrial stromal sarcoma 629–30
endometrioid carcinoma/adenocarcinoma 250, 629
  ovary 637, 639
endometrium 617
  oestrogen response 627
  progesterone response 627
endoscopic retrograde cholangiopancreatography (ERCP) 590, 594
endoscopic surgery 1561–2
  transanal 1561–2
  upper gastrointestinal malignancies 1562
endoscopic ultrasound
  islet cell tumours 595
  laparoscopic staging 1559
  pancreatic cancer 589–90, 594
endostatin 241–2, 971, 1469
  angiogenesis inhibition by gene therapy 1073
  antineoplastic/antiangiogenic activity 1476
  mechanism of action 1469
endothelial cells 236, 1009, 1025
  adhesion molecules 1014–17
    adhesion assays 1016–17
    identification techniques 1015
    organ distribution 1016
  angiogenesis 237
    mathematical modelling 945–7
  antiangiogenesis strategies 1465
  cell culture 1013–17
    activation 1014
    isolation 1014
  chemokine receptors 232
  coagulation system interactions 1019
  degradative enzyme release 1018–19
  extracellular matrix–adhesive interactions 939–40
  heterogeneity 1019–20
  immortalized cell lines 979–80
  matrix metalloproteinase (MMP) secretion 241
  migration 237, 242–3, 938–9
  monolayer wounding model 1018
  motility factors 987, 1011–12
    in vitro studies 1017–18
  proliferation 243, 939
  therapeutic strategies (vascular targeting) 1465, 1471
  transferrin production 1012
  tumour cell interactions 1009–21
    antibody binding techniques 1020
    cell adhesion 984–6
    growth inhibitory effects 1012–13
    immunocytochemistry 1020
    model establishment 1013–19
    model limitations 1020–1
  tumour-associated 1465
endothelium 236, 1010
  barrier function 1019
  basement membrane 1010–12
  continuous 236, 1010
  discontinuous 236
  fenestrated 236
  intercellular communication 1010
  permeability regulation 1011
  regional variations 1010–11
  transcapillary transport regulation 1011
  tumour blood vessels 236, 1465
endovascular stents 1176
energy restriction 395
engine exhaust 300, 377–8, 414–5, 418, 512
eniluracil 1277
Enneking sarcoma staging system 742–3
ENT1 1272, 1278–9
ENT2 1272, 1279
entactin 236
Enteral Wallstent 1562
enterocytes 7
enteroendocrine cells 7
enteropathy-type T cell lymphoma 712
env 76
environmental carcinogens 33
  occupational carcinogens differentiation 414–15
environmental polution 300
eosinophil colony-stimulating factor (CSF) 1103
eosinophilic granuloma *see* Langerhans cell histiocytosis
eosinophils 1056
  bone marrow precursor cells 6
  haematopoiesis 1103

ependymal cells 789–90
ependymoma 792, 794, 820, 1123
  anaplastic 805
  myxopapillary 805
  papillary 804–5
  pathology 804–5
  xenografts 1127
Eph 171
Ephrin 171
epidemiological studies 259–71
  aetiological hypothesis formulation 260, 265
  analytical 260–2, 266
  bias 262, 269
  cancer causation inference 265–71, 382
    criteria for causality 269–70, 382
    individual exposure 270–1
  chance variation 268
  confounding factors 260, 262, 268–9
  data quality 269
  descriptive 260, 262–5
    personal characteristics 262–3
    place/time characteristics 264–5
    supporting experimental evidence 265
  ecological 266
  sample size 381–2
  statistical associations 260
    measures of strength 260
  see also molecular epidemiology
epidermal growth factor (EGF) 14, 25, 395, 1025, 1034
  angiogenesis 239, 1490
  epidermal growth factor receptor (EGFR) interaction 1486–7
  gap junction-mediated intercellular communication 1032
  invasion/metastasis 232, 1034–6
  oesophageal adenocarcinoma 542
  small intestinal epithelium proliferation stimulation 8
  v-*erbB* oncogene 955
epidermal growth factor (EGF) family 8
epidermal growth factor receptor (EGFR) 161, 170–1, 176, 495, 1030–1, 1087, 1486–7
  amplification
    bladder cancer 1488
    pancreatic cancer 1488
    PTEN mutations 30
    tumours/tumour cell clones 1488
  antisense oligonucleotide targeting 1443
  bispecific antibodies 1492
  C225 antibody target 1383, 1489
  central nervous system tumours 1123–4
    transgenic mouse models 1131, 1133
  dominant-negative mutants 1492–3
  inhibitors 1427–31, 1486, 1500
  intracellular signalling 1487
  knockout mouse studies 1486–7
  ligand interactions 1486–7
    heterodimerization 1487
  ligands 1486–7
  oesophageal adenocarcinoma 542
  single-chain antibodies 1492
epidermal proliferative units 8
epidermis 8–9, 477
  extracellular matrix 9
  layers 477
epidermodysplasia verruciformis 54, 337
epididymis 679
epigallocatechin 436, 1482
epigenetic genome changes 317–31
  cancer diagnosis 329
  carcinogenesis 278, 328–9
  cytosine methylation 318–20
  gene regulation 317–20
  histone acetylation 317–18
  molecular modulators 320–2
  therapeutic approaches 330–1
  tumour cells 322–5
epiglottis 503–4
epiphysis 721–2
epiploic appendages 545
epirubicin 1333
  mechanism of action 1338
  resistance 1353
epithelial cancers 466, 1081–6
  models 1086
epithelial membrane antigen (EMA)
  anaplastic large cell lymphoma 710
  angiosarcoma 729
  atypical teratoid/rhabdoid tumour 800
  bladder cancer 657
  brain tumours 795
  chordoma 727–8, 811–2
  desmoplastic small round tumour 759
  endolymphatic sac papillary tumour 848
  ependymal tumours 805
  epithelioid sarcoma 758–9
  leiomyosarcoma 751–2
  meningioma 810, 847
  nodular lymphocyte predominance Hodgkin's lymphoma 714
  renal cell carcinoma 652
  soft tissue tumours 745
  synovial sarcoma 758
  testicular germ cell tumours 683
epithelial tissue 3, 1081–2
  cell junctions 9
  large intestine 545–7
  respiratory tract lining 503–4
  small intestine 7
epithelial to mesenchymal transdifferentiation
  metastasis 190, 1083
  soft tissue neoplasms 733–4
epithelioid haemangioendothelioma 519, 754
  central nervous system 810–11
  immunohistochemistry 745, 754
  penis 676
epithelioid haemangioma 576
epithelioid sarcoma 469, 758–9
  immunohistochemistry 745, 758–9
  penis 676
epothilone A 1331
epothilone B 1331
epothilones 21
Epstein–Barr virus 49–50, 55–62, 75, 247, 339–41, 1081, 1108–9
  AIDS-related B-cell lymphomas 1104–5
    primary CNS lymphomas 704
  associated carcinomas 60–1, 339
  B lymphocyte growth stimulation/immortalization 55–6, 58–9, 339
  BARF0 (complementary strand transcripts; CSTs) 59
  BARF1 59
  BCRF1 59
  BHRF1 59
  Burkitt's lymphoma 55, 59–62, 335, 339–40, 1104–5
  classification 55–6
  cytotoxic T lymphocyte epitopes 61–2
  EBV1/EBV2 polymorphism 56
  EBV nuclear antigens (EBNAs) 56–8, 61–2

EBV-encoded RNAs (EBERs) 56–8
epidemiology 55
epithelial cell immortalization 59
gene nomenclature 59
genomic organization 56, 56–9, 339
immediate early (lytically related) genes 56–8
latent genes 56–8
lytic genes 56–8
historical aspects 55
Hodgkin's disease 339–40, 716
immune response/immune tolerance 61–2
immunosuppression-related lymphoma 339–40, 784
in vitro B lymphocytes transformation 1104–5
infectious mononucleosis 55, 58–62, 339
IR1 (*Bam*H1 W) 56
lymphoepithelioma 60–1, 498
lymphomas 339–41, 343, 718–19
animal models 1111
membrane antigens (LMPs) 56–9
minimal immortalizing region (MIR) 59
nasopharyngeal carcinoma 55, 59–61, 335, 339–40
IgA/viral capsid antigen (VCA) screening 340
pathogenesis 59–61, 339
post-transplant lymphoproliferative disorders 58, 340, 716–17
ribozyme targeting 1445
sarcoma pathogenesis 737
T cell large granular lymphocytic leukaemia 707
transmission 339
vaccine development 56, 61–2, 341
virion structure 56, 339
equilibrative nucleoside transport (ENT) system 1272
ERAP-160 *see* SRC-1
erb B
activation by integrated provirus 76
v-erbB 176, 955
Erb B receptors 171, 1031
ligand-receptor couples 169–70
therapeutic targeting 1486
ErbB1 (erbB1) *see* epidermal growth factor receptor (EGFR)
ErbB2 (erbB2) *see* HER2/neu
ErbB3 1487
ErbB4 1487
ERG-1 translocation 722–3
ERG/EWS fusion 738–9
erionite 767
ERK (Erk) signalling 168, 175, 211
invasion/metastasis 225–8
radiotherapy responses 1366
erythrocytes 1101
bone marrow precursor cells 6, 1103
erythromycin 1048
erythroplakia 493
erythroplasia of Queyrat 676
erythropoietin 6–7, 1103, 1360–1
Esb/Esb-MP cells 1057
metastasis model
gene tagging 1057–8
micrometastases detection 1058–9
vaccination studies 1059–61
Esb/ESb-MP lymphoma model, adoptive cellular immunotherapy (ADI) 1062–3
ESbL lymphoma model, tumour microenvironment studies 1063
escalation with overdose control (EWOC) 1570–1
Escherichia coli
ribonuclease P 1440
strain WP2 mutation assay 353
estragole 391
estramustine 1329
ethacrynic acid 1355–6
ethanol injection *see* alcohol, percutaneous injection
Ethiodol (Lipiodol) embolization 1256–8
ethyl carbamate 407
N-ethyl-*N*-nitrosourea (ENU)
neurocarcinogenesis 1130
rodent breast cancer models 925–6
ethylenimines (aziridines) 1298–9
etoposide (VP-16) 609, 756
combination therapy
cyclophosphamide 1052
interleukin 1α (IL-1α) 1118
p53 gene therapy 1528
mechanism of action 1315–17
resistance 1353, 1355
toxicity 1044–5
etretinate 1431
ETS domain 738
ETV6/*NTRK3* fusion 747
ETV-1 translocation 722–3
ETV-1/*EWS* fusion 738–9
euchromatin 317–8
eustacian tube 841–2
Evi-1 184–5, 190
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 469, 726–7, 737, 760
clinical management 727, 756
extraskeletal 755–6
heart 782
immunohistochemistry 745, 755–6
minimal residual disease (MRD) detection 475
molecular genetics 722–3, 738, 756
EWS fusions 722–3, 738–9, 756
pathology 726–7
EWS fusions/translocation
ATF-1 fusion 757
CHN fusion 757
CHOP fusion 738–9, 750
E1AF fusion 738–9
ERG fusion 738–9
ETV-1 fusion 738–9
Ewing sarcoma/peripheral neuroectodermal tumour (PNET) 722–3, 738–9, 756
FEV fusion 738–9
WT-1 fusion 759
ZSG fusion 738–9
excision assays 919
excision repair 1509–10
see also nucleotide excision repair (NER)
excretory urography 656
exemestane 1457–8
EXO1 94
experimental allergic encephalomyelitis (EAE) 1125–6
expert computer systems 362–3
exponential (Gompertz) growth curve 957
expression profiling technologies 1045–6
EXT-1 722
EXT-2 722
external ear 841
extracellular matrix 5, 209–21, 1025–6, 1084–5
angiogenesis induction 971
antiangiogenic therapy targeting 1468–9
apoptosis regulation 220
bone 722
cell adhesion 230, 939–40, 984
cell adhesion molecules 5, 8

extracellular matrix – *continued*
cell interactions 26, 209
cell migration processes 8, 9, 231
cell proliferation regulation 26, 211–14
cell shape changes induction 213–14, 216–17
cytoskeletal connections 214–15
degradation during invasion 231–2, 939–40, 985
mathematical modelling 947–8
differentiation regulation 209–10, 213, 959–60
growth arrest 213
epidermal basal layer 9
fibroblast interactions 1084
focal adhesion complexes (focal contacts) 213
growth factor receptors 213
growth arrest induction 213–14
integrins 210–14, 242, 1026
cancer 218–20
clustering 213
intercellular communications establishment 215–16
nuclear connections 1025–6
proteins 210
signalling 211–12
cell morphology integration 213–16
complex matrices 213
cytoskeletal responses 214
tissue polarity establishment 215–16
skin carcinogenesis 1088–9
small intestinal epithelium 8
therapeutic strategies 1475–82
functional domains of small molecules 1475–7
heparan sulphate structural mimics 1476–7
integrin antagonists 1477–8
matrix degradation inhibition 1478–81
matrix deposition modulation 1478
rationale 1475
synthetic peptides 1476–7
therapeutic agent selective delivery 1481
tissue organization 216–18
consequences of disruption 218
mammary epithelium model 216–17
transforming growth factor-β (TGF-β) interactions 180–1
tumours 1475
cell adhesive interactions 218–20, 985–6
see also basement membrane
extramedullary haematopoiesis 1102
extraocular muscles 825
extraskeletal myxoid chondrosarcoma 738–9
eye development 825–6
eye tumours 825–39
aetiology 826–27
classification 826
clinical management 837–9
epidemiology 826
pathology 826
prevention/screening 827
eyelid tumours 826–8
clinical management 837–8

F

F11782 1321
FACS analysis 886
factor VIII-related antigen
angiosarcoma 575–6, 729, 755
soft tissue tumours 745
vascular neoplasms 810–11
factor XIIIa 745
Fadd (Mort1) 123, 124, 153
faecal occult blood testing 551
fallopian tubes 617
false morel (*Gyromitra esculenta*) 391
familial adenomatous polyposis (FAP) 29, 36–7, 547, 568
APC mutations 205–6, 325, 556, 1090
chemoprevention 552, 1424–5, 1432
colonic adenomas 36–7
clonality analysis 252
COX-2 inhibitor response 443
morphogenesis 553
colorectal cancer 36–7, 43, 549–50
extraintestinal features 36–7
gastric cancer 542
genetic aspects 36–7, 44
genetic/predictive screening 551–2
prophylactic surgery 551–2
thyroid carcinomas 607
familial breast cancer 19, 29–30, 35, 42–3, 612, 897, 1425
ATM 42
BRCA1 19, 29–30, 42–3, 897
BRCA2 19, 29–30, 42–3
ovarian cancer associations 43–4
PTEN 42
familial cancers 33, 35–44
genetic syndromes 36
genetic testing 44, 462
tumour-suppressor genes 28, 955
familial cholestatic cirrhosis 568
familial mediterranean fever 1329
familial medullary thyroid carcinoma 601, 608
familial osteosarcoma 722
familial renal cell carcinoma 649–50
Fanconi anaemia 18, 87–8, 382
stem cell transplantation 1546–7
FAP-1 124–5
farnesyl transferase inhibition 430, 440, 1426, 1498–9
Fas (CD95)
apoptosis regulation 123, 131–2, 1047–8
mutations/defects 123
tumour immune surveillance avoidance 124
tumour resistance mechanisms 124–5
Fas-ligand (FasL) 152
apoptosis regulation 131–2
knockout mouse studies 1003
radiotherapy responses 1365
tumour immune surveillance 124
FAST1/2 184
fat, dietary intake 395–6, 549, 589, 591, 636, 672
fava beans 392
Fc receptors 1379–80
female reproductive tract development 617
female reptoductive tract tumours 113, 617–42
FEN1 (flap endonuclease) 90–1, 94
fenretinide 440–1
ferritin
haematopoiesis inhibition 1104
immunocytochemistry 570–1
ferulic acid 424, 427
fetal bovine serum (FBS) 881, 887
fetal haemopoiesis 1101–2
FEV/*EWS* fusion 738–9
Feyrter (neuroendocrine) cells 504
FGF-1 *see* acidic fibroblast growth factor
FGF-2 *see* basic fibroblast growth factor
FHIT 410
Fibonacci dose escalation scheme 1570–1

fibrin deposits 1019
fibrin/fibrin degradation products, urothelial carcinoma screening 656
fibrinogen
  extracellular matrix 210, 242
  hepatocellular carcinoma 570–3
fibroblast growth factor (FGF) 1025, 1034
  differentiation regulation 959–60
  see also acidic fibroblast growth factor (aFGF); basic fibroblast growth factor (bFGF)
fibroblast growth factor receptor (FGFR) 1486
fibroblastic tumours 746–8
fibroblasts 1025, 1084
  extracellular matrix interactions 1084
  skin 477
  TGF-β response 181
fibrohistiocytic tumours 748–9
fibrolamellar hepatocellular carcinoma 572–3
fibroma
  heart 777
  nasopharynx 503
  ovarian sex-cord stromal tumours 640
  testis 682–3
fibromatoses 746–7
  deep (desmoid tumours) 746
  imaging 743–4
fibromyxoid sarcoma 747
fibronectin 8, 210, 215, 236
  angiogenesis induction 971
  antibody-mediated targeted drug delivery 1481
  antineoplastic synthetic peptides 1476
  basement membrane 987, 1010–12
  extracellular matrix 210, 212
    degradation during malignant invasion 231–2
    endothelial cell interactions 939–40
    TGF-β regulation of expression 181
  fragments
    angiogenesis-inhibiting 241
    antineoplastic 1475–6
    retroviral-mediated gene transfer enhancement 1481
  haematopoietic cell interactions 1104
  integrin binding 211, 242, 1026
    cell growth regulation 26
    endothelial cell interactions 939–40, 945
  tumour cell production 984
fibrosarcoma 747
  congenital 747
  multidrug resistance proteins (MRPs) expression 1353
  seminal vesicle 678
  variants/related tumours 747
fibrosis 181
fibrous dysplasia
  orbit 835
  pathology 723
  plain radiography 1148
fibrous histiocytoma
  benign 723
  orbit 834
field theory of carcinogenesis 250, 967–8, 1403
  oral squamous cell carcinoma 493–4
FIGO staging, endometrial cancer 633–4
filamen 214
filter sterilization 879
finasteride 673, 1434
fine-needle aspiration biopsy (FNAB) 461, 464, 1250–1
  breast lesions 1564
  colon cancer 1250
  hepatocellular carcinoma 569
  imaging guidance 1247–8
  liver lesions 1249–50
  lung tumours 1248
  soft tissue tumours 744–5
  technique 1247–8
  thyroid cancer 601–2
  see also percutaneous biopsy
finger clubbing 516
Fisher 344 rat 365
FK506 1048
FKBP12 182–3
FKHR 739
FKHR/PAX3 fusion 753
flavonoids 424, 429, 432, 437–8, 443
  dietary 391, 1482
flavopiridol 21–2, 1499, 1517
Flexner–Wintersteiner rosettes 829–31
FLI-1 translocation 722–3, 738, 756
Flip antiapoptotic proteins 124–5
flk-1 *see* vascular endothelial cell growth factor receptor 2 (VEGFRI-2)
flow cytometry 255, 615, 958–9, 964
floxuridine *see*5-fluoro-2′-deoxyuridine
Flp–frt gene targeting system 905
flt-1 1 *see* vascular endothelial cell growth factor receptor (VEGFRI-1)
FLT-3 ligand (flt3l) 1379
flt-4 3 *see* vascular endothelial cell growth factor receptors (VEGFRI-3)
fludarabine 1281–4, 1552–3
  mechanism of cytotoxicity 1281–4
  metabolism 1281–3
fluorescence spectrometry 299
fluorescent *in situ* hybridization (FISH)
  central nervous system tumours 820
  genomic instability measurement 899–900
  lymphomas 692–3
  RB gene aberrations 35–6
  soft tissue tumours 746
5-fluoro-2′-deoxyuridine (floxuridine) 1273
  metabolism 1273–5
  transport 1273, 1275
5-fluorocytosine 1470
  cytosine deaminase-mediated suicide gene therapy 1069, 1536–7, 1541
    bystander effect 1069, 1536–7
    clinical trials 1537
    combined immunotherapy 1537
[^{18}F]fluorodeoxyglucose (^{18}FDG) imaging 1227–31
  brain tumours 1216
  breast cancer 1217
  colorectal tumour recurrence 1216
  head and neck tumours 1217
  image interpretation 1231
  image registration 1230–1
  lung cancer staging 1228
  lymphoma staging 1228
  melanoma staging 1228–9
  principle 1227–8
  protocols 1230
  solitary pulmonary nodule 1228
5-fluoropyrimidines 1273–7
  orally bioavailable 1276–7
fluoroscopic percutaneous biopsy guidance 1247
5-fluorouracil 20–1, 559–60, 594, 1071, 1273–6, 1421–3, 1425, 1470
  ^{18}F positron emission tomography (PET) 1231
  cytosine deaminase activation (suicide gene therapy) 1069, 1536–7
  intracellular metabolism 1273–5
    RNA processing effects 1275
  leucovorin (folinic acid) modulated treatment 1273
  liver tumour chemoembolization 1256

5-fluorouracil – *continued*
mechanism of action 1275, 1536–7
neuroendocrine tumour treatment 595, 609
orally bioavailable analogues 1276–7
resistance mechanisms 1275–6
structure 1273
transport 1273
fluorspar miners 512
flutamide 1426
FNCLCC sarcoma grading system 741–2
focal adhesion kinase (FAK) 173, 211
chemokine receptor signal transduction 156–7
paxillin interaction 985
focal adhesions 173, 214–15
integrins 173
clustering 213
folate
antigenotoxic properties 424
dietary 329, 589, 1287
endogenous compounds 1287–9
transport 1288–9
folinic acid *see* leucovorin
follicular dendritic cells 689–90
follicular dendritic sarcoma 737
follicular lymphoma 698–700
clinical features 698
genetic features 699–700
t(14;18)(q32;q21) translocation 692, 699–700
grading 699
historical aspects 693
immunophenotype 699
minimal residual disease (MRD) detection 475
morphology 698–9
folylpolyglutamate synthase 1288–9
Food and Drug Administration (FDA)
FDA/CDER Carcinogenicity Assessment Committee (CAC) 372–3
IND (investigational new drug) applications 1041–2, 1570
NDA (new drug application) submission 1042, 1574
role in drug development 1041–2, 1570
food preservation 525
food supplements 423
foregut 545
Forkhead-family transcription factors 131
formaldehyde 414, 505
fixation 468–9
formalin fixation 464
formestane 1457
fos (*fos*)
antisense targeting 1441
cell growth signalling 15
ribozyme targeting 1445
FRAF-6 124
fragmin 1468
free radical-scavengers 422, 428
free radicals
DNA damage 278–9, 302
gastric cancer pathogenesis 530
ionizing radiation generation 1359–61
skin cancer pathogenesis 478
ultrasound generation 311
Frizzled gene family 198
Frizzled transmembrane receptor proteins 196, 198–200
Frizzled-Related Protein (FrzB; FRP) 199
frozen cells
basic culture technique 883
storage facilities 877
fruit and vegetables 528–30, 549, 589, 591
antioxidants 428
cancer chemoprevention 410, 423–4, 549
DNA adduct formation influence 379
ftorafur 1276–7
Fuchs adenoma 831–2
FUdR, liver tumour chemoembolization 1256
Fuhrman nuclear grading 649
fumagillin 971
functional magnetic resonance imaging (MRI), brain tumours 1202
fungizone 882
FUS/AFT-1 fusion 748–9
FUS/TLS fusion 738–9
Fyn 166–8, 172–3

G

G cells 524–5
G1 checkpoint 17–18, 953–5, 1324, 1508, 1512, 1526
antitumour compound molecular targeting 1044
Cdc25A 1515–17
DNA damage-related activation 18
p53 896, 1515–16, 1518
radiotherapy responses 1363
G2 checkpoint 17–18, 1324, 1508
Cdc25C 1515, 1517
DNA damage-related activation 18
radiotherapy responses 1363
therapeutic approaches 21–2, 1044
G-protein-coupled receptors 25, 155, 171
GADD45 905–6
GADD153 495–6, 896
gag 76
galactography (ductography) 1241–3
contraindications 1241–2
preoperative localization procedures 1242
galectin 1016
galectin 3 558
gall bladder cancer 262–3, 576–7
gall bladder development 563
gallic acid 427
gallium-67 citrate scanning 1222–3
indications 1222
interpretation 1222
principles 1222
procedure 1222
gamma camera 1215
γ rays 1359–60
cellular senescence 111
DNA damage 89, 310
linear energy transfer (LET) 307
γ-tubulin 1323
γ_c receptor subunit 147
γ_c-dependent cytokine receptors 147
immunosuppression-related defects 149
JAK–STAT signalling 148–9
γ_c-dependent cytokines 147
ganciclovir, HSV-TK-mediated suicide gene therapy 1069–70, 1534–6, 1541
adverse reactions 1536
bystander effect 1069, 1534–5
clinical trials 1535–6, 1541
combined immunotherapy 1536
gap junction-mediated sparing effect 1535
gangioneuroblastoma 815
gangioneuroma 815
gangliocytoma 806

dysplastic of cerebellum (Lhermitte–Duclos disease) 806, 820
gangliglioma 805–6
desmoplastic infantile 807
ganglion 744
ganglioneuroblastoma 800
gap junctions 9, 983, 1010, 1031–2
cytotoxic 'bystander effect' 1069, 1535
intercellular communication
abnormalities in cancer cells 1032
measurement 1031–2
metabolic cooperation assays 1032
progression of cancer 869–70
structure 1031
therapeutic approaches
functional enhancement 430, 441
genetic prodrug activation therapy 1535
Gardner syndrome 36–7, 252, 737, 746, 831
gastric acid secretion 524–5, 528–9
gastric cancer 523
aetiology 528–32
autoimmune gastritis association 531
chemoprevention 1432
clinical features 533
dietary factors 395–6, 526–31, 1432
diffuse tumours 528, 532
E-cadherin gene mutations 1082
epidemiology 264, 526–8
Epstein–Barr virus 60–1
familial clustering 35, 527–8, 542
glutathione-*S*-transferase (GST) isoform expression 1355
Helicobacter pylori 335, 344, 527–8, 530, 532
intestinal metaplasia 527–9
intestinal-type cancer 527–8
Japan 262
community screening 533
mismatch repair gene defects 30, 550
molecular genetics 542
mortality 344
mutiple tumours, clonality analysis 250–1
pathology 537
post-surgical gastric remnant 531
radiation-induced 531–2
staging 541, 1558–9
TGF-β receptors
expression 187
TβRII mutations 186–7
tobacco association 402
treatment 541
endoscopic surgery 1562
gastric glands 524
gastric lymphoepithelial carcinoma 340–1
gastric lymphoma
Helicobacter pylori 344
MALT type 548
marginal zone B cell lymphoma 700
gastric pits 524
gastric polyps 250
gastrin 524–5
gastrinoma 587
clinical features 586
liver tumour chemoembolization 1259
gastrointestinal tract
development 545
gross appearance 545
multidrug resistance proteins (MRPs) 1352
organization 523
small intestine 8
wall structure 523, 545–7
gastrointestinal tract tumours 547–8
epithelial 547
lower gastrointestinal tract 545–60
lymphoid 548
models 1090–1
metastasis 1091
xenografts 1091
neuroendocrine 609
secondary (metastatic) 548
stromal tumours 548
chemoembolization 1259
immunohistochemistry 745
telomerase activity 113, 116
upper gastrointestinal tract 523–43
GATA-1 137, 140
Gaucher's disease 1545
GBP/Frat 196, 199, 203
GCK 154
gelatinases 231–2
Gelfoam embolization 1256
gelsolvin 120–1
gene silencing by hypermethylation 324, 331
gemcitabine 21, 594, 1491
clinical trials in pancreatic cancer 1574–5
mechanism of action 1279
metabolism 1279
structure 1277
transport 1279
gemtuzumab ozogomycin 1386
gender-related cancer risk 262–3
gene expression arrays 591
gene expression profiling 1511–12
gene gun 1397
gene targeting 894–5, 903, 904–5
Cre-loxP system 905, 910
multiple mutations 910
targeting vector 903–4
tumour-suppressor genes 896–7
gene therapy 473, 1521–30
antiangiogenic 1073–4, 1455
brain tumours 1135
cancer-directed strategies 1072, 1524–5
clinical trials 1521–22, 1529
combined treatment approaches 1523–7
conventional protein drug therapy 1521
cytokines 1074
dendritic cell targets 1075–6
drug resistance genes 1076–7
gene replacement strategies 1524–5
gene transfer 1523
head and neck tumours 496
historical aspects 1521–2, 1533
early targets 1521–2, 1533
host-directed strategies 1077
MDR1 transfer 1047
models 1069–77, 1523
non-viral vectors 1523, 1528
oncogenes 1527
oncolytic viruses 1071–2
prodrug activation therapy *see* genetic prodrug activation therapy
promoter sequence/protein-coding sequence
recombinant DNA constructs 1523
suicide genes 1069–70, 1135
techniques 1522–23
transfection 1523
transgene/knockout models 1523

gene therapy – *continued*
tumour-suppressor genes 1070–1, 1307
replacement 1526–7
vector design 1521–2, 1528–9
vector targeting 1528–9
vector-generated side effects 1529
viral transduction efficiency 1525
viral vectors 1523, 1529
genes 1522
genetic counselling 462
genetic drift 367
genetic prodrug activation therapy 1533–41
bystander effect 1533–4
carboxylesterase–irinotecan 1540
carboxypeptidase G2–CMDA 1539
cytochrome P450 1A2–acetaminophen (paracetamol) 1539–40
cytochrome P450 4B1–2-aminoanthracene 1539
cytochrome P450 4B1–4-ipomeanol 1539
cytosine deaminase–5-fluorocytosine 1536–7
enzyme parameters 1533–4
linamarase–linamarin 1540–1
nitroreductase–CB1954 1538–9
principle 1533
prodrug properties 1533–4
purine nucleoside phosphorylase–p-(β-D-2-deoxyerythropentofuranosyl)-6-methylpurine 1539
thymidine kinase–ganciclovir 1069–70, 1534–6
thymidine phosphoylase–5′-deoxy-5-fluorouridine 1539
VZV thymidine kinase–araM 1539
genetic screening 462
inherited colorectal cancer syndromes 551–2
susceptibility markers 384–5
genistein 430, 439, 449–50, 920, 1477–8
genomic imprinting 324–8
cytosine methylation 325–6
histone acetylation 325–6
genomic instability 17–18, 87–104
apoptosis defects 119
centrosome abnormalities 88–9
human genetic disorders 87–8
ionizing radiation induction 310–11
measurement 899–900
microsatellite instability 88, 94
mouse models (chromosome 11 deletions) 897–8
role in carcinogenesis 87
tumours 88
chromosome instability 88
colorectal cancer 556–7
genotoxicity testing 351–63, 426
chemopreventive drug development 450
guidelines 351, 360–1
high-throughput screening 362
historical aspects 351
interpretation of results 360–2
risk–benefit continuum 361
thresholds for exposure 361
metabolically competent cell lines 363
new techniques 362–3
predictive expert computer systems 362–3
primary test systems 352–7
chromosomal damage in cell culture 355
chromosomal damage in rodents 356
in vitro mammalian cell tests 354–5
in vitro metabolic activation 352
mutation in bacteria 352–4
rat liver unscheduled DNA synthesis 355–7
supplementary test systems 357–60
aneuploidy tests 358–60
comet assay 358
transgenic models 357–8
gentamicin 882–3
geographical incidence 264
germ cell tumours
central nervous system 792, 812–13
heart/pericardial space 777
ovary 635
epidemiology 635
paediatric tumours 640
pathology 640
testis 1192–3
clinical management 685
epidemiology 683
immunohistochemistry 683
molecular genetics 684–5
pathology 679–83
precursor lesions 679–80
prognostic factors 685
risk factors 683–4
tumour markers 473
see also non-seminomatous germ cell tumours
germinal centres 689–90
B cell activation 690
progressive transformation 690
germinoma 813
GF109203X 23
GFAP-v-*src* transgenic mouse model 1133
GFP (green fluorescent protein)
knock-in mouse studies 905
metastasis *in vivo* models 1000–1, 1003–4
giant cell fibroblastoma 748
immunohistochemistry 745
giant cell tumour of bone 724
plain radiography 1148
GLI 740
glial cells 3, 789
development 789–90
glial fibrillary acidic protein (GFAP) 467
astrocytomas 802–3
atypical teratoid/rhabdoid tumour 800
brain tumours 795, 1123–4
choroid plexus papilloma 808
dysembryoplastic neuroepithelial tumour 806–7
endolymphatic sac papillary tumour 848
ependymal tumours 805
glioblastoma 802
oligodendroglioma 803
primitive neuroectodermal tumours 798–9
salivary gland carcinoma 498
see also GFAP-v-*src* transgenic mouse model
glial tumours *see* gliomas
glioblastoma 801–2
giant cell variant 802
PTEN mutations 30, 907
glioblastoma multiforme 1123–5
animal models 1128–9
astrocytoma progression 819, 821, 1123–4
cytokine secretion 1124
xenografts 1127
gliomas 792, 794, 819, 1123
C6 model 1130–1
cell culture 1126
established cell lines 1126
clonality analysis 251
CNS-1 model 1131

mixed 792, 805
molecular genetics 1123–4
multidrug resistance proteins (MRPs) expression 1353
nasal 506
optic nerve 835
pathology 800–5
retina 830–1
spontaneous animal tumours 1127–8
TGF-β TβRII mutations 186–7
thallium-201 scanning 1220–1
see also mixed neuronal–glial tumours
gliosarcoma 802
9L model 1130
Glivec (imitanib) 1497, 1500–1
glomus tumour 503
glottic carcinoma 508, 1177–8
glottis 503
glucagon 580
glucagonoma syndrome 609
glucagonomas 587, 595
clinical features 586
glucocorticoid hormones 135
glucocorticoid receptor 135
glucosinolates 424
glutathione 427, 438
antioxidant properties 422, 438
inducers 438
MRP-mediated drug resistance 1353–6
glutathione peroxidase 422, 438
selenium dependent/non-selenium-dependent 428, 438
glutathione-*S*-transferase (GST) 422, 427, 438
gastric cancer-associated polymorphism 542
inducers 428, 438
isoform expression in tumours 1355
multidrug resistance mediation 1354–5
in vitro studies 1355–6
transfection studies 1356
glycogen storage disease 568
Glycogen Synthase Kinase 3 (GSK3)
axin binding 200–1
β-catenin phosphorylation 199–201
GBP/Frat inhibition 203
Wnt signalling pathway (β-catenin destruction complex formation) 196, 199–200
glycosaminoglycans
extracellular matrix 210, 1026
tyrosine kinase ligand availability regulation 170–1
glycyrrhetinic acid 429
goblet cells 7
large intestinal epithelium 547
respiratory tract epithelium 503–4
Goldenhar syndrome 827–8
Gompertz (exponential) growth curve 957
gonadal dysgenesis (gonadoblastoma) 637
gonadal steroid hormone secreting tumours 600
good clinical practice (GCP) 1041
good laboratory practice (GLP) 1041
good manufacturing practice (GMP) 1041
Gorham disease 723–4
Gorlin–Goltz (basal cell naevus) syndrome 39, 482, 636, 820, 827, 1425
medulloblastoma 39
ovarian cancer 44
PTC mutations 483–5
skin tumours 478
goserelin 1455
gout 1329
gp100 tyrosinase 1392
grade, tumour 471
graft versus host disease 1548, 1550–1
cord blood transplantation 1549
treatment 1548
graft versus leukaemia effect 1550–2
graft versus tumour effect 1551
stem cell transplantation 1552
gramicidin D 1048
granular cell tumour
heart 783
larynx 508
granular epidermal layer 8
granulocyte colony-stimulating factor (G-CSF) 1103–4
angiogenesis 237
antitumour immune response 938, 1074
retinoids combined treatment 1411
granulocyte–macrophage colony-stimulating factor (GM-CSF) 1103, 1124
antitumour immune response 938, 1074
antitumour therapy 1379
cellular vaccine production 1391
haematopoiesis 6
in vitro dendritic cells generation 1378
granulocytes 1103
granulocytic sarcoma
orbit 835
testis 683
granulosa cell tumour 632
ovary 640
testis 682
granzyme B 121–2
Grawitz tumour *see* renal cell carcinoma
gray (Gy) 308
Grb2 166–8, 173, 1486–7
Ras-MAP kinase pathway initiation 168
see also SOS–Grb2 complex
GRIP1 *see* SRC-2
GRO-α 1088–9
GRO-β 239
Gross murine leukaemia virus 27, 335
Groucho repressors 204
growth arrest, extracellular matrix interactions 213–4
growth factor receptor blockade 1485–93
bispecific antibodies 1492
dominant-negative mutant receptors 1492–3
methods 1488
monoclonal antibody therapy 1488–91
receptor tyrosine kinase inhibitors 1491–2
single-chain antibodies 1492
growth factor receptors 953
focal adhesion complexes (focal contacts) 213
tyrosine kinases 1486
growth factors 25, 953, 1084–5, 1089
altered forms in oncogenes 955
angiogenesis 229, 239, 971
cell cycle regulation 14–15, 955, 1526
haematopoiesis 1102, 1104
hepatocellular carcinoma 571
metastasis 987–8, 999–1000, 1011–12
oesophageal adenocarcinoma 542
pleiotropic effects 1084
proliferation stimulation 999–1000
small intestinal epithelium 8
TGF-β regulation of expression 181
therapeutic antagonism 20
chemopreventive strategies 429–30, 439
growth fraction 255, 956–7

growth fraction – *continued*
rodent chemically-induced mammary carcinogenesis 932–3
growth hormone 25, 276–7
angiogenesis-inhibiting fragments 242
growth regulating gene mutations 25
growth, tumour 247, 253–5, 937–8, 953–68
angiogenesis dependence 235–6
cell loss rate 254–5
cell production (birth) rate 254
early detection 955–6
early stages 938
growth curves 253–4
in vitro cell culture models 955–6
measurement 957–9
parameters 956–7
kinetic 254–5
Gsα mutations 606–7
insulinomas 606–7
pituitary adenomas 606–7
thyroid adenomas 606–7
GSK3-β signalling 224–8
GSTM1 polymorphism 304
GSTP1 polymorphism 304
GSTT1 polymorphism 304
GTL2 327–8
guanylate kinase 1272–3
gypsetin 1048
Gyromitra esculenta (false morel) 391
gyromitrin 391

H

H&E (haematoxylin and eosin) stained sections 463–5
H19 737
genomic imprinting 326–7
H69AR 1351, 1353
H-*ras* 1469, 1525
antitumour compound molecular targeting 1044
malignancy-related changes 299, 1087
papilloma phenotype induction 1087
ribozyme targeting 1444–5
Ha-*ras*
cellular senescence 111
transgenic mouse model 368–9
HaCaT cell line 1087
malignant conversion 1088
microenvironmental influences 1088
haemangioblastoma 820
central nervous system 811
haemangioma
bone 723–4
central nervous system 810–11
imaging 743–4
larynx 508
haemangiopericytoma 755
central nervous system 811
orbit 834
haemangiosarcoma 286
haematological malignancies 1101–2
differentiation relationship 1101, 1106–7
cell culture studies 1110
dissemination 1101
models
animal 1110–12
blood cell sources 1108–9
cell culture techniques 1109–10
cellular differentiation 1110
clinical relevance 1112
host immune responses 1111–12
limitations 1111–12
pathogenesis 1108
principles 1109–11
safety precautions 1109
transgenic 1110–11
oncogenes 1102–7, 1105
preferential organ colonization 1101
stem cell transplantation 1545–7
telomerase activity 113
tumour viruses 1104–5
tumour-suppressor genes 1105, 1108
haematoma 743–4
haematopoiesis 6, 1101–4
blast forms 1102
bone marrow 1102–3
mouse models 1103
cell adhesive interactions 1104
children 1102
cytokines 6–7, 1102
differentiation 1102–4
extramedullary 1102
fetal 563, 1101–2
growth factors 1102, 1104
progenitor cells 1102
in vitro quantitation 1102
regulation 1104
responses to environmental stimuli 1102–3
haematopoietic stem cells 1102–3
bone marrow ‘homing’ 1104
drug resistance gene transfer 1076
haematopoietic/lymphoid 1103
haemopoietic system reconstitution 1102–3
therapeutic DNA repair enhancement 1512
haematopoietic system 6–7
cell lineages 6, 7
haemochromatosis
genetic 568
hepatocellular carcinoma 566, 568
haemocytometer
cell viability assessment 884
slide 878
haemoglobin adducts 377–8
hair follicles 477
hairpin ribozyme 1440–1, 1445
hairy cell leukaemia 705, 1383
human T cell leukaemia virus type 2 (HTLV-2) 1104–5
immunotoxin therapy 1386
halichondrin B 1328
halofuginone 1478
chronic graft versus host disease 1548
Ham F-12 880–1
hamartin 836–7
hamartoma
colorectal polyps 547
lung 1161, 509–10
hammerhead ribozyme 1440, 1445
hand-assist laparoscopic surgery 1561
Hand–Schuller–Christian disease 730
Handport System 1561
Hank’s balanced salt solution 880
hard palate carcinoma 492
hardwood dusts 505
harmonic ultrasound imaging 1187
hepatic masses 1195–6
hazardous pathogens 876

HBME-1 604
thyroid papillary carcinoma 605
HCT16 1091
hDlg 82
hE10 *see* Bcl-10
head and neck cancer
^{18}FDG positron emission tomography (PET) 1217
chemoembolization 1256, 1260
chemoprevention 496, 1427–31
E-cadherin aberrant expression 1082
human papillomavirus (HPV) 337
interleukin 2 (IL-2) therapy 1375
magnetic resonance imaging (MRI) 1203–4
molecular genetics 494–5
p53 gene therapy 1070–1
percutaneous ablative therapy 1261
second primary tumours 493–4
telomerase activity 113–15
tumour-selective oncolytic virus therapy 1071
see also oral cancer; salivary gland carcinoma
head and neck carcinoma *in situ* 113
heart development 777
heart tumours 777–86
aetiology 783–4
classification 778
clinical management 786
epidemiology 783
in utero development 777
molecular genetics 785
pathology 778–83
prognostic factors 785
heat-induced epitope retrieval (HIER) 468–9
heavy metals 414
Heffner tumour *see* endolymphatic sac papillary tumour
helical endoluminal CT colonography, colon cancer screening 1159
Helicobacter pylori 344
cytotoxin-associated gene A (*cagA*) strains 344, 530
epidemiology 336, 344, 530
eradication 344, 700, 1432
gastric cancer 335, 344, 526–30, 530–2, 1432
gastric non-Hodgkin's lymphoma 344
MALT type marginal zone B cell lymphoma 700
vaccines 344
heliotrine 390
Heliotropium 390
helminth infection-related cancers 345–7
helper T cells
cytokine gene therapy targets 1074
DNA vaccine response 1398
hemidesmosomes 9, 215, 216, 218–19
basement membrane-induced mammary epithelium differentiation 217
hemihypertrophy 573
HEp2 cell line 1351
Hep par1
cholangiocarcinoma 575
hepatocellular carcinoma 570–1
Hepadnaviridae 66, 338, 566
heparan sulphate 210, 1010
haematopoietic cell interactions 1104
therapeutic potential of structural mimics 1477
heparanases
malignant invasion/metastasis 231–2, 985
therapeutic inhibition 1480–1
heparin 170–1, 1468
heparin-binding EGF-related growth factor (HB-EGF) 1486–7
heparin-binding proteins 1441–2
heparin-like molecules 1477
hepatic artery 563–4
embolization 572, 1262
hepatic atypical adenomatous hyperplasia 250
hepatic duct 563
tumours 577
hepatic focal nodular hyperplasia 250
hepatic veins 563–4
hepatitis A virus 567
hepatitis B virus 49–50, 66–71, 75, 337–9, 566
blood products screening 338–9
cellular transformation 68–70
chronic infection 337–9, 566
core antigen (HBcAg) 66, 338, 566
e antigen (HBeAg) 66, 338, 566
epidemiology 336
genome organization 67–8, 338, 566
geographical distribution 68, 337–8
hepatocellular carcinoma 66, 70, 250, 260, 262–3, 265–6, 335, 337, 339, 565–6, 568
hepatitis C co-infection-related risk 338–9, 567
persistent infection-related risk 338–9, 567
prevention 338–9, 566–7, 569, 1433
historical aspects 66
immune response 71–3
insertional mutagenesis 69–70, 566–7
occupational exposure 414–15
pathogenesis 68–70
serology 68
surface antigen (HBsAg; Australia antigen) 66, 338, 566–7
transcriptional products 67–8
transmission 338, 566, 1109
preventive strategies 339
vaccines/vaccination programmes 70–1, 339, 566–7, 569, 1417
virion life cycle 66, 566
genome replication 67, 338, 566
RNA pregenome (pRNA) reverse transcription 67, 338, 566
virion structure 66, 338
X protein (transactivator protein) 68, 339, 566–7
hepatitis C virus 75, 342–3
blood products screening 342–3
chronic infection 342–3
epidemiology 342
hepatocellular carcinoma 66, 250, 335, 342–3, 565–9, 1433
hepatitis B co-infection-related risk 339, 567
non-Hodgkin's lymphoma 342
occupational exposure 414–15
oral cancer 342
prevention 343, 569
subtypes 342, 572
transmission 342, 567, 1109
hepatitis D virus 567
hepatitis E virus 567
hepatitis G virus 567
hepatitis TT virus 567
hepatitis viruses 1081
hepatobiliary system 563–77
hepatoblastoma 563, 573
α-fetoprotein levels 573
associated congenital anomalies 573
prognosis 564–5
hepatocellular adenoma 576
hepatocellular carcinoma 563–73
aetiology 566
aflatoxin B_1 273–4, 389, 390
induction mechanism 390
aflatoxins exposure 338–9, 379, 565–6, 568

hepatocellular carcinoma – *continued*
alcohol abuse association 565, 566, 567, 568
α-fetoprotein production 563, 572
tumour detection 569
Axin mutations 200–1, 206
chemical/drug exposure associations 294–5, 568
chemoembolization 572, 1257–9, 1262
percutaneous alcohol injection 572, 1176, 1264,
radiofrequency ablation combined treatment 1257–9
chemoprevention 1433
chemotherapy 572
computed tomography (CT) 1166
congenital abnormality associations 568
definition 565
diagnosis 572
epidemiology 564, 565–6
Epstein–Barr virus 60–1
fibrolamellar type 572–3
gallium-67 citrate scanning 1223
grading 570
hepatitis B virus 66, 70, 250, 260, 262–3, 265–6, 335, 337–9, 565–8
prevention 339, 566–7, 1433
hepatitis C virus 54–5, 250, 335, 342–3, 565–6, 566, 567, 568, 1433
imaging 572, 1149
immunocytochemistry 570–1
inferior vena cava obstruction 568
liver transplantation 572
metabolic disorder associations 566, 568
mixed cholangiocarcinoma 575
molecular genetics 571
multicentricity 569–70
pathology 569–71
cytological variants 570
macroscopic 569–70
microscopic 570–1
percutaneous ablative therapy 1261–2
percutaneous microwave coagulation 1265
precancerous changes 568–9
clonality analysis 250
presentation 572
prevention 569
prognosis 564–5, 572
radiofrequency ablation 1257–9, 1264
spontaneous regression 572
spread 569
vascular 569–70
surgical resection 572
treatment 572
ultrasound 569, 1196
hepatocyte growth factor (HGF) 1084
angiogenesis 237, 239
invasion/metastasis 231–2
hepatocyte growth factor receptor (c-Met) 1486
hepatocytes 10, 563–4
plate arrangement 563–4
zones 564
regenerative potential 10
hepatoma *see* hepatocellular carcinoma
hepatosplenic γ/λ T cell lymphoma 708
HEPES (*N*-2-hydroxyethylpiperazine-*N*-2-ethanesulphonic acid) 880, 881
HepG2 363
HER *see* epidermal growth factor receptor (EGFR)
HER2/*neu* 1383
antisense oligonucleotide targeting 1443
cellular vaccine target 1392–3
epidermal growth factor receptor (EGFR) heterodimer formation 1487
gene therapy 1523
monoclonal antibody *see* trastuzumab (Herceptin)
mutations
pancreatic cancer 592
pancreatic intraepithelial neoplasia (PanIN) 581–2
overexpression
breast cancer 175–6, 467, 472–3, 612, 966, 615, 966, 1092–3, 1443
cholangiocarcinoma 575
epithelial carcinomas 1089
gall bladder carcinoma 577
gastroenteropancreatic endocrine tumours 607
lung adenocarcinoma 512
small-cell lung carcinoma 512
soft tissue sarcomas 740, 758
therapeutic strategies 966, 1486, 1489
tumour marker 966
herbal remedies 390
Herceptin *see* trastuzumab
hereditary haemorrhagic telangienctasia 568
hereditary intraepithelial dyskeratosis, benign 827
hereditary non-polyposis colorectal cancer (HNPCC) 43–4, 1425, 1508, 1510–1
colorectal cancer 550
DNA mismatch repair defects (*hMLH1/hMSH2* mutations) 30, 43–4, 96, 324, 550, 556
gastric cancer 542
genetic/predictive screening 551–2
ovarian cancer 44, 550
TGF-β TβRII somatic mutations 186–7, 550
herpes simplex virus 1071–2
oncogenic activity 71
oncolytic activity 1071
safety modification for gene therapy 1071–2
herpes simplex virus thymidine kinase (HSV-TK)
ganciclovir prodrug activation 1534–6
mutants 1070
suicide gene activity 1069–70
herpesviruses 49–50
gene therapy vectors 1528
brain tumours 1135
genomic organization 56–8
heterochromatin 317–18
replication 321
heterocyclic amines
dietary 392–3
DNA adduct detection 298–9
DNA interactions 296–7
heterogeneity of tumour tissue/cells 863–4, 960–1, 1029, 1086
central nervous system tumours 1124
metastatic potential 1001, 1004
progession of cancer 866–7, 870
tumour cell diversification/clonal dominance 866–7
N, N′-hexamethylene bisacetamide (HMBA) 966–7, 1408
clinical studies 1412–3
retinoid combined treatment 1411
hexamethylmelamine 1298
metabolic activation 1298–9
HFE mutations 568
Hgprt (hypoxanthine–guanine phosphoribosyl transferase) 353
HHF35 *see* muscle-specific actin (MSA)
HICI (hypermethylated in cancer) 325
Hif-1 239
high endothelial vessels 690, 1009–10
high-altitude flight 308
high-density cell culture 884–5
high-efficiency particulate air (HEPA) filters 873–5
high-throughput screening 362

drug screening strategy 1045
single-nucleotide polymorphisms (SNPs) 377
hindgut 545
Hiroshima 308, 531–2
histiocytosis X *see* Langerhans cell histiocytosis
histochemical staining 464–5
histological examination 464–5
special stains 464–5
histone acetylation 317–18
genomic imprinting 325–6
histone acetyltransferase (HAT) 317–18, 320–1
histone deacetylase (HDAC) 317–18, 320–1, 321–2
Rb protein complexes/cell cycle regulation 1514
therapeutic strategies 330–1
histone deacetylase 2 (HDAC2) 321
histone deacetylase inhibitors 1404–7
all-*trans*-retinoic acid combined treatment 1411–12
clinical studies 1412–13
histones 317–18
HLA-DR 696
HMB-45 466–7
uveal melanoma 834
HMEC-1 cell line 980
HMGIC (*HMGI-C*) 469–70, 750
HMT-3522 218
Hodgkin cells 713–16
Hodgkin's disease 262, 713–16
antigen expression 710, 715
bone involvement 730
central nervous system 812
classical Hodgkin's lymphoma 713, 714–16
lymphocyte-depleted type 714–15
mixed cellularity type 714–15
nodular lymphocyte-rich type 713–15
nodular sclerosis type 714
classification 713
cytokine-mediated immunotherapy 1551
diagnostic criteria 713
donor lymphocyte infusion therapy 1551
Epstein–Barr virus 60–1, 339, 340, 716
gallium-67 citrate scanning 1222
historical aspects 693
human immunodeficiency virus (HIV) 343
iatrogenic lymphoproliferative disorders 718
sinonasal tract 507
stem cell transplantation 1546–7
WHO classification 694
hollow-fibre assay 1044
homeobox 960
homeobox genes 10, 1410
homeotic genes 960
Homer Wright rosettes 726–7, 755–6, 795–8, 807, 808, 815, 829–30
homing 236
homologous recombination 97–100, 1362
DNA double-stranded break repair 97–100, 1362
Rad51 system 99–100
knockout generation procedure 903–4
single-strand annealing 100
hormone replacement therapy 630, 632, 1139
hormone signalling
chemopreventive approaches 429–30, 439
gap junction-mediated intercellular communication 1032
hormone therapy 473, 1449–63
endocrine ablation 1449
hormones 599
chemical carcinogens 276–7
host surveillance mechanisms 869
Housefield units (HU) 1156
Hox genes 10
haematopoietic cell differentiation 1107
HOX11 1107
HOXD13 10
HOXDA9 10
hPMS1 550
hPMS2 94–5
hereditary non-polyposis colorectal cancer (HNPCC) 550, 557
Hrk 129–30
HsMAD1 82–3
hSNF5/INI1 mutations 759–60
atypical teratoid/rhabdoid tumour 800
Hsp70 137
Hsp90 137, 1452
therapeutic inhibition 1502–3
HT29 1091
HT-1080 1000–1
HTLV-1-associated myelopathy/tropical spastic parapesis (HAM/TSP) 76–7, 80
human anti-mouse antibodies (HAMA) 1224
human chorionic gonadotrophin (β-hCG) 515, 966
choriocarcinoma 813
germ cell tumour marker 473
seminomas 681
testicular germ cell tumours 683
human endogenous viruses 76
human foamy viruses 76
human herpesvirus-4 (HHV-4) *see* Epstein–Barr virus
human herpesvirus-6 (HHV-6) 71
human herpesvirus-8 (HHV-8) *see* Kaposi sarcoma-associated herpesvirus (KSHV)
human immunodeficiency virus (HIV) 76, 336, 339, 343–4, 508, 965
associated cancers 343, 783
blood products screening 343
Castleman's disease 691
chemokine receptor coreceptor activity 155
conjunctival squamous cell carcinoma 343
DNA vaccine trials 1399
genome organization/replication 343
historical aspects 343
Hodgkin's disease 343
Kaposi's sarcoma 343, 676, 810–11
lung cancer 512
lymphomas 784, 1104–5
central nervous system 704, 792–4, 812, 1220–2
Epstein–Barr virus association 340, 343, 704
non-Hodgkin's 343
plasmablastic 704
primary effusion 704
occupational exposure 414–5
ribozyme therapy 1440
soft tissue tumours 737
leiomyosarcoma 343
transmission 343, 1109
prevention 344
human immunodeficiency virus 2 (HIV-2) 343
human papillomavirus (HPV) 75, 336–7, 478, 1081, 1089
anogenital tumours 336–7
anal canal carcinoma 547
warts 54
cervical cancer 54, 336–7, 617–20, 626, 1435
hybrid capture technique 623–5
in situ hybridization 625–6
screening 337, 349, 621, 623–5
cervical intraepithelial neoplasia (CIN) 336–7
classification 50

human papillomavirus (HPV) – *continued*
conjunctival lesions 826
squamous cell carcinoma 337, 343
detection of infection 54
E6 protein 337, 620, 626–7, 1089
p53 interactions 53, 337, 620, 626–7
E7 protein 337, 620, 626–7, 1089
Rb-p105 interaction 53, 337, 620, 626–7
endometrial cancer 633
epidemiology 336
epidermodysplasia verruciformis 54, 337
haematological malignancies 1104–5
high-risk types 54, 336–7
immunology of infection 54–5
nasopharyngeal squamous papillomas 505
oesophageal cancer 337, 528
oral cancer (squamous cell carcinoma) 493
pathogenesis 54
koilocytosis 54
malignant progression 54
penile cancer 676–7
ribozyme targeting 1445
transmission 54, 336
vaccines 337, 627, 1435
virus-like particle (VLP)-based 337
human papillomavirus 1 (HPV1) 52–3
human papillomavirus 5 (HPV5) 54
cutaneous squamous cell carcinoma 337
human papillomavirus 6 (HPV6) 623
bronchial papilloma 510
laryngeal papilloma 507
warts 336–7
human papillomavirus 8 (HPV8) 54
cutaneous squamous cell carcinoma 337
human papillomavirus 11 (HPV11) 623
bronchial papillomas 510
laryngeal papilloma 507
warts 336–7
human papillomavirus 53, 54, 336, 337, 677
anal canal carcinoma 547
cervical cancer 335, 337, 1435
nasopharyngeal squamous cell carcinoma 505
human papillomavirus 18 (HPV18) 53, 54, 336
anal canal carcinoma 547
cervical cancer 337, 1435
endometrial cancer 633
nasopharyngeal squamous cell carcinoma 505
human papillomavirus 31 (HPV31) 336, 337
human papillomavirus 33 (HPV33) 336, 337
human papillomavirus 36 (HPV36) 481–2
human papillomavirus 45 (HPV45) 336
cervical cancer 337
human papillomavirus 51 (HPV51) 336
human papillomavirus 52 (HPV52) 336
human papillomavirus 58 (HPV58) 336
human papillomavirus 59 (HPV59) 336
human T cell leukaemia virus (HTLV) group 76–7
human T cell leukaemia virus type 1 (HTLV-1) 75, 76, 335, 341–2
adult T cell leukaemia/lymphoma 76–7, 80, 342, 707–8, 1104, 1108
molecular pathogenesis 80–4, 1108
disease associations 76–7, 341–2
epidemiology 77, 342
genomic organization 77, 342
HTLV-1-associated myelopathy/tropical spastic parapesis (HAM/TSP) 76–7, 80, 341–2
in vitro infection 78
infection prevention 84–5
proviral genome 78–9
pX region 78–9
Rex protein 78–9
RNA splicing regulation 79
RNA processing 79
serological diagnosis 84–5
Tax protein 78–9
trans-activation of transcription 80–1
trans-repression of transcription 81–2
transcriptional activation 79, 83
transcriptional inhibitors down-regulation 81
tumour-suppressor proteins inhibition 82–3
transmission 78, 85, 342
breast-feeding 78, 85
virion characterisitcs 342
human T cell leukaemia virus type 2 (HTLV-2) 77
hairy T cell leukaemia 1104–5
humanized antibodies 1380
Humatrix 1030–1
Hus1 103
Hus1:Rad1:Rad9 sliding clamp model 103
hus1 knockout mice 103
Hutchinson–Gilford progeria 108–9, 117
hyalinizing spindle cell tumour with giant rosettes 747
hyaluronic acid/hyaluronidase, urothelial carcinoma screening 656
hydrazines
mushroom 273–5, 391
tobacco smoke 407
hydrocarbons
hepatocellular carcinoma 568
lung cancer 512
hydrolytic reactions 277, 427
4-hydroperoxycyclophosphamide 1298
8-hydroxydeoxyguanosine 278–9
8-hydroxyguanosine 89
hydroxyphenanthrenes 528
4-hydroxytamoxifen, oestrogen receptor interaction 1454
hydroxyurea 966–7, 1291–2
mechanism of cytotoxicity 1292
metabolism 1291–2
hypercalcaemia
lung cancer 515
renal cell carcinoma 648
hypermethylated in cancer *see* HICI
hypernephroma *see* renal cell carcinoma
hyperparathyroidism, primary 600
hyperplasia 1085, 1087, 249–50
carcinogenesis 10–11
clonality analysis 251–2
hypertension 632, 648, 783
hypertrophic pulmonary osteoarthropathy 516, 1152
finger clubbing 516
hypoplasia 5
hypoxanthine–guanine phosphoribosyl transferase (HGPRT) 1272, 1284–7
hypoxia 1360–1
angiogenesis induction 237, 980, 1469–70
retinopathy of prematurity 977
therapeutic strategies 1470
tumours 237
vascular endothelial cell growth factor (VEGF) expression regulation 239, 1470
hypoxia inducible factor (HIF-1) 1360–1
hypoxia inducible factor 1α (HIF-1α) 980, 1470
hypoxia inducible factor 2α (HIF-2α) 1470
hypoxia-response element (HRE) 239

I

^{131}I imaging 1215, 1218–20
indications 1218
interpretation 1220
principle 1218
procedure 1218
whole-body scanning 1218–20
thyroid imaging 1218
see also MIBG scanning
I-Flice 124–6
IARC Monograph programme 259, 413–14
ibuprofen 432
ICAD 120–1
ICH (International Conference on the Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use) 366–73
ICH guidance
S1A 367–8
S1B 367–70
S1C 370–1
ICI 182, 780 1459–60
idarubicin 1333, 1552–3
mechanism of action 1338–9
IEX-1L 131
ifosfamide 756, 762–3, 1296
metabolic activation 1298
structure 1296
IGF-BP3 896, 905–7
IGF-IIR 737
frame-shift mutations 94
colorectal cancer 556–7
IκB 154
IκBα 1363
IKKα 131, 154
IKKβ 154
IKKγ (NEMO) 154
ileum 545
IMDM 880
imitanib (Glivec; CGP57148B; ST1571) 1497, 1500–1
immortalization 17, 959, 1403
cultured cell lines 1087
endothelial lines 979–80
establishment 961
telomerase 110–11
in vitro models 117
immune response 937–8, 1056, 1389–90
antiangiogenesis 1472
brain tumours 1123
cancer vaccine strategies 1390–91
chemopreventive strategies 430, 441
co-stimulatory molecules 1389–90
cytokines, gene therapy-mediated enhancement 1074–5
Epstein–Barr virus 61–2
evasion by tumours 1390
haematological malignancy models 1111–12
hepatitis B virus 70–1
human papillomavirus (HPV) infection 54
antibody response 54–5
cellular immune response 55
mathematical modelling 941–2, 950
memory 1056
immune-deficient rodent xenograft recipients 915–18
immunoassay, DNA adducts detection 298–9
immunoblastic sarcoma 1101
immunoglobulin gene rearrangements
lymphomas 473
multiple myeloma 730
immunoglobulin heavy chains
lymphoma 692
multiple myeloma 729–30
immunoglobulin light chains
lymphoma 691–2
multiple myeloma 729–30
precursor B cell lymphoblastic leukaemia/lymphoma 696
immunoglobulin-like adhesion molecules 984
immunohistochemistry 466–9
acoustic neuroma 846
angiomatoid fibrous histiocytoma 748–9
angiosarcoma 575–6
antigen–antibody complexes detection 467–9
atypical alveolar hyperplasia 513
atypical teratoid/rhabdoid tumour 800
bladder cancer 657
brain tumours 795
cardiac myxoma 784
cholangiocarcinoma 575
chordoma 727–8
desmoplastic small round tumour 745, 759
endolymphatic sac papillary tumour 848
epithelioid haemangioendothelioma 745, 754
epithelioid sarcoma 745, 758–9
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 727, 745, 755–6
false positive/negative staining 469
fibrosarcoma 747
hepatocellular carcinoma 570–1
islet cell tumours 586
jugulotympanic paraganglioma 846
Kaposi sarcoma 754
lymph node follicles 690
meningioma 810, 847
mesothelioma 770–1
micrometastases detection 474–5
middle ear adenoma 845
multiple myeloma 729–30
neuroendocrine tumour 602
positive/negative controls 469
primitive neuroectodermal tumours 798–9
prostate cancer 671
pulmonary pleomorphic carcinoma 519
renal cell carcinoma 652
rhabdoid tumour 745, 759–60
rhabdomyosarcoma 745, 752–3
salivary gland carcinoma 498
seminal vesicle carcinoma 678
sentinel lymph node biopsy 471–2
small-cell lung carcinoma 518
soft tissue tumours 745
squamous cell carcinoma 481
synovial sarcoma 745, 758
testicular tumours 683
thyroid papillary carcinoma 605
uveal melanoma 834
see also immunophenotype
immunoliposomes, antisense oligonucleotide delivery 1443
immunophenotype
adult T cell leukaemia/lymphoma 708
anaplastic large cell lymphoma 710
angioimmunoblastic T cell lymphoma 708
antibodies 691
B cell prolymphocytic leukaemia (B-PLL) 697
B type chronic lymphocytic leukaemia/small lymphocytic lymphoma 696
Burkitt's lymphoma 704

immunophenotype – *continued*
diffuse large B cell lymphoma 703
follicular lymphoma 699
hairy cell leukaemia 705
hepatosplenic γ/λ T cell lymphoma 708
Hodgkin's disease 710, 715
nodular lymphocyte predominance 714
lymphomas 467, 691
prognostic significance 694–5
marginal zone B cell lymphomas 700–2
nasal NK/T cell lymphoma 712
plasmacytoma 705
precursor B cell lymphoblastic leukaemia/lymphoma 696
precursor T cell lymphoblastic/acute lymphoblastic leukaemia (precursor T cell lymphoblastic lymphoma) 706
T cell large granular lymphocytic leukaemia 707
T cell lymphoma, peripheral, not otherwise specified 709
T cell prolymphocytic leukaemia (T-PLL) 706
immunoproliferative small intestinal disease (α heavy chain disease) 548
immunosuppression/immunocompromised patient
cutaneous malignancies 478, 481–2
Epstein–Barr virus-induced lymphoma 339–40, 784
JAK–STAT signalling/γ_c-cytokine receptor defects 149
lymphoproliferative disorders 716–19
transforming growth factor-β1 (TGF-β1) in tumorigenesis 191–2
immunosuppressive agents
iatrogenic lymphoproliferative disorders 718
P-glycoprotein interaction 1048
immunotherapy 473, 1055–66
active specific (ASI) 1061–2
adoptive cellular 1062–3, 1550–1
biological limitations 1063
bladder cancer 660
cytokine-mediated 1551–2
overcoming tolerance 1063
p53 gene therapy combined treatment 1070
immunotoxins 1386
^{111}In imaging 1215
^{111}In labelled monoclonal antibodies 1223
^{111}In pentetreotide (Octreoscan) scan 1224–5
in vitro cancer models 864–5
in vivo cancer models see animal models
Ina D 214
incidence data 261–2
incidence rate ratio 260, 261–2
incubator 877
incus 841–2
indole-3-carbinol 428, 438
indomethacin 423, 428, 432, 440
inert materials implantation 275–6
infection-related cancers 335–49, 389
bacteria 344
epidemiology 265, 335–6
helminths 345–7
historical background 335
preventive approaches 336
viruses 336–44
infectious mononucleosis 55, 58–61, 339–40
inflammation
arachidonic acid metabolites 442–3
chemopreventive intervention targets 436, 442–3
early tumour growth stages 938
radiotherapy response 1365
inflammatory bowel disease
colorectal cancer association 550
colorectal dysplasia 554
inflammatory myofibroblastic tumour 747–8
larynx 508
information bias 380
informed consent, clinical drug trials 1573–4
infusion pumps 1135
inhibitor of apoptosis (IAP) proteins
caspase inhibition 122–3
expression in cancers 122–3
initiation events 32, 1085–6
chemical carcinogenesis 286
ionizing radiation 311
multistage carcinogenesis 286–6, 1086–7
INK4 Cdk inhibitors 1108, 1365, 1513–14
deletion/mutation 1514–15
inner ear 841–2
inositol hexaphosphate (IP6) 1087
institutional review boards 1574
insulin 579–80
insulin receptor 161, 1486
insulin receptor substrate (IRS) complex formation 166–8
insulin receptor substrates (IRSs) 166–8
insulin-like growth factor I (IGF-I) 14, 395, 999–1000, 1031
antiapoptotic activities 1486
breast tissue regulation 9–10
transgenic mouse skin tumorigenesis 1089
insulin-like growth factor I receptor (IGF-IR)
antisense oligonucleotide targeting 1443
cell proliferation 965–6
therapeutic strategies 965–6
insulin-like growth factor II (IGF-II) 1031
genomic imprinting 326–7
promoter hypermethylation 326–7
hepatocellular carcinoma 571
insulin-like growth factors 276–7
metastasis 229
insulinomas 587, 595
clinical features 586
Gsα mutations 606–7
k-*rev*-1 mutations 607
liver tumour chemoembolization 1259
Rb mutations 607
int1 76
int2 76
integrin antagonists 1477–8
integrin receptors 26
integrins 5, 984, 1034
activation (inside-out signalling) 211
adherens junctions 215
angiogenesis 980
antibody inhibition 1384, 1477
antiangiogenic therapy 1468–9
basement membrane binding 213,
cell adhesion 230, 939–40, 983
cell migration processes 8, 242–3
clustering 213–14, 985
cytoskeletal connections 214, 243, 985
expression patterns 211
extracellular matrix 210–13, 939–40, 1026
basement membrane-induced mammary epithelium differentiation 216–7
cancer 218–20
cell proliferation stimulation 211–12
fibroblast interactions 1084
focal adhesions 173, 213
haematopoietic cell interactions 1104
hemidesmosomes 9, 216
ligand binding 210–11
malignant invasion 230, 232

organ-specific metastasis 985
matrix metalloproteinase expression regulation 241
RGD motif recognition 211, 242–3
ribozyme targeting 1445
signalling 211–12, 985
α subunit 212
β subunit 211
growth arrest 213
therapeutic approaches 1477–8
subunit structure 210–11, 230
heterodimers 210
targeted drug delivery 1469, 1481
TGF-β regulation of protein expression 181
intercellular adhesion molecules (ICAMs) 558, 985, 1390
malignant invasion 230
intercellular communication
cell growth regulation 25–6
contact-dependent interactions 1031–4
gap junctions 1031–4
juxtacrine signalling 1034
cytotoxic ‘bystander effect’ 1069–70
endothelium 1010
extracellular matrix functions 215–16
preventive therapeutic approaches 430, 441
progression of cancer 869–70
interdigitating cells 690
interferon α (IFN-α) 1410
antitumour immune response 938
haematopoiesis inhibition 1104
signalling pathway 151–2, 1410–11
interferon α (IFN-α) therapy 1371–2, 1377, 1466
immunotherapy following stem cell transplantation relapse 1551
melanoma 488
retinoids combined treatment 1411
toxicity 1377
interferon α/β receptors (IFNRA) 151–2, 1410–11
interferon β (IFN-β) 1410
differentiation induction 1410–11
mezerein combined treatment 1412
signalling pathway 151–2, 1410–11
interferon γ (IFN-γ) 1410
antitumour immune response 938, 1074
antitumour therapy 1377–8
differentiation induction 1410–11
endothelial cell activation 1013
haematopoiesis inhibition 1104
integrin expression inhibition 1477
retinoids combined treatment 1411
signalling pathway 150–1
interferon γ receptor 1 (IFNGR1) 150–1
interferon γ receptor 2 (IFNGR2) 150–1
interferon γ receptors 1410–11
interferons 150
angiogenesis inhibition by gene therapy 1073
differentiation induction 1410–11
retinoid combined treatment 1411, 1415
proliferation inhibition 1410
signalling pathways 150, 1410–11
type I/type II 150
interleukin 1 (IL-1) 8
angiogenesis 239
antitumour immune response 938, 1074
antitumour therapy 1377–8
radiotherapy response 1364–5
interleukin 1α (IL-1α) 1124
etoposide combined treatment 1118
interleukin 1β (IL-1β) 1124
interleukin 2 (IL-2)
antitumour immune response 938, 1074
γ_c receptor subunit 147
properties 147
STATs activation 149
structure 147
interleukin 2 (IL-2) therapy 1371–6
antibody combined therapy 1376
chemotherapy combinations 1376
clinical experience 1375
combined adoptive therapy
lymphokine-activated killer (LAK) cells 1372, 1375
tumour-infiltrating lymphocytes (TIL) 1372, 1375–6
cytokine combined therapy 1376
immunotherapy following stem cell transplantation relapse 1551
renal cell carcinoma 653
toxicity 1375
tumour vaccine studies 1059–60, 1376
interleukin 3 (IL-3) 6–7, 938
interleukin 4 (IL-4) 8
antitumour immune response 938, 1074
antitumour therapy 1378
γ_c receptor subunit 147
haematopoiesis inhibition 1104
in vitro dendritic cells generation 1378
properties 147
STATs activation 149
structure 147
interleukin 5 (IL-5) 7, 938
interleukin 6 (IL-6) 8, 1124
antitumour immune response 938, 1074
antitumour therapy 1378
endothelial cell production 1013
radiotherapy responses 1365
retinoids combined treatment 1411
interleukin 7 (IL-7)
antitumour immune response 938, 1074
γ_c receptor subunit 147
properties 147
STATs activation 149
structure 147
interleukin 8 (IL-8) 239
angiogenesis 239, 971, 1028–9
inhibition by epidermal growth factor receptor (EGFR) blockade 1490
antitumour immune response 938
interleukin 9 (IL-9)
γ_c receptor subunit 147
properties 147
STATs activation 149
structure 147
interleukin 10 (IL-10) 1124
angiogenesis inhibition 1471
antitumour immune response 938
haematopoiesis inhibition 1104
interleukin 12 (IL-12)
antitumour immune response 938, 1074
antitumour therapy 1378–9, 1471
tumour vaccine studies 1059–60
interleukin 13 (IL-13) 938
interleukin 15 (IL-15)
antitumour immune response 938
γ_c receptor subunit 147
properties 147
STATs activation 149
structure 147
interleukin 18 (IL-18) 1074

interleukins
cell growth regulation 25
differentiation regulation 959–60
endothelial cell activation 1014
intermediate filaments 466, 469
integrin β1 subunit interactions 214
internal dose markers 377–9
internal mammary lymph nodes 612
International Conference on the Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 366–72
intersex syndromes 684
interstitial collagenases 231–2
interstitial laser photocoagulation 1261, 1264–5
interstitial thermotherapy 1565
interventional magnetic resonance (IMR) 1563–5
intestinal crypts 7, 545–7
cell proliferation rate 252–3
clonal cell populations 252–3
intestinal metaplasia
clonality analysis 252
diet-related 531
progression to gastric cancer 526–7, 528–9
intracranial pressure elevation, brain tumours 1124–5
treatment-related 1125–6
intradermal angiogenesis assays 974
intraduct papilloma 250
intrahepatic bile ducts 563–4
intraluminal ultrasound 1198
intraocular tumours 826
intratubular germ cell neoplasia, unclassified type (IGCNU) 679–80
intravascular (angiotropic) B cell lymphoma 704
intravital microscopy 973–4
intravital videomicroscopy 1000–1, 1005
intrinsic factor 524
invadopodia 231–2
invasion 230, 461, 999–1000, 1034–6
angiogenesis/neovascularization 229–30
apoptosis abnormalities 228
cell adhesion 230
cell migration 231, 985–6
cell selection processes 225–8
cell surface interactions 232
colorectal cancer 555
extracellular matrix degradation 231–2, 939, 940, 985, 993
host responses 228, 232
in vitro assay 886, 992–3
mathematical modelling 947–50
metastasis 225
physiological 229–30
transforming growth factor-β (TGF-β) activation 180
tumour–host communication interface 232
inverted microscope 877
investigational new drugs (INDs) 1041
FDA application 372–3, 1041–2, 1570
ion channels 175
ionizing radiation 307, 1359
background exposure levels 309
brain tumours 794
in vivo experimental effect 1134–5
induction in animal models 1128
breast cancer models 926–7
carcinogenesis 311–14
animal models 312
dose–response relationship 312
definition 307
DNA damage 310, 1360–2, 1508
dose equivalents 308
dose/activity units 308
electromagnetic 307
energy spectrum 307
exposure risk assessment 314
free radical generation 1359–61
gastric cancer 531–2
leukaemia 265
linear energy transfer (LET) 307, 310, 312
lung cancer 512
particulate 307
protection standards 314
radioisotope decay 308
salivary gland carcinoma 498
skin cancer 478, 481
soft tissue sarcoma 736–7, 783
sources of exposure 309
stress response 310, 1360
thyroid follicular carcinoma 601
see also radiotherapy
IP-10 239
Iressa (ZD1839) 1497, 1500
irinotecan sulphate (CPT-11) 1317–19
carboxylesterase activation 1540
drug combinations 1320, 1500
metabolism 1319
resistance 1319–20
iris 825–6
iris melanoma 832–3, 838
iris tumours 832–3
iron 275, 300
ISG Viewing Wand 1564
islet cell tumours 586–7, 820
clinical features 595
clinical management 595
hormone production 586–7
immunohistochemistry 586
malignancy 586–7
multiple endocrine neoplasia syndrome type 1 (MEN1) 587
poorly differentiated 587
islets of Langerhans 580
isoflavones 424, 439, 449–50
rodent chow diets 920
isorhamnetin 391
ITAMs 172–3
Ito cells 563–4

J

J chain 714
J-107088 1318
JAK1 148, 150–152, 1410–11
JAK2 148, 150–1, 1410–11
JAK3 148
JAK–STAT signalling pathway 148–9, 1410
immunosuppression-related defects 149
interferon α/β (IFNα/β) 151–2, 1410–11
interferon γ (IFN-γ) 150–1, 1410–11
tyrosine kinase signalling 162–4, 171–2
janus kinases (JAKs) 148
jaundice 580
JC virus 71, 794
experimental brain tumour induction 1128–9
jejunum 545
JNK pathway 175, 1047–8, 1364–5
joules (J) 308

jugulotympanic paraganglioma 845–6
treatment 846
jun 607
juvenile laryngeal papillomatosis 507
juvenile polyposis 40, 547
juvenile xanthogranuloma 832, 838
juxtacrine signalling 953, 1034
antigen presentation to T cells 172–3
receptor tyrosine kinases 169, 171

K
K19 495
K-1735 cells 1116
K-*ras* 288, 1525
antisense oligonucleotide targeting 1443
antitumour compound molecular targeting 1044
colorectal cancers 551, 556–8
gene therapy 1525
lung cancer 407–9, 512, 518
O^{6}-*MGMT* promoter region hypermethylation 324
pancreatic cancer 592
use as screening marker 590
pancreatic intraepithelial neoplasia (PanIN) 581–2
pancreatic osteoclast-like giant-cell tumour 583
k-*rev*-1 mutations 607
kaempferol 391, 424
KAI1 1095–6
KAM17.1 590
kanamycin 882–3
Kaposi sarcoma 62, 341, 486–7, 519, 754
central nervous system 810–11
classic (sporadic) 62, 754
clonality of multiple deposits 251
endemic (African) 62, 341, 754
epidemic (AIDS-related) 62, 341, 754
epidemiology 265, 341
human immunodeficiency virus (HIV) 343, 486–7, 676, 810–11
immunohistochemistry 754
immunosuppression-associated (transplant-related) 62, 754
Kaposi sarcoma-associated herpes virus association 65, 737
pathology 486–7
thallium-201 scanning 1221–2
treatment 754
Kaposi sarcoma-associated herpesvirus (KSHV; HHV-8) 49–50, 62–5, 335, 341, 343, 478, 691, 783, 784
Castleman's disease association 65, 341
early genes 64
genome organization 64
historical aspects 62–4, 341
immediate early genes 64–5
Kaposi sarcoma association 65, 486–7, 737
LANA (latency associated antigen; ORF 73) 64–5
late genes 64–5
latent genes 64–5
multiple myeloma association 722–3
ORF K72 65
pathogenesis 65, 341
phylogenetic tree 64
prevalence 341
primary effusion lymphoma association 65, 341, 704
transmission 65, 341
v-FLIP (ORF 71) 65
virion characteristics 64, 341
karyotyping *see* chromosomal abnormalities
KDR *see* vascular endothelial cell growth factor receptor 2 (VEGFRI-2)
keratin pearls 1403
keratinocytes 8
cell lines 1087
epidermal proliferative units 8
terminal differentiation 8, 9
keratins 8, 466–7, 469, 477
angiosarcoma 729
chordoma 727–8
desmosomes 9
mesothelioma 771
salivary gland carcinoma 498
sentinel lymph node biopsy 471–2
squamous cell carcinoma 477, 481
see also cytokeratins
keratoma *see* cholesteatoma
keratosis, larynx 507–8
Kerley B lines 504–5
Ki-67 426, 495, 958, 255
astrocytoma 800–1
brain tumours 795
breast cancer 615
glioblastoma 802
lymphoma 691
primitive neuroectodermal tumours 798
soft tissue sarcomas 745, 760
Ki-*ras* activation 32–3
kidney
anatomy 645
development 645
multidrug resistance proteins (MRPs) 1352
Pgp expression 1349
tumours 646–53
Kiel classification 693
KIP/Cip Cdk inhibitors 1364, 1513
KM12 1091
knock-in mouse studies 905
knockout mouse studies 865–6, 903–10
angiopoietins 244
apoptotic protease activating factor-1 (Apaf-1) 127
applications 905–9
brain tumours 1131–3
Brca1 897–8, 1094
breast cancer 1094
caspase-9 127
chemopreventive agent testing 425, 445, 1426
conditional knockouts 894, 897–8
costs 910
Cre-loxP gene targeting system 905, 910
embryonic mesenchymal–epithelial interactions 1027
epidermal growth factor receptor (EGFR) 1487
Fas-ligand (FasL) 1003
gene therapy models 1523
hus1 103
knockout generation procedure 895, 903–5
blastocyst microinjection 904
chimaeric mice generation 905
selection procedures 903–4
targeting vector 903–4
limitations 901, 909–10
mdr-1 1044, 1046–7
medulloblastoma 909
metastasis models 1003
MRP1 1352
MSH-2 95
multidrug resistance (MDR) 1044–5
multiple mutation targeting 910
Nf1 1131

knockout mouse studies – *continued*
NQOI 1090
oestrogen receptorα 136, 1451
oestrogen receptorβ 136, 1451
oncogenes/tumour-suppressor genes 891–902
p27SKIP1s 908
p53 395, 905–7, 1131, 1364–7
P-glycoprotein 1044, 1337–8
PML (promyelocytic leukaemia gene) 909
Ptc 1131
PTEN 907–8
radiotherapy response pathways 1366
Rb 908–9, 1131
squamous cell carcinoma 1089–90
STATs 149
vascular endothelial cell growth factor (VEGF) 238
vascular endothelial cell growth factor receptors 239
XPA 300, 368–9
Knudson's two-hit hypothesis 26–7, 31, 593
see also multistep carcinogenesis
koilocytosis 54
Kostmann's syndrome 1545
KP-1 583
Krause glands 826, 827
KRIB 1118
Krukenberg tumours 533, 640
Ku 97, 1361–2
WRN interaction 102
Kupffer cells 563, 564

L
9L gliosarcoma 1130
L1210 lymphocytic leukaemia model 915, 1043, 1130
L5178Y mouse lymphoma assay 354
L778123 1498–9
L-selectin 1016–17
malignant invasion 230
labyrinth
bony 841–2
membraneous 841–2
Lac-Z reporter gene
knock-in mouse studies 905
metastasis *in vivo* models 1003–4
Muta Mouse genotoxicity test system 357–8
LacI 379
transgenic genotoxicity studies 357–8
lacrimal gland tumours 835–6
clinical management 838–9
prognosis 837
lacrimal glands 826
lacrimal sac tumours 826, 835–6, 838–9
lactation 9
lactoperoxidase 278
Lambert–Eaton myasthenic syndrome 516
laminar flow clean benches 873–5
laminarin 1477
laminin 8, 210
angiogenesis induction 971
antineoplastic synthetic peptides 1476
basement membrane 210, 236, 987, 1010–11
extracellular matrix 5, 214–15, 216
basement membrane-induced mammary epithelium differentiation 216–17
degradation during malignant invasion 231
functional domains therapeutic potential 1475–6
integrins binding 210–11, 1026
phosphorylation
mitosis initiation 1324
mitotic arrest 1325
laminin B2 1325
laminin-1 210
laminin-2 210
laminin-5 210–11
Land and Weinberg model 31–2
Langerhans cell histiocytosis 730
central nervous system 812
orbit 835
Langerhans cells 8, 477, 1389–90
sunlight exposure response 481–2
lanreotide 609
laparoscopic surgery 1559–61
advantages 1560
colorectal cancer 1559–60
costs 1560
disadvantages 1560–1
hand-assist 1561
oncological principles 1559
pneumoperitoneum 1561
port-site recurrence 1560–1
laparoscopic ultrasound 1559
laparoscopy 1559–62
staging 1559
large scale cell culture 884–5
large-cell lung carcinoma 510, 518
laryngeal carcinoma
adenocarcinoma 508
chemically induced tumour models 1092
occupational exposure-related 417
squamous cell carcinoma 508
genetic changes 508
laryngeal keratosis/dysplasia 507–8
laryngeal papillomas 507
laryngeal tumours 507–8
haemangioma 508
lymphoma 508
metastatic 508
neural 508
polyclonal processes 250
soft tissue 508
larynx 503–4
laser imaging fluorescence endoscopy (LIFE) 513
lead candidate compounds 414, 1041, 1043
leather dust 505
Ledderhose disease (plantar fibromatosis) 746
Lef-1 204
LEF/TCF transcription factors 185
β-catenin binding 203
lefluonomide (SU101) 1468, 1500–1
leiomyoma
benign metastasizing 734
bladder 654
imaging 1149
myometrium 629, 634
nasopharynx 503
leiomyosarcoma 469, 751–2
bladder 654
cutaneous 752
Epstein–Barr virus 340–1
human immunodeficiency virus (HIV) 343
immunohistochemistry 745, 751–2
intra-abdominal 752
intramuscular 752

larynx 508
myometrium 629–30
nose 503
pancreas 587
penis 676
seminal vesicle 678
subcutaneous 752
vascular 752
letrozole 1457
Letterer–Siwe disease 730
Leu-7 *see* CD57
Leu-M1 657
leucocyte common antigen (LCA) 481
leucocytes 1101
leucotriene C_4 1352
leucovorin (folinic acid) 559, 1288–9, 1291
5-fluorouracil modulated treatment 1273
leukaemia 465
acute childhood 262, 265–6
electric/magnetic field exposure association 313
aetiology
infectious agents 347
ionizing radiation 311–12
antisense therapy 1445–6
developmental field effects 1403
epidemiology 1101
Fas (CD95) somatic mutations 123
historical aspects 693
JAK–STAT signalling defects 149
MDR1 expression 1351
multidrug resistance proteins (MRPs) expression 1353–4
pancreatic involvement 587
Philadelphia chromosome positive, transgenic studies 1110–11
stem cell transplantation 1545, 1546–7
testicular involvement 683, 1193
viruses 76
leukaemia, acute lymphocytic (ALL) 1101, 1382–3
allogenic stem cell transplantation 1546–7
molecular classification 470–1
p15 promoter hypermethylation 323–4
transgenic studies 1110–11
leukaemia, acute myeloid (AML)
allogenic stem cell transplantation 1546–7
immunotoxin therapy 1386
interleukin 2 (IL-2) therapy 1375
molecular classification 470–1
p15 promoter hypermethylation 323–4
telomerase activity 116
tobacco association 402
leukaemia, acute promyelocytic (APL)
all-*trans*-retinoic acid/high dose retinoic acid treatment 473, 1106–7 1112, 1414
differentiation induction mechanism 1414–15
resistance 1112, 1414
retinoic acid syndrome 1414
PML knockout mouse studies 909
leukaemia, adult T cell see adult T cell leukaemia/lymphoma
leukaemia, chronic lymphocytic (CLL) 1382–3
allogenic stem cell transplantation 1546–7
B type (small lymphocytic lymphoma) 696–7
genetic features 696–7
immunophenotype 696
transformation 697
immunotoxin therapy 1386
liver involvement 1101
pretherapeutic cellular DNA repair profiling 1511–12
leukaemia, chronic myelogenous (CML)
allogenic stem cell transplantation 1546–7
Philadelphia chromosome 174
Abl protein aberrant cellular localization 33, 1106
antisense oligonucleotide targeting 1443
transgenic studies 1110–11
leukaemia, precursor B cell lymphoblastic/acute lymphoblastic (precursor B cell lymphoblastic lymphoma) 695–6
leukaemia, precursor T cell lymphoblastic/acute lymphoblastic (precursor T cell lymphoblastic lymphoma) 706
leukaemia, prolymphocytic B cell 697, 1101
leukaemia, prolymphocytic T cell (T-PLL) 706
leukaemia, T cell acute lymphocytic 1101
HOX11 aberrant expression 1107
leukaemia, T cell large granular lymphocytic 707
Lewis lung carcinoma 1092
Leydig cell tumour
immunohistochemistry 683
testis 682
Leydig cells 678
LFA-3 1390
Lhermitte–Duclos disease (cerebellar dysplastic gangliocytoma) 806, 820, 907
Li–Fraumeni syndrome 19, 33, 41–2, 636
lung cancer 511–12, 514
osteosarcoma 42, 722, 724
p53 knockout mouse studies 905–7
p53 mutations 19, 41–42, 896
soft tissue sarcomas 737
lifestyle-related cancer risk 263, 421
light-scattering spectroscopy (LSS) 1420
lignans 424
limit dose 371
limonene 410, 424, 428, 430, 440
linamarase–linamarin prodrug activation system 1540–1
bystander effect 1540–1
toxicity 1541
linoleic acid 395
lip 491
lip cancer
epidemiology 491
pathology 491
ultraviolet exposure association 493
lipid peroxidation 302–3
lipid-associated sialic acid 636
Lipiodol (Ethiodol) embolization 1256, 1257–9
lipoma
cytogenetic alterations 469–70
imaging 743–4
magnetic resonance imaging (MRI) 1210
lipopolysaccharide 152
endothelial cell activation 1014
liposarcoma 733, 743, 749–51
cytogenetic alterations 469–70
dedifferentiated 750–1
heart 782
imaging 743–4, 1210
larynx 508
myxoid 750
pancreas 587
pleomorphic 751
round cell 750
seminal vesicle 678
well-differentiated 750
liposome vectors 1523, 1528
antisense oligonucleotide delivery 1442–3
ribozyme delivery 1443
lipoxygenase (LOX) inhibitors 1426, 1431

liver
  anatomy 563–4
    lobes 563
    segments 563
  blood supply 563
  cell generation/regeneration 10
  development 563–4
  functional unit (acinus) 563
  haematological malignancies 1101
  microscopic structure 563–4
    'streaming liver' concept 564
  multidrug resistance proteins (MRPs) 1352
  Pgp expression 1349
  stem cells 10
liver flukes 347
  cholangiocarcinoma association 347, 573–4
    precancerous changes 574
  life cycle 347, 573–4
  treatment 347
liver metastases 229, 229, 575, 999–1000, 1001–2
  chemoembolization 1259
  computed tomography (CT) 1165–6
  interstitial thermotherapy 1565
  magnetic resonance imaging (MRI) 1205
  neuroendocrine tumours 609
  percutaneous therapy 1255
    ablative 1261
    microwave coagulation 1265
  rodent models 1002
  tumour–host interactions 1063–5
    cell isolation techniques 1064
    flow cytometric analysis 1065
    gene expression studies 1065
  ultrasound 1196–7
liver tumours 564
  2-acetylaminofluorene carcinogenesis 282
  benign 576
  chemoembolization 1256
    alcohol injection 1256
    chemotherapeutic/embolic agents 1256
    contraindications 1259
    indications 1256–9
    mode of action 1256
    neuroendocrine tumours 1259
  classification 564
  computed tomography (CT) 1161
  epidemiology 564
  magnetic resonance imaging (MRI) 1205–6
  percutaneous ablative therapy 1261–2
  percutaneous biopsy 1247, 1249–50
  schistosomiasis 345
  telomerase activity 113
  ultrasound 1195–7
    contrast imaging 1190
LM609 242
LMB-2 1386
lobular carcinoma *in situ* 614
logarithmic-phase growth 888
lomustine (CCNU) 1300–1
loop and electrocautery procedure (LEEP) 627
loss of heterozygosity (LOH) 88, 1522
Louis–Barr syndrome *see* ataxia telangiectasia
low molecular weight heparins 1477
lower oesophageal sphincter 523–4
lpr/lpr mouse 123
LS174T 1091
lucifer yellow 1031–2
lung
  anatomy 504–5
  lymphatic drainage 504–5
lung cancer 511–19, 1091
  adenocarcinoma 401–2, 511, 517
    treatment/prognosis 517
  adenosquamous carcinoma 518
  aetiology 511–12
  Bcl-2 overexpression 131
  bronchoalveolar carcinoma (alveolar carcinoma) 517
  bronchoscopic dysplasia treatment 1431
  chemoprevention 1092, 1431–2
  chromosomal abnormalities 512
  clinical presentation 515
  computed tomography (CT)
    needle biopsy guidance 1166, 1168
    staging 1162–5
  drug resistance 1351
  epidemiology 264–5, 266, 401–2, 402–3, 511
  erbB-2/neu expression abnormalities 512, 966
  Fas (CD95) somatic mutations 123
  ^{18}FDG positron emission tomography (PET) 1228
  genetic predisposition 511–2
  histological heterogeneity 514
  human papillomavirus (HPV) 337, 510
  hypertrophic pulmonary osteoarthropathy 1152
  interleukin 2 (IL-2) therapy 1375
  large-cell carcinoma 518
  magnetic resonance imaging (MRI) 1205
  metachronous tumours 514–5
  metastases 515
    adrenal 1165
    brain 1125
    liver 575, 1196–7
    middle ear 849
    vagus nerve entrapment 503
  models 1091–2
    chemically induced tumours 1092
    in vitro 1092
    Lewis lung carcinoma 1092
    xenografts 1092
  molecular genetics 1314–15
  multidrug resistance proteins (MRPs) expression 1353, 1354
  neurological syndromes 516
  occupational exposure-related 311–12, 417, 419, 512
    spiral CT/biomarker screening 419
  oxidative DNA damage 89
  p53 gene therapy 1071, 1527
  p53 mutations 29, 299, 511–12, 518
  Pancoast tumours 515
  paraneoplastic syndromes 515
  percutaneous ablative therapy 1261
  pleomorphic carcinoma 518–19
  premalignant lesions 512–14
    atypical alveolar hyperplasia 513–14
    diffuse idiopathic neuroendocrine hyperplasia (DIPNECH) 514
    preinvasive squamous lesions 513
    tumourlets 514
  Rb mutations 511–12, 908
  screening
    computed tomography (CT) 1150, 1158, 1159
    Early Lung Cancer Action Project (ELCAP) 1158
    plain radiography 1150, 1153
  squamous cell carcinoma 401–2, 516–17
    basaloid 516
    prognosis 517
    treatment 516

synchronous tumours 514
telomerase activity 113–15
thoracoscopic resection 1563
tobacco smoke association 33, 400–1, 401–2, 407, 511
duration of exposure 384, 401
mechanism of tumour induction 407–10
passive smoking/environmental exposure 260, 402–3, 511
risk assessment from DNA adducts 303
see also small-cell lung carcinoma
lung carcinoma *in situ* 513
telomerase activity 113
lung lymphoepithelial carcinoma 340–1, 512
lung metastases 228–9, 519, 999, 1001
in vivo models 1000–1
rodent models 1002
lung tumours 509–19
benign 509–10
classification 509, 515
computed tomography (CT) 1161
glutathione-*S*-transferase (GST) isoform expression 1355
lymphomas 519
malignant see lung cancer
neuroendocrine 510–1
percutaneous biopsy 1247–9
pneumothorax 1248–9
sarcomas 519
secondary *see* lung metastases
lung-derived growth factor 1012
luteinizing hormone-releasing hormone (LHRH) 1455
luteinizing hormone-releasing hormone (LHRH) agonists 1454–5
LY335979 1350
lycopene 672–3
lyl-1 1107
lymph nodes
anatomical structure 689–90
mantle zone 689–90
marginal zone 689–90
medulla 690
paracortex 690
primary follicles 689
secondary follicles 689
biopsy
CT-guided needle 1166–8
percutaneous 1247
functional aspects 690
imagimg 1149
immunohistochemistry 690
metastases 465
magnetic resonance imaging (MRI) 1212
sentinel lymph node biopsy 471
tumour staging 471
reactive patterns 690–1
lymphangioma
bone 723–4
orbit 834, 838
lymphangiosarcoma 737
lymphatic spread 999–1000
lymphocytes
activation 1389
cell culture
diploid stem cells 1109
isolation from blood 1108–9
permanent cell lines 1109
short-term for genetic analysis 1109
immune response 1056
stromal cells 1025
trafficking (homing) 1056
lymphocytic and histiocytic cells 713
lymphoepithelial carcinoma 340–1, 512
lymphoepithelioma 60–1, 340–1, 498, 512
lymphoid hyperplasia, benign
lachrymal gland 836
orbit 835
lymphoid infiltration 937–8
lymphoid interstitial pneumonia 519
lymphoid stem cells 6
lymphoid tissue
intestinal tract 548
nasopharynx 503
lymphokine-activated killer (LAK) cells, interleukin 2 (IL-2) combined therapy 1372, 1375
lymphoma 465, 691–706
anaplastic large cell 710–11
B cell lineage *see* B cell lymphoma
bone 730
central nervous system 792–794, 812, 836–838
cIAP-2 expression 122–3
classification 693–6
historical aspects 693–4
Kiel 693–4
REAL 467, 694, 812
WHO 694
Working Formulation 693–4
clonality analysis 473
cytokine-mediated immunotherapy 1551–2
donor lymphocyte infusion therapy 1551
epidemiology 1101
Epstein–Barr virus 60–1, 339–40, 1104–5
see also Burkitt's lymphoma
Fas (CD95) somatic mutations 123
^{18}FDG positron emission tomography (PET) 1228
follicular *see* follicular lymphoma
gallium-67 citrate scanning 1222
genotypic studies 473, 692–3
clonality analysis 692
karyotyping 692–3
heart 778, 782, 784–6
human immunodeficiency virus (HIV)/AIDS patients *see* human immunodeficiency virus (HIV)
imaging 1149
immunohistochemistry 467
immunophenotyping 467, 691
infectious aetiology 347–9
JAK–STAT signalling defects 149
lachrymal gland 836
liver 1196–7, 576, 1101
lung 519
lymphoplasmacytic 702
magnetic resonance imaging (MRI) 1212
mantle cell *see* mantle cell lymphoma
mediastinum 1249
ocular adnexa 836, 838
ocular tissues 836–8
orbit 835
pancreas 587
penis 676
prognostic factors 694–695
genetic features 694–5
grade 694–5
immunohistochemistry 694–5
radiotherapy sensitivity 1359
retina 831
rituximab therapy 1382
small lymphocytic *see* small lymphocytic lymphoma

lymphoma – *continued*
spleen 1250
stem cell transplantation 1545–6
stomach 548
T cell lineage *see* T cell lymphoma
testis 683, 1193
thyroid 601–2
tumour markers 472–3
tumour-suppressor gene knockout mouse studies 899
see also Hodgkin's disease; non-Hodgkin's lymphoma
lymphoplasmacytic lymphoma 702
clinical features 702
t(9;14)(p13;q32) 702
lymphoproliferative disorders
atypical polymorphous 718
iatrogenic non-transplantation-induced 718
immunocompromised patient 716–19
post-transplantation 716–18
sporadic atypical 719
lymphotoxin-α see tumour necrosis factor β (TNF-β)
lymphotoxinβ 152–3
lynoestrenol 921
lysozyme 7, 504

M

Mac-1 (CD11b/CD18) 1441
Mac-387 583
McCune–Albright syndrome 641, 723
Mcl-1 129
macrophage colony-stimulating factor (M-CSF) 1103
macrophages
bone marrow precursor cells 6
isolation from blood 1108–9
lymph nodes 690
follicles (tingible body macrophages) 689
stromal cells 1025
TGF-β response 181
tumour angiogenesis 239–40
Mad2 1517
Mafucci syndrome 722
MAGE 1 1392
MAGE 3 1392
MAGI-2 215–16
magnetic fields 266, 307, 308
activity units 308–9
biological effects 311
carcinogenesis studies 313–14
brain tumours 1128
exposure limits 314
sources of exposure 309
magnetic resonance cholangiopancreatography (MRCP) 1206
magnetic resonance endoscopy 1558–9
magnetic resonance imaging (MRI) 1147, 1201–12, 1558
adrenal tumours 1207
astrocytoma 800–1
biliary tumours 1206
bladder cancer 1209–10
bone tumours 1211–12
breast lesions 1209, 1238–41
cancer staging 1241
nipple discharge 1242
central nervous system tumours 792, 1202–3,
cervical cancer 1208
colorectal cancer 1208
contraindications 1202
endometrial cancer 1208
endoscopic 1558–9
head and neck tumours 1203–4
liver tumours 1205–6
lung cancer 1205
lymph node imaging 1212
metastatic tumours 1149
ovarian cancer 1208–9
pancreatic tumours 589, 1206
percutaneous ablative therapy guidance 1262
principle 1201–2
prostate cancer 1210
renal tumours 1206
soft tissue tumours 743, 1210
surgical interventional guidance 1565
testicular cancer 1210
MAGUK 214–6
MAID chemotherapy 762
major histocompatibility complex (MHC) molecules 1056–7, 1389–90
dendritic cell expression 1391
T cell receptor interactions 172
malaria 264
male genital tract tumours 113, 665–85
malignant fibrous fibrosarcoma 728
malignant fibrous histiocytoma 722–3, 728, 736–7, 749
larynx 508
pancreas 587
penis 676
seminal vesicle 678
malignant melanoma *see* melanoma
malignant peripheral nerve sheath tumour 38, 466, 756, 792
cytogenetic studies 756
electron microscopy 746
heart 782
imaging 743–4
immunohistochemistry 745
liver 743
pathology 817
malignant tumours 461
malleus 841
malnutrition 492–3
MALT lymphoma (MALToma) 548
central nervous system 812
marginal zone B cell lymphoma 700–1
t(11;18)(q21;q21) 700–1
small intestine (α heavy chain disease; immunoproliferative small intestinal disease) 548
mammary gland *see* breast
mammography 1235
adjuvant ultrasound 1238–41
calcifications/microcalcifications 1236, 1238, 1241
diagnostic 1236–41
digital 1235–43
screening 1235–7, 1243
efficacy 1235
elderly women 1236
false negative 1235
interpretation 1235
randomized controlled trials 1235–6
Mammotome 1564
Manihot esculenta 1540
mannitol 1134
mantle cell lymphoma 697–8
immunophenotype 698
t(11;14)(q13;q32) translocation 692, 698
MAPK (mitogen-activated protein kinase) 204, 1410, 1486
cell growth signalling 15

extracellular matrix signalling 213
integrins 212–14
JNK cascade/TGF-β signalling convergence
non-Smad signalling 186
Smad signalling 185–6
oestrogen receptor phosphorylation 141–2, 1451, 1453
radiotherapy response pathways 1364–6
see also Ras-MAP kinase
marginal zone B cell lymphoma 700–2
extranodal of MALT type 700
nodal 701–2
splenic 701
margins examination 465
marimastat (BB-25160) 1470, 1479
marital status 263
MART-1 1075
MART-q/Melan-A 1392
masquerade syndrome 828
mass spectrometry 299
mast cells 6
mathematical modelling 937–50
angiogenesis 945–7, 950
immune response to cancer 941–2, 950
invasion/metastasis 947–50
principles 940–1
Matrigel
invasion assays 992
plug assay 976–7
stromal–epithelial interactions 1030–1
tubulogenesis model 977–9
matrilysin 215
matrix implant angiogenesis models 975–7
cannulated sponge model 976
disc angiogenesis system 976
Matrigel plug assay 976–7
matrix metalloproteinase 1 (MMP-1) 241
matrix metalloproteinase 2 (MMP-2) 231–2, 241
malignant invasion 232, 1034–6
therapeutic inhibition 1470, 1479
transforming growth factor-β (TGF-β) activation 180, 191
matrix metalloproteinase 3 (MMP-3), transgenic mouse studies 220
matrix metalloproteinase 9 (MMP-9) 231–2, 241, 1088
malignant invasion 1034–6
ribozyme targeting 1445
therapeutic inhibition 1470, 1479
transforming growth factor-β (TGF-β) activation 180, 191
matrix metalloproteinase 10 (MMP-10) 1088
matrix metalloproteinases (MMPs) 231, 626–7, 987, 1030–1, 1083
angiogenesis 231–2, 241, 1470
gelatinases 231–2
interstitial collagenases 231–2
malignant invasion/metastasis 231, 940, 985, 999–1000, 1034–6, 1088
stromelysins 231–2
therapeutic inhibition 1470, 1478–80
transforming growth factor-β (TGF-β) activation 180, 191
maturation promoting factor (MPF) 14
mature cystic teratoma *see* dermoid cyst
maximum feasible dose 371
maximum tolerated dose (MTD) 370, 1570
rodent carcinogenicity studies 365
rodent toxicological studies 921
maytansine 1329
MBD2 320
MCC (mutated in colon cancer) 29, 1090
central nervous system tumours 1123–4
familial adenomatous polyposis (FAP) 37
gastric cancer 250–1
MCF-7 cells 1029–31, 1082, 1094
MCL5 363
MCP-1 239
MDA cell lines 1094
MDM2 (Mdm2) 99–100, 1123–4, 1307, 1363
p53 regulation 1516, 1518, 1526
skin carcinogenesis 1087–8
MDM2 (*mdm2*) 750, 896, 905–7, 1087–8
amplification
osteosarcoma 724
soft tissue sarcomas 740, 745, 760
antitumour compound molecular targeting 1044
MDR1 1341, 1348
expression in disease 1046–7
expression in normal tissues 1046–7, 1349
expression in tumours 1350–1
gene product-related multidrug resistance 1046
gene transfer into haematopoietic stem cells 1076
species differences 1348
therapeutic applications 1047
transgenic mouse studies 1047
mdr-1
antisense targeting 1441
antitumour compound molecular targeting 1044
knockout mouse studies 1044, 1046–7
ribozyme targeting 1445
mean graph 1043
meat 377–8, 392, 549
mechanistic chemopreventive efficacy assays 444–5
mechlorethamine (nitrogen mustard) 89, 1295–6
Meckel's cartilage (first branchial arch) 841
MeCP1 320
MeCP2 320–1
median-effect plots 1051–2
mediastinal germ cell tumours 777
mediastinal lymph node biopsy 1166–8
mediastinal mass 1248–9
mediastinal (thymic) lymphoma 704
medroxyprogesterone acetate 921
medullary pancreatic carcinoma 582–3
medullary thyroid carcinoma 601–2
carcinoembryonic antigen (CEA) 608
epidemiology 600
hepatic metastases 1196–7
multiple endocrine neoplasia 2 (MEN-2) 601
prognostic factors 608
ret mutations 601, 606
medulloblastoma 39, 794–8
cell lines 1126
chromosomal abnormalities 820
desmoplastic 798
experimental induction 1128–9
mouse knockout model 909
xenografts 1127
medulloepithelioma 795
megakaryocytes 1103
meibomian glands 826
Mel-A 466–7
melanin pigment 477
photoprotective effect 478, 481
melanocytes 8, 477
melanoma 466
borderline 488
brain 792, 1127–8
cellular vaccines 1391, 1393–4, 1435
tumour-specific peptides 1392
DNA vaccine clinical trials 1399

melanoma – *continued*
drug resistance 1351
eyelids/conjunctiva 827
familial predisposition (heritable melanoma syndrome) 479, 487
^{18}FDG positron emission tomography (PET) 1228–9
gallium-67 citrate scanning 1223
immunohistochemistry 466–7
larynx 508
malignant of soft parts *see* clear cell sarcoma
metastases 1001, 1115
brain 1001, 1125
chemoembolization 1259
interleukin 2 (IL-2) therapy 1375
liver 1196–7
middle ear 849
models 1127–8
animal brain tumours 1115–17
cell lines 1116
chemical carcinogens/UV radiation 1116–17
human artificial skin reconstructs 1116–17
human/mouse xenografts 1116
mouse 1116
transgenic 1116
multidrug resistance proteins (MRPs) expression 1354
nasal cavity/paranasal sinuses 506
orbit 835
penis 676
premalignant lesions 1115
clonality analysis 250
dysplastic naevi (BK moles) 479–80
prevention 479
progression 1115
experimental studies 1115
sentinel node localization 1232–3, 1563
skin
acral/mucosal lentiginous 485
chemoprevention 1435
clinical features 485
clinical management 488
lentigo maligna 485
metastasis 485, 488
pathology 485–6, 1115
prognosis 488
radial/vertical growth 485, 488, 1115–16
superficial spreading 485
stem cell transplantation 1546–7
sunlight exposure association 33, 313, 478–9, 485, 1115
testis 683
uvea *see* uveal melanoma
melanosis
congenital 826–7
primary acquired, eyelids 826, 828, 837
melanotic neuroectodermal tumour of infancy 506
melphalan 301, 1296
carrier system 1298
resistance 1356
structure 1296
membrane type 1 matrix metalloproteinase (MT1-MMP) 232, 241
menin mutations 601, 607
meningioangiomatosis 820
meningioma 792, 794
aggressive behaviour 810
anaplastic 810
atypical 810
clonality analysis 251
electron microscopy 810
epidemiology 792–4
histological subtypes 810
immunohistochemistry 810, 847
middle ear 847
optic nerve 835, 838
pathology 808–10
telomerase activity 116
temporal bone 847
menopause 9
menstrual cycle breast changes 9
MER-25 1451
mercaptopurine 1284–7
allopurinol interaction 1287
mechanism of cytotoxicity 1287
metabolism 1284–7
resistance 1287
Merkel cell tumour 827–8
Merkel cells 8
mesenchymal cells 733
stem cells 10
see also epithelial to mesenchymal transdifferentiation
mesenchymal chondrosarcoma 726, 745, 757
mesenchymal tissue *see* connective tissue
mesenchymal–epithelial interactions 1083–4
embryonic organogenesis 1026–8
mesoderm 1081
mesothelial cell development 767–8
mesothelial reactive hyperplasia 769
mesothelioma 767–74
aetiology 768–9
asbestos association 413, 417, 767–70
fibre dimensions 275–6
clinical features 769–70
clinical management 774
electron microscopy 771
epidemiology 768
erionite dust association 767–8
genetic prodrug activation therapy 1536
histopathology 770
immunohistochemistry 770–1
molecular genetics 771–3
occupational exposure-related 417
pericardium 778, 782–5
prevention 769
prgnostic factors 773
risk factors 767
SV40 association 767–9, 771
testis 683
Met (hepatocyte growth factor receptor) 1486
MET mutations, familial papillary renal cell carcinoma 649–50, 652–3
meta-analysis 268
clinical drug trials 1573
metachromatic leukodystrophy 1545
metaphase arrest technique 958
metaphysis 721–2
metaplasia 1085
clonality of intestinal lesions 252
metastases
adoptive cellular immunotherapy (ADI) 1062–3
chemoembolization 1259
immunotherapy 1055
plain radiography 1149
staging laparoscopy 1559
metastasis 26, 225, 461, 474, 938–9, 999–1006, 1034–7, 1055–6, 1086
angiogenesis 225, 971–2, 999–1000, 1086
apoptosis defects 119, 1055–6
cell adhesion 983, 1055
cell selection processes 225–8, 1056

experimental models 1001
cell–cell adhesion loss 984–5
distribution 228–9, 1001
efficiency 1000–1
quantitation 1003–4
endothelial cell–tumour cell interactions 1009
epithelial to mesenchymal transdifferentiation 1083
TGF-β induction 190
gastrointestinal tumour models 1091
gene expression changes 868–9
heterogeneity in cellular properties 1001, 1004, 1055
host factors 1111
surveillance mechanisms 869
in vivo assay 886
in vivo models 864–5
establishment 1002–4
experimental 1002–4
inoculation sites 918
interpretation of results 1004–5
knockout mouse studies 1003
limitations 1005
spontaneous tumours 1002
technical factors affecting assays 1005–6
transgenic mouse studies 1003
xenografts 920, 1003, 1005
intravital videomicroscopy 1000–1, 1005
mathematical modelling 947–50
oncogene expression changes 868
ribozyme therapy 1445
signalling pathways 1055–6
'soil and seed' hypothesis 1001–3, 1009, 1036, 1485–6
specific organ preferences 236, 983, 985, 987–8, 1001–2, 1009–2, 1036, 1485–6
stages 938–9, 983, 999–1000, 1034–6, 1485
suppressor genes 868–9
tumour cell lines 1056–7
tumour cell phenotypic characteristics 1000–1
metastat (COL-3) 1480
METH-1/ADAMTS1 1482
METH-2/ADAMTS8 1482
methallothionein 570–1
methanesulfonate (MMS) 295
methotrexate 660, 718, 1077, 1289–91
mechanism of cytotoxicity 1290
resistance 1290–1, 1347
side effects 1151–2
structure 1289–90
transport 1288–9
methyl-CCNU (semustine) 1300–1
methylazoxymethanol 390–1
methylcholanthrenes (MCA) 273
L1210 leukaemia induction 1130
neurocarcinogenesis 1129
murine tumours 1129–30
rodent breast cancer models 925–6, 934
methylene diphosphonate (MDP) 1216–17
4, 4′-methylenebis-2-chloroaniline (MOCA) 301
methylenetetrahydrofolate reductase (MTHFR) polymorphism 549
N-methylnitrosourea (MNU) 295
neurocarcinogenesis 1130
rodent breast cancer models 92–7, 929–30, 933–4, 1426
methylprednisolone 1548
5-methyltetrahydrofolate 1287–9
mezerein 1412
MG-63 1117–18
MGMT (O^6-alkylguanine-DNA alkyltransferase) 1076–7, 1509
chemotherapeutic drug resistance 1510
therapeutic inhibition 1512
MGMT 1511
gene transfer into haematopoietic stem cells 1076–7, 1512
promoter hypermethylation 324
MIBG (meta-iodobenzylguanidine) scanning 1226–7
accuracy 1226–7
imaging protocol 1226
indications 1226
principles 1215
therapeutic potential 1227
MIBI1 *see* Ki-67
MIC2 gene product
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
synovial sarcoma 758
microarray analysis
DNA methylation profiles 329
gene expression 900–1
haematological malignancy classification 470–1
skin tumorigenesis 1089
microcarrier bead culture 884–5
microfilaments 214
microglia cells 789–90
microinvasive carcinoma 225
micrometastases 474–5
active specific immunotherapy (ASI) 1061
detection 1058–9
magnetic resonance imaging (MRI) 1059
micronucleus test
in vitro 358–60
rodent bone marrow 355
microsatellite instability 94, 900
colon cancer 96, 324
hereditary non-polyposis colorectal cancer (HNPCC) 96, 324, 550
hMLH1 disruption 324, 556–7
pancreatic medullary carcinoma 582–3, 593–4
PCR assay 900
polyclonality marker 248
preneoplastic lesions 249–50
tumour mismatch repair (MMR) mutator phenotype 94
tumours 88
microtubule-active drugs *see* antimicrotubule agents
microtubule-associated proteins (MAPs) 1323
microtubules 1323
microwave radiation 308–9
carcinogenesis studies 313
dose/activity units 308–9
exposure limits 314
genotoxicity 310–11
percutaneous coagulation therapy 1261, 1265
sources of exposure 309
middle ear (tympanic cavity) 841
middle ear tumours
adenocarcinoma 848–9
adenoma 844–5
benign 844–7
metastatic 849
squamous cell carcinoma 848
midgut 545
Midostaurin (CGP 41251) 1501–2
migrant population studies 264
Min mouse model 1091
adenoma clonality studies 252
chemopreventive agent testing 1426
Min (Multiple Intestinal Neoplasia) 1091

mineral wools 414
minerals, dietary 423–4
minimal access surgery 1557–8
  robotic surgical systems 1565–6
Minimal Essential Medium (Minimal Essential Medium Eagle; MEM) 880
minimal residual disease (MRD) 474–5
minocycline 1480
minor salivary glands 491
mismatch repair 94–5, 280, 1509–10
  apoptosis induction 95
  colon cancer 96, 550, 556
  defects 436, 1510
    apoptosis avoidance 95
    chemotherapy resistance 95
    microsatellite instability 88, 550
  genes 30, 293
  insertion/deletion loops 94–5
  pancreatic cancer 591–4
  single base-pair mismatches 94–5
  TGF β-RII 96
misoprostol 449
mithramycin 966–7
mitochondria
  caspase activation pathways (apoptosis) 126–8, 130
  cytochrome *c* release 126, 130
    Bcl-2 protein regulation 128, 130
    necrosis (caspase-independent death)-related 130
  DNA damage 94
  DNA repair 94
mitomycin C 301, 660, 1298, 1333, 1341
  clinical activity 1341–2
  combination drug therapy 1342
  liver tumour chemoembolization 1256
  mechanism of action 1341
  metabolic activation 1298–9
  pharmacology 1342–3
  resistance 1343–4
  toxicity 1343
mitomycins 1341–4
mitosis 13–14
  anaphase 1324
  initiation 1325
  metaphase 1324
  microtubule function 1323
  regulation 17, 1324
mitotic arrest 1324–5
  antimicrotubule drugs 1324–5, 1327–8
  cell death 1325–6
  chemotherapeutic approaches 21
  proteins phosphorylated on serine 1325
mitotic catastrophe 21–2
mitotic count 957–9
mitotic index 957–9
mitotic (M-phase) checkpoint 18, 1517
  centrosome abnormalities 88
  Tax protein effects 82, 83
mitotic slippage 1324–7
mitotic spindle 14, 361, 1323–4
  assembly 1324, 1517
mitoxantrone 1333, 1339
  clinical activity 1339
  mechanism of action 1315–16
  resistance 1339
  structure 1339
  toxicity 1339
  transport 1045
mixed epithelial–stromal tumours, prostate 671–2
mixed function oxidase inducers 438
mixed germ cell tumours
  central nervous system 813
  testis 682
mixed mullerian tumour 629–30, 634
mixed neuronal–glial tumours 792, 805–7
mixed sex cord–stromal tumour 682
MK571 1352–3
MKK4/SEK1 (MKK4) 593, 1095–6
MLH1 (hHMH1) 30, 94–5, 104, 557
  chromosomal locus 1314–15
  DNA mismatch repair 94–5
  hereditary non-polyposis colorectal cancer (HNPCC) 43–4, 96, 324, 550, 556–7, 1314–15
  hypermethylation-related silencing
    pancreatic cancer 591
    sporadic colon cancer 96, 324, 556–7
  non-small cell lung cancer 1314–15
  tumour cisplatin resistance 95
  see also MutSβ complex
MLH2 (*hMLH2*) 30
MLH3 (*hMLH3*) 94
MLT1 700–1
MMAC1 (mutated in multiple advanced cancer) *see PTEN*
MNNG/HOS 1118
models of cancer 863–5
  in vitro 864–5
  in vivo 913
  see also animal models
modified continual reassessment method (MCRM) 1570–1
molecular epidemiology 377–85
  advantages/disadvantages 385
  early response markers 377, 378–9
  goals 377
  internal dose markers 378–9
  study design 377
  susceptibility markers 377, 380–4
molecular genetics
  anaplastic large cell lymphoma 711
  atypical alveolar hyperplasia 513–14
  bladder cancer 657–8
  brain tumours 1123–4
  breast cancer 612, 867–8, 869, 1092–3
  bronchopulmonary neuroendocrine tumours 511
  cervical cancer 626–7, 1314–15
  colorectal cancer 36–7, 556–7, 867–8, 1090
  diagnostic techniques 473–4
    costs 474
  diffuse large B cell lymphoma 703–4
  endocrine tumours 606–8
  endometrial cancer 633
  Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 722–3, 738, 756
    EWS fusions 722–3, 738–9, 756
  follicular lymphoma 699
  gastric cancer 542
  heart tumours 785
  hepatocellular carcinoma 571
  lung cancer 1314–5
  lymphoma 692–3
    prognostic significance 694–5
  MALT type marginal zone B cell lymphoma 700–1
  mantle cell lymphoma 692, 698
  mesothelioma 771–3
  micrometastases detection 475
  multiple myeloma 730

nephroblastoma (Wilms' tumour) 30, 40–1, 248
  nervous system tumours 817–21
  oesophageal cancer 541–2
  oral cancer (squamous cell carcinoma) 494–6
  osteosarcoma 724
  ovarian cancer 641
  pancreatic cancer 591–4
    medullary carcinoma 582–3
  pancreatic intraepithelial neoplasia (PanIN) 581–2
  penile cancer 677
  precursor B cell lymphoblastic leukaemia/lymphoma 696
  prostate cancer 673
  renal cell carcinoma 652–3
  salivary gland carcinoma 498–9
  skin cancer 487
  soft tissue tumours 737–40, 746
  testicular cancer 684–5
  tumour classification 469–71
molecular targeting, antitumour compound screening 1044
Moll glands 826–7
molybdenum, dietary 424
Mom-1 1091
momab (murine antibodies) 1380
monobutyrin 1406
monoclonal antibodies 1379
  nuclear medicine 1223–4
    imaging accuracy 1223–4
    imaging protocol 1223
    precautions 1224
    principle 1223
    scan interpretation 1223
    therapeutic potential 1224
  therapeutic growth factor receptor blockade 1488–9
  see also antibody-based therapy
monoclonal gammopathy of unknown significance 730
monocyte-derived antigen-presenting cells 1392
monocytes 1056
  haematopoietic stem cells 1103
  isolation from blood 1108–9
  TGF-β response 181
monodispersed cultured cells 887
monolayer culture, antitumour compound screening 1044
monounsaturated fats 395
MOPP chemotherapy 414–15
Mort1 (Fadd) 123, 124–5, 153
mortality data 264
mos 1325
motility factors 987–8
mouse (*Mus musculus*) 964
  cornea model 1028
  genetic manipulations 891–5
    limitations 901
    see also knockout mouse studies; transgenic models
  lymphoma genotoxicity assay 354
  plasmacytomas 1105–6
  prostate reconstruction (MPR) model 1097
  skin chemical carcinogenesis model 1086–7
    inositol hexaphosphate (IP6) actions 1087
  see also rodent carcinogenicity studies; rodent inbred strains
mouse mammary tumour virus (MMTV) 895–6, 923, 924, 1093
  c-*myc* transgenic mouse studies 896
mouse sarcoma virus 1128–9
Mre11 97, 104
Mrit 125
MRP1 1347, 1352
  expression in tumours 1354
  knockout mouse studies 1352
  multidrug resistance 1353
  physiological functions 1352
  substrates 1352
MRP2 1347
  multidrug resistance 1353
  physiological functions 1352
MRP3 1352
  multidrug resistance 1353
  physiological functions 1352
MRP4, expression in tumours 1354
MRP5 1352
  expression in tumours 1354
MRP6 1352
MRPs *see* multidrug resistance-associated proteins
MSG1 184–5
MSH2 (hMSM2) 11, 94–5, 104, 557
  hereditary non-polyposis colorectal cancer (HNPCC) 43–4, 96, 550, 557
  knockout mouse studies 95
  see also MutSα complex
MSH3 (hMSM3) 94
  see also MutSβ complex
MSH6 (hMSH6) 94–5, 104, 557
  hereditary non-polyposis colorectal cancer (HNPCC) 43–4, 550, 557
  see also MutSα complex
MTA1 (*mta1*) 869
  metastasis-related expression changes 869
MTS1 (*mts1*) 1445
  central nervous system tumours 1123–4
  metastasis-related expression changes 869
MTT assay 964
MUC1 1392–3
mucociliary escalator 504
mucosa-associated lymphoid tissue (MALT) 548
  lymphoma *see* MALT lymphoma (MALToma)
  nasopharynx 503
Muir-Torre syndrome 550, 636
Mullerian system 617
multicentric angiofollicular lymphoid hyperplasia *see* Castleman's disease
multidetector CT scanning 1165–6
multidrug resistance (MDR) 1046–8, 1347
  ABC transporters 1046, 1337–8, 1348
  apoptosis 1046, 1047–8
    mismatch repair defects 95
  atypical 1046
  BCRP/MXR/ABCP transporter 1045, 1339
  cell lines production 1044
  knockout models 1044–5
  multidrug resistance proteins (MRPs) 1347, 1351–4
  non-P-glycoprotein-mediated 1046, 1347
  P-glycoprotein-mediated 1046, 1047, 1347, 1348–51
  transport proteins 1347–8
multidrug resistance (MDR) reversal 1042, 1048–53, 1356
  agent evaluation 1048–51
  drug combinations 1051
    additive effect 1051
    combination index (CI) 1052
    experimental design 1052
    median-effect plots 1051–2
    quantitation of effects 1051–2
    therapeutic goals 1051
    see also antagonistic drug combinations; synergistic drug combinations
multidrug resistance proteins (MRPs) 1347, 1351–4
  anthracyclines resistance 1337–8
  expression in normal tissues 1352
  expression in tumours 1354

multidrug resistance proteins (MRPs) – *continued*
gene loci 1351
knockout mouse studies 1044–5
modulators 1352–3
multidrug resistance 1353
role of glutathione 1353–4
physiological functions 1352
protein family 1352
substrates 1352
Vinca alkaloid transport 1328
multiple endocrine neoplasia (MEN) 40, 601
multiple endocrine neoplasia 1 (MEN-1) 40, 44, 601, 607
genetic aspects 40
islet cell tumours 587, 595
parathyroid lesion clonality analysis 251
multiple endocrine neoplasia 2 (MEN-2) 40, 44, 595, 601
medullary thyroid carcinoma 40, 601
multiple endocrine neoplasia 2A (MEN-2A) 40, 601, 608
multiple endocrine neoplasia 2B (MEN-2B) 601, 608
multiple exostoses 40, 722
multiple lipomatosis 40
multiple myeloma 729–30, 1382–3
bone lesions
false-negative bone scan 1218
radiography 1148–9
clinical features 729–30
clinical management 730
Fas (CD95) somatic mutations 123
immunohistochemistry 729–30
Kaposi sarcoma-associated herpesvirus (KSHV) association 722–3
molecular studies 730
pathology 729–30
stem cell transplantation 1546–7
variants 730
multistep carcinogenesis 31–3, 41, 87, 435, 955, 960–1, 967–8, 1086
apoptosis defects 119
Barrett's model 25
chemical carcinogens 286–8, 289
dietary factors 396
initiation 286–7, 1086–7
Knudson's hypothesis 26–7, 593
Land and Weinberg 31–2
mouse skin model 1086–7
progression 288, 1086–7
promotion 287–8, 1086–7
Vogelstein's model (Vogelgram) 32–3
Mus musculus *see* mouse
muscle tissue 3
muscle-specific actin (MSA; HHF35)
rhabdomyosarcoma 752–3
soft tissue tumours 745
musculoskeletal plain radiography 1147, 1148–9, 1152–3
musculoskeletal tumours 789
chemoembolization 1256
percutaneous biopsy 1247, 1250
mushroom hydrazines 273–5, 391
mustard gas 89, 414, 1295
mutagenic activity assays 280–1
bacterial 352–4
mutated in colon cancer *see MCC*
mutated in multiple advanced cancer (*MMAC1*) *see PTEN*
mutation 19–20, 1522
chemical carcinogen induction 280–2, 379
mutational fingerprints 281, 379–80
chemopreventive intervention targets 436
cytosine methylation site susceptibility 322, 379
DNA damage-related 1507
DNA adducts 279
DNA low-fidelity polymerase 1509
early response markers 377, 379
mutational fingerprints 281, 379, 380
MutSα complex (MSH2:MSH6) 94–5
MutSβ complex 94–5
MXR (ABCP/BCRP) transporter 1045, 1337–8, 1339
MYB (*myb*)
antisense oligonucleotide targeting 1443, 1446
pancreatic cancer 592
Myc (myc)
apoptosis induction 1516
cell growth signalling 15
gene transcription regulation 28
Ras combined action in tumour formation 31–2
TGF-β-induced growth arrest 181, 190
MYC (myc) 215, 1525
activation by integrated provirus 76
Burkitt's lymphoma 339–40, 1105–6
gastroenteropancreatic endocrine tumours 607
mouse plasmacytomas 1105–6
multiple myeloma 730
small-cell lung carcinoma 512
soft tissue sarcoma 760
transgenic mouse studies 896
Wnt target genes 203–4, 206–7
Mycobacterium avium intracellulare 508
mycophenolate mofetil 1552–3
mycoplasmal cell culture contamination 883
mycosis fungoides (cutaneous T cell lymphoma) 486
clinical features 486
pathology 486
treatment 486
mycotoxins, dietary 389–90, 528
myd 960
myelodysplastic syndromes 1546–7
myeloid stem cells 6
myeloperoxidase 278
myeloproliferative leukaemia virus (MPLV) 1111
myoD 960
MyoD1
rhabdomyosarcoma 752–3
soft tissue tumours 745
myofibroblastic sarcoma 778
epidemiology 783
low grade (myofibrosarcoma) 747
pathology 780–1
myofibroblastic tumours 746–8
myofibroblasts 1025
Myogenic Factor 3 *see* MyoD1
myogenin
rhabdomyosarcoma 752–3
soft tissue tumours 745
myogenin 960
myoglobin 966
rhabdomyosarcoma 752–3
soft tissue tumours 745
myometrium 617
myositis ossificans 743–4
myotube differentiation 964
myricetin 424
myxofibrosarcoma 749
myxoid chondrosarcoma
chromosomal translocations 757
electron microscopy 746
extraskeletal 757
myxoid liposarcoma 469–70, 750

chromosomal translocations 750
EWS fusions 738–9
myxoma
eyelid 827–8
heart *see* cardiac myxoma

N

N-*ras* 1525
antitumour compound molecular targeting 1044
naevi
eyelids/conjunctiva 826–7
naevocellular 250, 479
uvea 826–7
naevoid basal cell carcinoma syndrome *see* Gorlin–Goltz syndrome
naevus of Ota (oculodermal melanocytosis) 826–7
Nagasaki 311, 531–2
naked cuticle 204
2-naphthylamine 273, 380, 413–14, 419
Napoleon 35, 527–8
nasal cavity cancer 262
nasal glioma 506
nasal NK/T cell lymphoma 711–12
nasal sinuses 503
nasopharyngeal angiofibroma 507
nasopharyngeal carcinoma 505–6
clinical management 506
clonality analysis 251–2
environmental/dietary factors 340, 506
Epstein–Barr virus 55, 59–61, 335, 339, 340, 506
IgA/viral capsid antigen (VCA) screening 340
global distribution 59–60, 340
pathology 506
nasopharyngeal neurilemmoma 506
nasopharyngeal neurofibroma 506
nasopharyngeal neuroma 506
nasopharyngeal squamous papilloma 505
nasopharyngeal tumours 505–7
clinical features 505
lymphoid 506–7
metastatic 503
polyclonal processes 250
vascular 507
NAT1 polymorphism 549
NAT2 polymorphism 304, 549
National Cancer Institute (NCI)
Cancer Drug Screen 1043
sarcoma grading system 741–2
natural killer (NK) cell lymphomas *see* T cell/natural killer (NK) cell lymphomas
natural killer (NK) cells 999–1000
beige/nude (*bg/nu*) mouse models 917
cytokine activation 1371
NB-506 1318–19
NB/70K 636
NBS1 104
Nck 166–8, 1487
NcoA-1 *see* SRC-1
NcoA-2 *see* SRC-2
NCoR (RIP 13) 140–1
NCTN 109 880
Nd:YAG laser photocoagulation 1264–5
NDA (new drug application) 1042, 1574
necrosis 957
mitochondrial cytochome *c* release 130
NEMO (IKKγ) 154
NEMO-like kinase (NLK) 204
neo 894
neonatal hepatitis/biliary atresia 568
neoplasia 461
neoplastic cells 25
Neovastat 1479–80
nephroblastoma (Wilms' tumour) 40–1, 646–8
epidemiology 647
genetic aspects 40–1
Wt1 mutations 30, 40–1
hepatoblastoma association 573
immunohistochemistry 648
loss of heterozygosity (LOH) 248
management 648
pathology 646–7
gross features 647–8
microscopic features 648
presentation 647
prognostic factors 648
screening high-risk patients 647
staging 647
syndromic associations 647, 737
nephron 645
nervous system development 789–90
nervous system tumours 789–822
classification 790
see also central nervous system tumours
nervous tissue 3
neu see HER2
neural cell adhesion molecule (NCAM) *see* CD56
neural crest 789
eye/ocular adnexa development 825
neural microvessel angiogenesis 979
neural tube 789–90
neuregulins 1487
neurilemmoma
larynx 508
nasopharynx 506
neuroblastic tumours 792
adrenal gland/sympathetic nervous system 794
neuroblastoma 799–800, 815
animal models 1127–8
developmental field effects 1403
differentiation into ganglioneuroma 1403
experimental induction 1128–9
MDR1 expression 1351
MIBG imaging 1226–7
multidrug resistance proteins (MRPs) expression 1353
olfactory 506
stem cell transplantation 1546–7
telomerase activity 113–6
neurocutaneous syndromes (phakomatoses) 817–18, 836–7
neurocytoma 807
neuroendocrine cell adenoma 599–600
neuroendocrine (Feyrter) cells 504, 599
neuroendocrine tumours 466–7, 599–600
classification 602–3
clinical management 608–9
somatostatin analogues 609
diagnosis/imaging 608
immunohistochemistry 602
liver tumour chemoembolization 1259
lung 510–11
malignant potential 602–3
pathology 602–3
pentetreotide (Octreoscan) scanning 1224–5
prognostic factors 608

neuroendocrine tumours – *continued*
somatostatin receptor expression 608–9
neuroepithelial tumours 789
biological categories 791–2
neuronal/mixed glial–neuronal tumours 805–7
primary of mature neuroepithelium 800–8
primary of primitive neuroepithelial tissue 795–800
retina 831
neurofibroma 792, 815–16, 1123
diffuse 834
heart 783
imaging 743–4
isolated 834
larynx 508
malignant transformation 734
nasopharynx 506
orbit 834
plexiform 819, 834, 838
neurofibromatosis 568, 734
retinal tumours 830–1
neurofibromatosis type 1 (NF1; von Recklinghausen disease) 737, 815–17, 836–7
clinical features 38
genetic aspects 38, 836–7
gliomas 819
optic 38
malignant peripheral nerve sheath tumours 756
pathology 819
plexiform neurofibroma 819
orbit 834
neurofibromatosis type 2 (NF2) 38–9, 816, 836–7
acoustic neuromas 38, 819–20, 846
genetic aspects 38–9, 819, 836–7
pathology 819–20
neurofilament proteins
brain tumours 795, 1123–4
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
neuroblastic tumours 815
neuroglial cell adhesion molecules (NG-CAMs) 230
neuroma, nasopharynx 506
neuron-specific enolase
desmoplastic small round tumour 759
endolymphatic sac papillary tumour 848
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
islet cell tumours 586
neuroblastic tumours 815
neuroendocrine tumours 602
neuronal stem cells 10
neuronal tumours 792, 805
nasopharynx 506
neurons 3, 789
development 789–90
neuroreceptor markers 1231–2
neurotrophin growth factor receptor 1486
neutrophils 1056
bone marrow precursor cells 6
new drug applications (NDAs) 1041, 1574
NF1 30, 44, 836–7
central nervous system tumours 1123–4
neurofibromatosis type 1 38, 737, 819
Nf1 transgenic/knockout mouse studies 1131
NF2 30, 44, 836–7
central nervous system tumours 1123–4
mesothelioma 772–3
neurofibromatosis type 2 38–9
NF-κB 137, 140
antisense oligonucleotide targeting 1443
apoptosis regulation 122–6
mechanisms 131
radiotherapy response pathways 1364–7
therapeutic approaches 131
TNF receptor signalling 153–4
NF–Noonan syndrome 38
nfnB 1538
nickel 262, 275, 407, 505, 512
epigenetic mechanisms of carcinogenesis 328
nicotine 404–5, 509
addiction 407, 410, 509
patch 410
nicotine-derived nitrosamino ketone (NNK) 405–7, 1092
cancer chemopreventive strategies 410
lung cancer causation 410
metabolic activation 407–8
metabolite biomarkers of tobacco smoke uptake 408
nidogen 210, 216–7
nifedipine 1048
P-glycoprotein inhibition 1350
Nijmegen breakage syndrome 19, 87–8, 104
NIK 154
Nip3 129
nipple discharge 1241
galactography (ductography) 1241–3
nitrates, dietary 529–531
nitric oxide 391–2, 395, 530–1
nitric oxide synthase 391–2, 395
nitrites, dietary 391–2, 531, 568
nitrogen mustard (mechlorethamine) 89, 1295–6
nitrogen mustards 1295–8
DNA adduct formation 1296–8
1-nitropyrene 273
nitroreductase 1386–7
nitroreductase–CB1954
genetic prodrug activation therapy 1537–9
bystander effect 1539
nitrosamines 285–6, 377–8, 380–1, 530
bladder cancer 655, 660
dietary 391–2
DNA adduct formation 90
prevention of formation 427, 437–8
tobacco smoke 301, 401–2, 405–7
nitrosatable compounds, dietary 391–2
N-nitroso compounds 391–2, 438, 794
chemopreventive strategies 437–8
neurocarcinogenesis, animal models 1130–1
oesophageal cancer 525, 528
N-nitrosodialkylamines 391–2
N-nitrosodiethylamine 405–7
N-nitrosodimethylamine 391–2
N-nitrosonicotine (NNN) 407
nitrosoureas 1300–2
DNA adduct formation 90
experimental neurocarcinogenesis 1130
Nix 129
NK-109 1316
nm23 868–9
nocodazole 1331
Nod (CARD4) 127
nodular fasciitis 743
nofetumomab merpentan (Verluma) 1223
non-Hodgkin's lymphoma
anti-CD20 radioisotope-labelled antibodies 1385
antisense therapy 1446

bone 730, 1382–3
central nervous system 812
classification 694
cytokine-mediated immunotherapy 1551
donor lymphocyte infusion therapy 1551
hepatitis C virus 342
human immunodeficiency virus (HIV)/AIDS patients *see* human immunodeficiency virus (HIV)
iatrogenic lymphoproliferative disorders 718
larynx 508
lung 519
MDR1 expression 1351
prognostic factors 694–5
stem cell transplantation 1546–7
stomach 344
Waldeyer's ring tumours 506–7
non-homologous end joining 97
non-ionizing radiation, energy spectrum 308
non-obese diabetic (NOD) mouse 917–18
non-ossifying fibroma 723
non-seminomatous germ cell tumours 679–80
clinical management 685
pathology 681–3
non-small cell lung cancer *see* lung cancer
non-steroidal anti-inflammatory drugs (NSAIDs) 1426–7
antigenotoxic/chemopreventive activity 423, 428–30, 436, 443
apoptosis stimulation 430–2
bladder cancer prevention 1434
colorectal cancer prevention 144, 549, 552, 1432
combined treatment strategies 449
lung cancer prevention 1431
oesophageal cancer prevention 1432
oral cancer prevention 1427–31
pancreatic cancer prevention 591
skin cancer prevention 1435
nose 503
NPK/*ALK* fusion 691
NPM/*ALK* fusion 711
NQOI knockout mouse studies 1090
NTRK3/*ETV6* fusion 747
nuclear envolope 1324
nuclear matrix 1025–6
nuclear matrix proteins, urothelial carcinoma screening 656
nuclear medicine 1149, 1215–33, 1558
bone scanning 1216–18
gallium scanning 1222–3
gamma camera 1215
iodine scanning 1218–20
isotopes/radiotracers 1215
MIBG (meta-iodobenzylguanidine) 1226–7
monoclonal antibodies 1223–4
peptides 1224–5
scintimammography 1225
sentinel node localization 1232–3
thallium scanning 1220–2
see also positron emission tomography (PET)
nuclear receptor (NR) box 139–40
nuclear receptor (NR) family 132
nucleosidases 1273
nucleoside analogues 361
nucleoside diphosphate kinase 1273
nucleoside monophosphate kinase 1272–3
nucleosides
metabolism 1272–3
nomeclature 1271
transport 1272
nucleosomes 317–8
nucleotide excision repair (NER) 91, 94, 1509–10
global genome repair (GGR) 91–2
transcription-coupled repair (TCR) 91–4
nucleotides
nomeclature 1271
roles in cellular metabolism 1271–2
synthetic pathways 1272
nude (*nu/nu*) mouse studies 886, 967
hollow-fibre assay 1044
melanoma models 1116
metastasis models 1003, 1005
model characteristics 916–17
mouse cornea assays 1028–9
transformed cell tumorigenicity testing 961–2, 964
xenograft models 915–17
orthotopic sites 1088
see also beige/nude mouse studies; beige/nude/X-linked immunodeficiency mouse studies
nude (*nu/nu*) rat studies 917

O

obesity 632, 1139
occludens junctions *see* tight junctions
occupational cancers 263, 413–19
at risk occupations 415
bladder cancer 417, 419
carcinogens *see* occupational carcinogens
CAREX database 418
epidemiological studies 413
estimation of total burden attributable to occupation 416–18
developing countries 418
Europe 418
hepatocellular carcinoma 568
historical aspects 413
laryngeal cancer 417
lung cancer 417, 419, 512
spiral CT/biomarker screening 419
mesothelioma 417
nasopharyngeal squamous cell carcinoma 505
preventive strategies 418–19
primary prevention 418
industrial hygiene 418
legal initiatives 418
salivary gland carcinoma 498
sinonasal cancers 417, 505
tobacco smoke contribution 417–18, 512
workplace surveillance/monitoring 418–19
occupational carcinogens 273, 413–16
assessment difficulties 415–16
DNA adduct detection in tissues 300
environmental carcinogens differentiation 414–15
established carcinogens 413–14
IARC Monograph programme 414
animal studies 414
biological data 414
descriptive categories 414
levels of evidence of carcinogenicity 414
ionizing radiation 311–12
legal restrictions 418
personal protective equipment 418
probable/possible carcinogens 414
substitition of less hazardous materials 418
see also chemical carcinogens
Octreoscan *see* pentetreotide
octreotide

octreotide – *continued*
  neuroendocrine tumours 609
  side effects 609
  thyroid imaging 1220
ocular adnexal development 825–6
ocular adnexal tumours 825–39
  aetiology 826–7
  clinical management 837–9
  epidemiology 826
  pathology 826
  prevention/screening 827
oculodermal melanocytosis (naevus of Ota) 826–7
odds ratio 260–2
oesophageal cancer 523
  adenocarcinoma 525–6, 532–3, 540–2
    pathology 535–7
  aetiology 528
  alcohol association 528
  chemoprevention 1432
  clinical features 532
  cytological screening programmes 532
  dietary factors 525–6, 528
  drug resistance 1351
  epidemiology 525–6
  erbB-2/neu expression abnormalities 966
  human papillomavirus (HPV) 337, 528
  molecular genetics 541–2
  multidrug resistance proteins (MRPs) expression 1354
  prevention 532–3
  prognosis 540–1
  squamous cell carcinoma 525, 532, 537–42
    gross pathology 534–5
    progression from dysplasia 534–5
  staging 537, 1558–9
  telomerase activity 113–15
  tobacco smoking association 528
  treatment 537–41
oesophagus
  anatomy 523–4
  development 523
oestrogen 135
  animal model influences 921
  Bcl expression regulation 132
  breast cancer promotion 276–7, 1449
    treatment strategies 1455–8
  breast tissue responses 9–10, 1029–30
  cell growth regulation 17, 25
  cumulative lifetime exposure 33
  differentiation regulation 959–60
  endometrial cancer induction 630–2, 1139
    animal models 1140–2
  endometrial response 627
    hyperplasia 628–9
  mechanism of action 1452
  oestrogen receptor interaction 136–7, 1452
  promoting activity 287
  terminal differentiation induction 966–7
oestrogen receptor 135, 966
  antioestrogens mechanism of action 1453
  chaperone proteins 137
  expression
    breast cancer 135, 142–3, 467, 472, 1449–50
    female/male organs 136
    immunohistochemical detection 467, 469
    normal breast 142–3
  heterodimer formation 1451
  hormone resistant breast cancer 144
  in vivo function 136
  molecular genetics 135
  mutations 143–4
  oestrogen interaction 136–7, 1452
  promoter hypermethylation 324
  regulation of function 141–2
    phosphorylation 141–2, 1451
    proteasome-mediated degradation 142
  transcription activation 136–42, 1451–2
    basal transcription factor interactions 137, 1452
    coactivator interactions 138–40, 1452–3
    corepressor interactions 140, 1452
  variant forms 143
oestrogen receptorα
  basal transcription factor interactions 137–8, 1451
  breast cancer therapeutic target 1451
  coactivator interactions
    CBP/p300 140
    SRC proteins 139–40
  expression
    normal breast 142–3
    premalignant/malignant breast lesions 142–3
  heterodimer formation 1451
  knockout mouse studies 136, 1451
  molecular genetics 1449
  mutations 131
  phosphorylation 141–2, 1451
  proteasome-mediated degradation 142
  splice variants 143
  structure 1451
    functional domains 136, 1451
    ligand-binding domain 136–7, 1451–4
  transcription activation 137, 1451
oestrogen receptorβ
  breast cancer therapeutic target 1451
  expression
    breast tumours 143
    normal breast 143
  heterodimer formation 1451
  knockout mouse studies 136, 1451
  molecular genetics 136
  phosphorylation 142
  SRC protein interactions 140
  structure 1451
    functional domains 136, 1451
    ligand-binding domain 136–7, 1451
  transcription activation 1451
  variant forms 143
oestrogen response elements (EREs) 137, 1451
oils exposure 273, 478, 481
OKT3 1384
oleic acid 395
olfactory esthesioneuroblastoma 791, 815
olfactory neuroblastoma 506
oligoastrocytoma 806
oligodendrocytes 789–90
oligodendroglioma 792, 1123
  anaplastic 803–4
  experimental induction 1131
  pathology 803–4
olive oil 395
Ollier disease 722
oltipraz 423, 438
Omenn syndrome 917–18
omeprazole 532–3
oncogene hypothesis 27
oncogenes 18, 27–8, 35, 41, 195, 288, 293, 335, 891, 955, 1084, 1524–5

angiogenesis stimulation 1469
antisense oligodeoxynucleotide (ODN) targeting 1443
cell morphological change induction 963–4
chemopreventive intervention targets 430, 436, 439–40
cultured cell transformation 962–3, 967
in vitro assays 964
expression assays 963–4
gene therapy 1524
haematological malignancy 1102–7, 1105
head and neck cancer 495
hepatocellular carcinoma 571
historical aspects 27
knockout mouse studies 904–5
multistep activation in carcinogenesis 31–2, 288
mutations 25–7, 195, 204–5
cellular vaccine targets 1393
dominant 27–8, 1524
malignant transformation 1524
post-transplantation lymphoproliferative disorders 718
pancreatic cancer 591–2
progression of cancer 867–8
transfection experiments 869–70
ribozyme targeting 1444–5
transgenic mouse studies 891
in vivo 895–6
see also proto-oncogenes
oncogenesis 25–33
cellular events 26–7
infectious agents 335–49
oncolytic viruses 1071–2
Oncoscint (satumomab pendetide) 1223–4
one-step DNA repair (OSR) 1509
ONYX-015 adenovirus 1071, 1427–31
opiates 594, 1255–6
Opisthorchis felineus 347
Opisthorchis viverrini 347, 573–4
optic glioma 38
optic nerve tumours 835
clinical management 838
optic vesicles 825
optical coherence tomography (OCT) 1420
oral cancer (squamous cell carcinoma) 491–6
aetiology 492–3
clinical management 496
epidemiology 491–2
hepatitis C virus association 342
molecular genetics 494–6
pathology 493–4
macroscopic appearances 493
microscopic appearances 493
variants 494
premalignant lesions/leucoplakia 491, 493, 1427–31
prgnostic factors 496
second primary tumours (field cancerization) 250, 496
TNM staging 496
tobacco use association 402–3, 407, 478, 492–4
oral cavity 491
oral contraceptives 630–2, 636
hepatocellular adenoma risk 576
hepatocellular carcinoma risk 568
oral hairy leukoplakia 56–8
oral mucosa, multidrug resistance proteins (MRPs) 1352
orbital spaces 826
orbital tumours 826, 834–5
children 826
clinical management 838
organochloride pesticides 568
ormplatin 1303
ornithine decarboxylase 429, 439–40
oropharyngeal cancer 402
OSI-774 (CP358774) 1500
ossification 722
osteoblastoma 722
osteoblasts 722
osteochondroma 722
osteoclast-like giant-cell tumour 583
osteoclasts 722
osteocytes 722
osteogenic sarcoma 827
orbit 835
osteoid 722
osteoid osteoma 723
osteoma 503
osteomyelitis 722–3
osteonecrosis 1150–1
osteopetrosis 1545
osteopontin 242
osteoporosis 1150–1, 1462–3
chemotherapeutic agent side effect 1151
plain radiography 1150–1
osteosarcoma 465, 724–5, 736–7, 1118
clinical management 724–5
extraskeletal 757
genetic predisposition 722
histology 724
metastases detection by bone scanning 1216, 1221
models 1117–9
canine spontaneous 1119
cell lines 1117–8
clinical relevance 1119
human–murine xenografts 1118
murine syngenic 1118–19
molecular genetics 724
nasopharynx 503
p53 knockout mouse studies 907
p53 mutations 29, 42
plain radiography 1149
predisposing factors 722–3
prognosis 724–5
retinoblastoma association 19, 36, 724, 837
surface 725
otic vesicle (otocyst) 841
oval cells 10
ovarian cancer
aetiology 636
BRCA1 44, 636, 641
BRCA2 636
clinical features 635
clinical management 642
DNA mismatch repair defects 96, 550
drug resistance 95, 1351
epidemiology 635
erbB-2/neu expression abnormalities 966
familial predisposition 44, 636
breast cancer associations 44, 1425
gene therapy 1070
hepatic metastases 1196–7
interleukin 2 (IL-2) therapy 1375
magnetic resonance imaging (MRI) 1208–9
molecular genetics 641
multidrug resistance proteins (MRPs) expression 1354
mutiple tumour clonality analysis 250–1
pathology *see* ovarian tumours
prognostic factors 641

ovarian cancer – *continued*
  satumomab pendetide 1223
  screening 636–7, 1195
  spread 228–9, 635
  staging 642
  stem cell transplantation 1547
  TGF-β TβRI mutations 186–7
  tumour markers 636
  ultrasound 1194–5
    angiogenesis Doppler imaging 1194
ovarian cyst 1194
ovarian cystadenocarcinoma 1149
ovarian tumours 635–42
  clinical features 635
  epidemiology 635–6
  epithelial 635
  germ cell 635, 640
  metastatic 640, 1194
  pathology 637–40
    adenocarcinoma 637
    adenofibroma 637
    benign 637–9
    borderline type 637–9
    clear cell 637, 639
    cystadenofibroma 637
    endometrioid 637, 639
    malignant 637, 639
    mucinous 637, 639
    serous 637, 639
    transitional cell (Brenner tumours) 637, 639–40
  peritoneal implants 637
  sex-cord stromal 635, 640
ovary development 617
oxaliplatin 1303–6
  structure 1303
oxidation reactions 277, 294, 380–1, 427
  chemopreventive intervention targets 436
oxidative stress 111
oxygen regulation of angiogenesis 977
  see also hypoxia
oxytocin 9
ozogomycin 1386
ozone layer 309, 479

P

^{32}P postlabelling, DNA adducts detection 298, 378–9
^{32}P treatment 1215
P values 268
p14ARF 1516
p15 alterations 1514–15
  promoter region hypermethylation 323–4
p15INK4b 83, 1513
  Tax protein binding/inactivation 82
  transforming growth factor-β (TGF-β) induction 180–1, 186
p16 alterations 1108, 1514–15
  astrocytoma 1123–4
  central nervous system tumours 1123–4
  cervical cancer 626–7
  head and neck cancers 495
  lung cancer 409–10
  mantle cell lymphoma 698
  pancreatic cancer 588, 592–3
  pancreatic intraepithelial neoplasia (PanIN) 581–2
  promoter region hypermethylation 31, 322–3, 591
  salivary gland carcinoma 499–500
p16–cyclin D1–pRb pathway
  bronchopulmonary neuroendocrine tumours 511
  cancer cell defects 1507
p16INK4a 30, 83, 1513
  cell cycle regulation 111, 1526
  Tax protein binding/inactivation 82–3
p18 alterations 1514–15
p18INK4c 1513
p19ARF 99–100, 1526
p19INK4d 1513
p21
  antimicrotubule agents mechanism of action 1327
  cyclin-dependent kinase (cdk) inhibition 15, 17, 1108, 1513–14
p21/WAF/CIP (WAF1/Cip1) 896, 1307, 1327, 1405, 1408–9
  antisense oligonucleotide targeting 1443
  head and neck cancers 495
  transforming growth factor-β (TGF-β) induction 180–1, 186
p21bcr/abl 1110–11
p21KIP-1 905–7, 1513
p21WAF 905–7
  methylation 31
p27 395, 1470
  cyclin-dependent kinase (cdk) inhibition 15, 1513
  oesophageal adenocarcinoma 542
p27KIP1 1513
p27KIP1 knockout mouse studies 908
p30/32 MIC2 *see* CD99
p38 175
  stress activated protein kinases (SAPKs) 1364–5
p53 140, 955, 959, 1108
  antimicrotubule agent mode of action 1326–8
  apoptosis regulation 19, 29, 30, 131, 896, 905–7, 1516
    dysfunction for tumour progression 1516
    mitochondria-dependent mechansisms 127–8
    radiotherapy response 1363–4
  cell cycle arrest 18, 896, 905–7, 1526
  DNA repair promotion 1508–9
  G_1 checkpoint 896, 1326–7, 1515–16, 1526
    therapeutic manipulation 1518
  G_2 checkpoint 1326–7
  HPV E6 protein interaction 53, 337, 620, 626–7
  immunohistochemistry 469
  nucleotide excision repair signalling 91
  p21 up-regulation 1513–14
  protein binding interactions 99–100
  Rad51 binding 99–100
  topoisomerase II 13
    activity modulation 1317
    inhibitor mode of action 1317
  transcriptional activation targets 896
p53 962–3, 967, 1525
  antisense oligodeoxynucleotide (ODN) targeting 1445–6
  antitumour compound molecular targeting 1044
  cell cycle regulation 109
  DNA damage-induced activity 18, 896–7
    cisplatin mode of action 1307
  gastric cancer-related polymorphism 542
  gene therapy 525, 1070, 1526–7
    clinical applications 496, 1527
    combined chemotherapy 1307, 1528
    preclinical studies 1526–7
  knockout mouse studies 368–9, 905–7, 1364
    brain tumour models 1131
    energy restriction effect 395
    production method 896–7
  loss during colon cancer progression 32–3
  mutations 11, 88, 131–2, 436, 1469

aflatoxin B_1-induced 288, 299, 379, 390, 1090
associated cancers 19, 29, 471
bladder cancer 657–60
bone tumours 722
breast cancer 29, 473, 613, 615
bronchopulmonary neuroendocrine tumours 511
cellular vaccine targets 1393
central nervous system tumours 1123–4
chemical-induced 288–9, 379
cholangiocarcinoma 575
chondrosarcoma 726
colorectal cancer 556
DNA adduct formation sites 299, 379
DNA repair defects 29, 1510–11
early response markers 379–80
endocrine tumours 607
endometrial cancer 633
familial cancers 41
gall bladder carcinoma 577
gastric cancer 250–1
head and neck cancers 495
immortalized cell lines 1087
L5178Y mouse lymphoma cells 354–5
laryngeal/pharyngeal tumours 250, 508
Li–Fraumeni syndrome 19, 33, 41–2, 722, 737, 896
lung cancer 299, 407–9, 511–12, 518
lymphomas 691
mantle cell lymphoma 698
multiple myeloma 730
oesophageal cancer 541–2
osteosarcoma 724
pancreatic cancer 592–3
pancreatic intraepithelial neoplasia (PanIN) 581–2
penile cancer 677
premalignant lesions 249–1, 253
salivary gland carcinoma 498–9
small-cell lung carcinoma 512
smoking-related dysplastic tracheobronchial mucosa 251
soft tissue sarcomas 739, 745, 760
thyroid carcinomas 607–8
ulcerative colitis neoplastic progression 557
ultraviolet radiation-induced 379
p53BP2 129–30
p53R2 1508–9
p55 153
see also tumour necrosis factor receptor 1 (TNFR1)
p57KIP-2 1513
p60 *see* tumour necrosis factor receptor 1 (TNFR1)
p70S6 kinase 1502
p73 324
p75 *see* tumour necrosis factor receptor 1 (TNFR2)
p80 *see* tumour necrosis factor receptor 2 (TNFR2)
p107
cell cycle regulation 17
progression across R-point regulation 1514
p130 1514
p130 cas 211
p300/CBP 99–100
P388 lymphocytic leukaemia model 915, 1043, 1351
P-glycoprotein 1076, 1341, 1348–51
anthracyclines resistance 1337–8
expression in tumours 1350–1
chemotherapy relapse 1351
soft tissue sarcomas 745
in vitro expression 1351
interactions with natural anticancer agents 1048
knockout mouse studies 1044, 1337–8
modulators 1350
multidrug resistance (MDR) 1046–7, 1337–8, 1348
reversal agents 1048, 1338
normal function 1046–7, 1349
structure 1348–9
substrates 1350
taxanes transport 1328
transport function 1349–50
Vinca alkaloids transport 1328
see also MDR1
P-selectin 1016–7
malignant invasion 230
p/CIP *see* SRC-3
paclitaxel (taxol) 21–2, 1070, 1329, 1471
combined drug treatment 23, 1327–8, 1489, 1491
β-lapachone 22
vinorelbine 1331
mechanism of action 1323, 1326–7, 1329–30
antitumour compound screening 1043
resistance 1076, 1328–9
side effects 1330
Paget disease
bone 722–3
nipple 614
penis 676
pain relief
chemoembolization 1255–6
pancreatic cancer 594
Pak1 130–1
palliative percutaneous ablative therapy 1261
palmar fibromatosis (Dupuytren disease) 746
pan 399, 492
Pancoast tumour 515
magnetic resonance imaging (MRI) 1205
pancreas
development 579
endocrine 579–80
exocrine 579–80
multidrug resistance proteins (MRPs) 1352
structure 579–80
pancreatic cancer
acinar cell carcinoma 583
aetiology 587–9
anti-epidermal growth factor receptor (EGFR) therapy 1489–90
clinical features 594
clinical management 594–5
gemcitabine clinical trials 1574–5
radioimmunoguided surgery 1565
dietary factors 589
DPC4 deletions 30, 592–3
ductal adenocarcinoma 580–2, 587
histological precursors (PanIN) 581–2
management 594–5
metastasis 580–1
E-cadherin gene mutations 1082
early detection 582, 589
epidemiology 587–9
epidermal growth factor receptor (EGFR) overexpression 1488
genetic predisposition 588–9, 591
intraductal papillary mucinous neoplasms 585
liver metastases 575
medullary carcinoma 582–3
metastatic tumours 587
molecular genetics 591–4
DNA mismatch repair genes 592–4
oncogenes 591–2
tumour-suppressor genes 591–3

pancreatic cancer – *continued*
mucinous cystadenoma/cystadenocarcinoma/borderline neoplasms 584–5
osteoclast-like giant-cell tumour 583
pain control 594
pancreatoblastoma (pancreatic carcinoma of infancy) 583, 587–8
percutaneous ablative therapy 1261
preoperative computed tomography (CT) 1169–72
prevention 591
prognostic factors 594
radiological detection 588–9
risk factors 587
screening 589–91
serous cystadenoma/cystadenocarcinoma 584
staging 1558–9
telomerase activity 113–15, 590
TGF-β TβRI mutations 186–7
tobacco smoking association 402, 407, 511, 588
transforming growth factor-α (TGF-α) 1488
tumour markers 590–1
pancreatic cyst 820
pancreatic cystadenoma
mucinous 584–5
serous 583–4
pancreatic ducts 579–80
pancreatic intraepithelial neoplasia (PanIN) 581–2
genetic changes 581–2
pancreatic polypeptide 595
pancreatic tumours 579–95
classification 580
endocrine 586–7, 595, 600
see also islet cell tumours
lymphoma/leukaemia 587
magnetic resonance imaging (MRI) 589, 1206
metastatic 587
non-endocrine
cystic 583–6
solid 580–3
pathology 580–7
percutaneous biopsy 1247
primary mesenchymal 587
solid-pseudopapillary 585–6
pancreatitis, chronic 589
pancreatoblastoma (pancreatic carcinoma of infancy) 583, 587–8
Paneth cells 7, 547
Papanicolaou test (Pap smear) 462, 621–3
atypical squamous cells of undetermined significance (ASCUS) 621–3, 626
endometrial cancer detection 632
high-grade squamous intraepithelial lesions (SILs) 623
low-grade squamous intraepithelial lesions (SILs) 623, 626
ovarian cancer detection 636
papillary renal cell carcinoma 651–3
papillary thyroid cancer 30
papilloma
bladder tumours 656
bronchus 510
Papillomavirus subfamily 50
papillomaviruses 19, 49–55, 335
benign lesions 53
classification 50
clinical aspects 50
E1 proteins/E1 ORF 51, 53
E2 proteins/E2 ORF 51–4
E3 ORF 52
E4 protein/E4 ORF 52–3
E5 proteins/E5 ORF 53
E6 proteins/E6 ORF 53–4
E7 proteins 53–4
E8 ORF 53
genome
integration into cellular DNA 53
replication 53–4
transcription 53–4
genomic organization 50–3
early region 50–1, 53–4
late region 50–1, 53–4
long regulatory region (LCR) 50–1
promoters 54
L1 protein 50
L2 protein 50–1
transforming activities 53
virion structure 50
see also human papillomavirus (HPV)
papovaviruses
cultured cell transformation 962
experimental brain tumour induction 1128–9
paracrine signalling 4, 6, 953, 1034
endothelium 1012
motility factors 1017
mesenchymal–epithelial interactions 1083–4
receptor tyrosine kinases 169
paraganglia 792
extra-adrenal tumours 602
paraganglioma 792
heart 771
jugulotympanic 845–6
larynx 508
nasopharyngeal 506
pathology 817
paraneoplastic syndromes
angiomatoid fibrous histiocytoma 743
lung cancer/small-cell lung carcinoma 515
renal cell carcinoma 648
parathormone-related protein (PTH-rP) 515
parathyroid adenoma 600
mutiple tumour clonality analysis 250–1
parathyroid carcinoma 600
Rb mutations 607
parathyroid hyperplasia 251
parathyroid tumours 599–600, 602
epidemiology 600
parietal cells 524, 528–9
parosteal osteosarcoma 725
parotid glands 491
PARP *see* poly(ADP-ribose) polymerase
passage number 887
passaging (subculture) 883–4
pathology 461–2
pathology reports 465
PAX3/FKHR fusion 753
PAX5 702
multiple myeloma 730
paxillin 214, 985
PBX1 10
PD153135 1491–2
PD168393 1491
PDGFβ/COL1A1 fusion 748
PDZ proteins 214–16
peanut agglutinin 652
pelvic ultrasound, female 1193
penacetin 260
penicillin 882, 1044
Penicillium 390

penile cancer 676–7
  clinical management 677
  effect of circumcision 677
  epidemiology 676–7
  histology 676
  human papillomavirus (HPV) 337, 676–7
  molecular genetics 677
  preinvasive lesions 676
  prognostic factors 677
  rare primary tumours 676
  screening 677
  squamous cell carcinoma 676
    subtypes 676
    superficial spreading 676
    verruciform 676
    vertical growth pattern 676
penile fibromatosis (Peyronie disease) 746
penis
  development 675
  structure 675–6
pentavalent DMSA, thyroid imaging 1220
pentetreotide (Octreoscan) 1224–5
  imaging protocol 1224
  indications 1224–5
  principles 1224
pentosan polysulfate 1468
pentostatin 1280
  adenosine deaminase inhibition 1280
  mechanism of cytotoxicity 1280
  structure 1279–80
pepsin 524
pepsinogen group I 524, 528–9, 531, 533
pepsinogen group II 524, 533
peptide hormones 599
percutaneous ablative therapy 1261
  goals 1261–2
  imaging guidance 1262
  target sites 1261
percutaneous biopsy 1247–51
  breast cancer 1238, 1241, 1247, 1250, 1564
  colon cancer 1250
  imaging guidance 1250–1
    computed tomography (CT) 1248
    fluoroscopy 1247
    ultrasound 1241, 1247, 1250
  liver tumours 1247, 1249–50
  lung tumours 1247–9
  mediastinal leions 1248–9
  musculoskeletal tumours 1247, 1250
  renal tumours 1247, 1250
  retroperitoneal structures 1248
  splenic mass 1247, 1250
  technique 1247–8
pericardial mesothelioma 778, 785
  asbestos exposure association 783–4
  epidemiology 783
pericytes 236, 243–4
perineuroma 792
periosteal osteosarcoma 725
peripheral nerve development 733
peripheral nervous system tumours 789
  biological categories 792
  pathology 814–17
  secondary (metastatic) 792
peripheral neuroectodermal tumour of bone (PNET) A8729:A8756 *see* Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET)
peritoneal tumours 767
  see also mesothelioma
peroxidase–antiperoxidase (PAP) system 468
peroxisome proliferator–activator receptor ligands 1407–8, 1426
  clinical studies 1413–14
peroxisome proliferator–activator receptors (PPARs) 1407–9
peroxisome proliferators 328–9
peroxynitrite 395
personal protective equipment 418
petasitenine 390
Petasites japonicus (coltsfoot) 390
Peutz–Jeghers syndrome 40, 44, 636
  colorectal hamartomatous polyps 547
Peyer's patches 548
Peyronie disease (penile fibromatosis) 746
PGP (Protein Gene Product) 120
  Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
  neuroendocrine tumours 602
phaeochromocytoma 40, 820
  epidemiology 600
  liver tumour chemoembolization 1259
  magnetic resonance imaging (MRI) 1207
  MIBG imaging 1215, 1226–7
  ret mutations 606
phakomatoses (neurocutaneous/ectomesodermal syndromes) 817–20, 836–7
pharmaceutical carcinogenicity assessment 365
  FDA/CDER Carcinogenicity Assessment Committee (CAC) 372–3
  implementation of ICH guidance 372–3
  see also rodent carcinogenicity studies
pharyngeal tumours *see* nasopharyngeal tumours
phase I clinical trials 448, 450–1, 1042–3, 1570, 1573
  escalation with overdose control (EWOC) 1570–1
  Fibonacci dose escalation scheme 1570–1
  maximum tolerated dose (MTD) determination 1570–1
  modified continual reassessment method (MCRM) 1570–1
  safe starting dose 1570–1
phase I metabolism 277, 294
  cancer risk-related enzyme polymorphisms 304, 380–1
  chemopreventive approaches 427, 438
phase II clinical drug trials 448–9, 1042–3, 1571–3
  recommended dose establishment 1571
  single-stage design 1571–2
  two-stage design 1571
phase II metabolism 277, 294
  cancer risk-related enzyme polymorphisms 304, 380–1
  chemopreventive approaches 427, 438
    enzyme inducers 428
phase III clinical drug trials 448–9, 1042–3, 1572–3
  analysis 1572–3
  equivalency design 1572
  ethics of randomization 1572
  factorial design 1572
  sample size 1572
  statistical considerations 1572
phenacetin 649–50, 653–4
phenethyl isothiocyanate (PEITC) 410
phenotype 1522
phenoxyacetic acids 737
phenyl isothiocyanate 428
phenylacetate 1406, 1408
  clinical studies 1413
phenylbutyrate 1406–7
  clinical studies 1413
Philadelphia chromosome 33, 1106
  see also bcr-abl fusion

phomopsin A 1329
phosphatase type 2A 1324
phosphate-buffered saline 880
phosphaturic mesencymal tumour 743
phospholipase C 155
phospholipase Cγ 166, 173, 175, 1487
phospholipase D 155–6
phosphorus, dietary 424
phosphorylation
  cell cycle regulation 15–17
  JAK–STAT signalling pathway 148
  oestrogen receptor function regulation 141–2
  tyrosine 161–2, 166–8, 173
    receptor tyrosine kinases 165–6, 169–70, 1486
    regulatory role 170
phosphotyrosine-binding proteins 165–6
phyllodes tumours
  breast 614–5
  clonality analysis 251
physical carcinogens 307–14
phytooestrogens 424, 439, 672–3
PI3 kinase inhibitors 1502
PI3 kinase pathway 166–8, 173, 175, 211, 1487
  invasion/metastasis 225–8
  oestrogen receptor phosphorylation 141–2
PI-4, 5-kinase 1487
pilocytic astrocytoma 1123–4
pilomatrixoma (calcifying epithelioma of Malherbe) 827
pineoblastoma 791
  pathology 808
pineocytoma
  experimental induction 1128–9
  pathology 808
pineocytoma rosettes 808
pinguecula 826–7
pioglitazone 1407–8
pipettes 878
pipetting devices 878
piroxicam 432
pituitary adenoma 606–7
  octreotide treatment 609
pituitary tumours 599–600
  epidemiology 600
  pathology 602
pivaloyloxymethyl butyrate (AN-9) 1406
pixels 1156
PKA 130–1
PKI166 1492, 1500
placenta growth factor 1 (PIGF-1) 238
placenta growth factor 2 (PIGF-2) 238
placenta, P-glycoprotein expression 1349
placental alkaline phosphatase (PLAP)
  germinoma 813
  testicular germ cell tumours 683
PLAG1 498–9
plain film radiography 1147–53
  abdominal neoplasms 1147, 1149, 1153
  lung neoplasms 1150, 1153
  musculoskeletal lesions 1147–9, 1152–3
plakoglobin 9
planar–polar solvent differentiation-inducing agents 1408
plantar fibromatosis (Ledderhose disease) 746
plasma cell leukaemia 730
plasmablastic lymphoma 704
plasmacytoma 705
  bone (solitary myeloma) 730
  c-*myc* activation 1105–6
  central nervous system 812
  larynx 508
  sinonasal tract 507
  testis 683
plasmin
  malignant invasion/metastasis 231–2
  transforming growth factor-β (TGF-β) activation 180
plasminogen activator system
  angiogenesis 241
  malignant invasion/metastasis 940, 985, 999–1000, 1019
  therapeutic inhibition 1480
Plasmodium 1291
platelet factor 4 (PF-4) 239, 971, 1468
platelet-derived endothelial cell growth factor (PD-ECGF; thymidine phosphorylase) 237, 239, 244
platelet-derived growth factor (PDGF) 25
  angiogenesis 237, 1073
  central nervous system tumours 1123–4
  v-*sis* oncogene 955
platelet-derived growth factor A (PDGFA) 1488
platelet-derived growth factor B (PDGFB) 1488
platelet-derived growth factor (PDGF) receptor 1486
  ligand-induced activation 169–70
platelet-derived growth factor (PDGF) receptor inhibitors 1500–1
platelets (thrombocytes) 1101
  bone marrow precursor cells 6
  tumour cell–endothelial cell adhesive interactions 1017
    co-culture techniques 1019
platinum-based drugs 1302–8
  combined drug therapy 1334–5
  DNA adduct formation 1295–6
  mechanism of cytotoxicity 1308
pleiotrophin gene, ribozyme targeting 1445
pleomorphic adenoma
  ceruminal gland 844
  lacrimal gland 835
pleomorphic carcinoma
  lacrimal gland 835–6
  lung 518–19
pleomorphic liposarcoma 751
pleomorphic rhabdomyosarcoma 752
pleomorphic xanthoastrocytoma 802–3
pleural fluid 467
pleural tumours 767
  see also mesothelioma
plexiform fibrohistiocytic tumour 748
plexiform neurofibroma 834, 838
pluripotent stem cells 5
  bone marrow 6
PML (promyelocytic leukaemia gene)
  knockout mouse studies 909
  RARα fusion 909, 1106
  tumour-suppressor activity 909
PML protein
  acute promyelocytic leukaemia (APL) pathogenesis 909, 1106–7
  nuclear bodies 909
PMS1 43–4
PMS2 43–4, 95
pneumatocytes, type I 504
pneumatocytes, type II 504
Pneumocystis carinii pneumonia 1291
POEMS syndrome 691
pol 76
Pol β
  DNA base-free site repair 90–1
  mitochondrial DNA repair 94
Pol δ 94

Polo-like kinase (Plk1) 1517–18
poly(ADP-ribose) polymerase (PARP) 99–100, 120–1
  DNA repair processes 1361
  topoisomerase II activity modulation 1317
polyamine metabolism inhibition 429, 440–1, 1426
polychlorinated biphenyls 568
polychlorophenols 414
polyclonal tumours 248–9
polycyclic aromatic hydrocarbons (PAHs) 273, 378, 438, 1086–7
  air pollution 379
  brain tumour models 1129–30
  cancer chemopreventive strategies 410
  dietary 300, 392
  DNA adducts formation 279, 296–300
    exposure assessment 378
  DNA damage 281
  lung cancer 512
    models 407, 409, 1092
  mechanism of carcinogenesis 284–5
  metabolic activation 278, 407
  occupational exposure 415
  tobacco smoke 301, 405, 407, 409–10, 509
polycythaemia 648
polymerase chain reaction (PCR) 473, 692, 717–18
  gene expression change assay techniques 900
  methylation-sensitive 329
  micrometastases detection 475
  transgenic mouse production procedures 892
  see also reverse transcriptase polymerase chain reaction (RT-PCR)
polymorphisms, metabolic
  methodological issues for epidemiological studies 381–2
  phase I metabolism 304
  phase II metabolism 304
  screening 384–5
  susceptibility markers 380–2, 384
    cause–effect relationships 382
    public health applications 382
polynucleotide immunization *see* DNA vaccines
polyoma middle T antigen (PMTA) 1003
polyoma virus 49–50, 962, 965
  brain tumours 794
  haematological malignancies 1104–5
Polyomavirus subfamily 50
polyphenols 429, 436–8
polyprenoic acid 1433
polyunsaturated fatty acids (PUFAs) 395
poly(vinyl) alcohol embolization 1256
Pontin 52 204
porphyria 568
porphyrins 422
porta hepaits clamping (Pringle manoeuvre) with radiofrequency ablation 1263
portal vein 563
positional cloning 1045–6
positive predictive value 462
positron emission tomography (PET) 1215, 1227–32, 1558
  blood flow studies 1231
  cell proliferation rate measurement 957–8
  DNA synthesis markers 1231
  drug markers 1231
  [^{18}F]fluorodeoxyglucose (^{18}FDG) imaging 1227–31
  fluoride bone scans 1231
  neuroreceptor markers 1231–2
  pancreatic cancer 590
  PET camera 1227
  principles 1227
  protein synthesis markers 1231
  thyroid cancer 1220
post-gastric surgery pancreatic cancer risk 589
postirradiation angiosarcoma 755
postmarketing drug surveillance 1041–2
post-transplantation lymphoproliferative disorders 716–18
  central nervous system 812
  clinical features 716–17
  Epstein–Barr virus 58, 716–17
  genetic alterations 718
  monomorphic 717
  pathological classification 717
  polymorphic 717
  sinonasal tract 507
pox viruses 335
pp60c-src 1034
PPAR-γ 604
PPAR-γ ligands 132
PPAR-γ mutations, thyroid follicular carcinomas 607–8
praziquantel 346–7
precursor B cell lymphoblastic leukaemia/lymphoma 695–6
precursor lesions *see* preinvasive/premalignant lesions
precursor T cell lymphoblastic/acute lymphoblastic leukaemia (precursor T cell lymphoblastic lymphoma) 706
pregnancy, breast development 9
preinvasive/premalignant lesions 1085, 250–2
  breast lesions 142–3
  cervical cancer 617
  clonality analysis 249–52
  colorectal cancer 552–4
  detection by screening 462
  lung cancer 512–14
  melanoma 250, 479–80, 1115
  oral cancer (squamous cell carcinoma) 491, 493, 1431
  p53 mutations 251, 253
  penile cancer 676
  preventive interventions 1418–19
  prostate cancer 668
  skin cancer 479–81, 1435
  testicular germ cell tumours 679–80
pretherapeutic cellular DNA repair profiling 1511–12
preventive strategies 259, 421, 1417–18
  blocking agents 421
  carcinogen formation prevention 421
  high-risk cohort studies 1423–4
  primary prevention 1417
  secondary prevention 1417
  suppressing agents 421
  tertiary prevention 1417
  see also antigenotoxins; chemoprevention
primary cell culture 885
primary cutaneous anaplastic large cell lymphoma 711
primary effusion lymphoma 704
  Kaposi sarcoma-associated herpes virus association 65, 341, 704
primitive neuroectodermal tumours 795–800
  electron microscopy 799
  immunohistochemistry 798–9
  posterior fossa (medulloblastoma) 795–8
  supratentorial 798–9
Pringle manoeuvre 1263
prinopastat (AG-3340) 1479
probability 268
procarbazide 301
procarcinogens 277, 294
  metabolic activation 278, 294, 427
    chemopreventive strategies 437–8
pro-caspase-1 124
pro-caspase-2 123

pro-caspase-3 126
pro-caspase-8 123, 124–5
pro-caspase-9
  mitochondrial caspase activation pathway 126–7
  organellar sequestration 127
pro-caspase-10 123, 124–5
pro-caspases 121
  recruitment to receptors 123–4
prodrugs
  antibody-directed prodrug enzyme therapy (ADEPT) 1386–7, 1533
  genetic activation therapy 1533–41
  suicide gene therapy 1069–70
  tumour-associated protease activation 1481
progenitor cells 1084
  haematopoiesis 1102
    cell culture 1111
    in vitro quantitation 1102
progesterone
  breast tissue response 9
  endometrial response 627
progesterone receptor 135
progestins 135
prognostic markers 472–3
programmed cell death 5
  see also apoptosis
progression of cancer 436, 863–70, 937, 1085–6
  angiogenesis 191–2, 235–6, 938–40, 971, 1088
  breast cancer models 865–7
    cell phenotypic diversification 866–7
    cell signalling 866
    genetic changes 865–6
    tumour cell diversification 866–7
  chemical carcinogenesis 288
  E-cadherin regulatory role 1082
  host surveillance mechanisms 869
  intercellular communication 869–70
  metastasis suppressor genes 869–70
  multistage carcinogenesis model 1086–7
  oncogenes 868
  skin cancer *in vitro* model 1087
  tissue organization disruption 217–18
  transforming growth factor-β (TGF-β) 1087
  see also invasion; mestastasis
prolactin
  angiogenesis-inhibiting fragments 242
  breast tissue responses 9
  cell growth regulation 25
prolactinoma 600
proliferating cell nuclear antigen (PCNA) 103, 426, 905–7
  DNA repair
    base-free (abasic) sites 90–1
    mismatch 94
    mitochondrial DNA 94
    nucleotide excision repair 91
  immunohistochemical detection 958
proliferative breast disease 613
proliferative verrucous leukoplakia 494
proline 427
promoting agents
  chemical carcinogens 287–8
    tobacco smoke 407
  concentration threshold 288
  multistage carcinogenesis 287–8, 1086–7
pronuclear injection technique 892–3
prostacyclins 1011
prostaglandin E_2 (PGE_2) 395
prostaglandin H synthetase 278, 443
prostaglandins
  angiogenesis 237
  haematopoiesis inhibition 1104
Prostascint (capromab pendetide) 1223–4
prostate
  anatomy 665
  blood supply 665
  development 665
  zones 665
    cancer susceptibility 666
prostate acid phosphatase 966
  prostate adenocarcinoma 671
prostate cancer 666–75, 1095
  aetiology 672
  androgen dependence 1449
  androgen receptors 1449
    mutations 135
  atypical adenomatous hyperplasia (adenosis) 668–1
  Bcl-2 overexpression 131
  capromab pendetide 1223
  cellular vaccines 1393–4
  chemoprevention 1434
  clinical management 674–5
    hormone therapy 1449
  cytological features 667–8
  detection on digital rectal examination 666
  dietary factors 33, 395, 672–3
  drug resistance 1351
  epidemiology 672–3
  erbB-2/neu expression abnormalities 966
  genetic prodrug activation therapy 1536
  genetic susceptibility 672
  Gleason grading system 471, 666–7
  human papillomavirus (HPV) 337
  immunohistochemistry 671
  magnetic resonance imaging (MRI) 1210
  mesenchymal tumours 671–2
  metastases
    bone 1001, 1216
    middle ear 849
    seminal vesicle 678
  micrometastases detection 475
  models 1095–7
    'Dunning' 1095–6
    gene therapy 1070
    rodent 1095–6
    transgenic 1096–7
    xenografts 1096
  molecular genetics 673
  multidrug resistance proteins (MRPs) expression 1354
  p53 mutations 29
  pathology 666–72
  percutaneous ablative therapy 1261
  precursor lesions 668
  prognostic factors 673–4
  prostatic intraepithelial neoplasia (PIN) 113, 225–8, 668–70
  PTEN mutations 30, 907
  Rb mutations 908
  screening 462, 673
  signal transduction abnormalities 225–8
  small-cell carcinoma 671
  staging 1559
  transrectal ultrasound 672
  ultrasound-guided biopsy 1197–8
prostate intraepithelial neoplasia (PIN) 669–70
  apoptosis suppression 225–8
  telomerase activity 113

prostate lesion biopsy 1247
prostate-specific antigen (PSA) 966
 antiangiogenic activity 1471
 cellular vaccine target 1392–3
 immunohistochemistry 467
 prostate cancer 671–2
  screening 462, 673
 tumour marker 472
protease inhibitors 438–40, 1478–80
 antiangiogenic therapy 1470–1
 local prodrug activation 1481
proteasome-mediated degradation
 β-catenin 196, 199–201
 hypoxia-inducible factors 1470
 oestrogen receptor 142
 therapeutic approaches 1502–3
protein adducts, internal dose markers 377
protein kinase A 162
protein kinase C 170–1, 1087
 antisense oligonucleotide targeting 1443
 chemopreventive strategies 429, 439
protein kinase C inhibitors 1501–2
protein lysate microarrays 232
protein phosphatase 2A (PP2A) 1325
proteoglycans
 basement membrane 210, 1010
 extracellular matrix 210, 1026
  degradation during malignant invasion 231
 haematopoietic cell interactions 1104
 Wnt receptors 199
proto-oncogenes 27, 288, 606, 1525
 chemical-induced damage 288
 cytoplasmic 28
 haematological malignancy-related translocation 1105–6
 in vitro site-directed mutagenesis 963
 nuclear 28
 protein product signal transduction 1525–6
 RNA tumour virus acquisition 27–8
 transforming retroviruses 962
 transgenic mouse studies 903
proximate carcinogens 278, 294
PS-341 1502
psammoma bodies 637–9
ptaquiloside 273–5, 390
PTB
 insulin receptor substrates (IRSs) 166–8
 receptor tyrosine kinase signalling 165–6, 171–2
PTC (*Ptc*) 30
 basal cell naevus syndrome 483–5
 knockout mouse studies 1131
 ret rearrangements, thyroid papillary carcinomas
   605, 607–8
PTCH 39, 820
PTEN 175, 213, 215–16
 normal function 30, 907–8
 knockout mouse studies 907–8
 mutations 30, 131
  associated cancers/syndromes 907
  Cowden disease 820
  epidermal growth factor (EGF) receptor amplification 30
  familial breast cancer 42–3
Pteridium aquilinum (bracken fern) 390
pterygium 826–7
publication bias 380
pulmonary fibrosis 512
pulmonary granuloma 1161
pulmonary hamartoma 509–10
 computed tomography (CT) 1161
pulmonary lymphangitic carcinomatosis 1178
pulmonary lymphatics 504–5
pulmonary lymphoid proliferations 519
pulmonary neuroendocrine tumours 510–11
purine analogues 1279–87
 structure 1279–80
purine nuceloside phosphorylasep-(β-D-2-deoxyerythropentofuranosyl)-
  6-methylpurine prodrug activation system 1539
purines 1271
 salvage pathway 1272
 synthetic pathways 1272
 transport systems 1272
pyrimethamine 1291
pyrimidine analogues 1273–9
pyrimidines 1271
 salvage pathway 1272–3
 synthetic pathways 1272
 transport systems 1272
pyrolizidine alkaloids 273–5, 390

Q
quercetin 391, 424, 1501
quiescence *see* dormancy
quinicrine 1048
quinidine 1048
quinine 1048
quinolones 1048

R
R115777 1498–9
rabbit corneal micropocket assay 242, 975, 980
RAC3 *see* SRC-3
Rac/Rho GTPases 173, 214
Rad1 103
Rad9 103
Rad17 103
 clamp-loader model 103
Rad50 104
 Mre11:Nbs complex 104
Rad51 98–9, 101
 DNA homologous recombination repair 99–100, 1362
Rad52 98–101, 1362
Rad54 99, 101, 1362
Rad checkpoint 103
rad unit 308
radiation exposure 307
 carcinogenesis 311–14
 definitions 304–8
 dose/activity units 308–9
 genotoxicity 309–11
 protection standards 314
 risk assessment 314
 transforming growth factor-β (TGF-β) activation 180
radiation pneumonitis 1181–2
radiofrequency ablation 1261, 1263–4
 chemoembolization combined treatment 1257–9
 hepatocellular carcinoma treatment 1264
 liver metastases treatment 1261–2
 principle 1263
 probes 1263
 Radionics device 1263–4
 Radiotherapeutics device 1263–4
 RITA device 1263

radiofrequency (RF) radiation 308
carcinogenesis studies 313
exposure limits 314
genotoxicity 310–11
sources of exposure 309
radioimmunoguided surgery 1564–5
Radioimmunoguided Surgery System (RIGS) 1565
radioisotopes 1215
decay, ionizing radiation dose/activity units 308
labelled antibodies 1385–6
see also nuclear medicine
Radionics radiofrequency ablation device 1263–4
radioprotective agents 1361
radiosensitizing agents 1361
Radiotherapeutics radiofrequency ablation device 1263–4
radiotherapy 473, 1359–67
apoptosis response 1361, 1363–5, 1367
bladder cancer 655, 660
brain tumours 822
breast cancer 615
cell survival–dose relationship 1361
ceramide induction 1364–5
complications 655, 736–7
CT-guidance 1176
DNA damage 1361, 1367
cellular radiosensitivity 1362–4
DNA repair 1361–2
fractionated dose delivery 1359
inflammatory response 1365
knockout mouse studies 1366
mesothelioma 774
mitochondrial damage 1364
oral cancer 496
postirradiation sarcoma 736–7
prostate cancer 674–5
radioisotope-labelled antibodies 1385–6
soft tissue sarcomas 761–2
treatment aims 1359
tumour sensitivity/resistance 1359
influence of hypoxia 1360–1
urethral tumours 662
radium workers 311–12
radon exposure 307, 309, 311–12, 414–15, 418, 512
Raf kinases 211–12
antisense oligonucleotide targeting 1443
cell growth regulation (Ras targets) 28
Raf-1 129–31
antisense oligonucleotide targeting 1443, 1446
mitotic arrest 1325–6
RAG-1 1361–2
Rag-1 mutations 917–18
RAG-2 1361–2
Rag-2 mutations 917–18
RAIDD 123, 154
raloxifene 1434, 1462
oestrogen receptor interaction 137, 1454, 1460–2
osteoporosis management 1462–3
raltitrexed 1291
rapamycin 1048
RARs *see* retinoic acid receptors
Ras 211–12
cell growth signalling 15, 28
therapeutic approaches 21
farnesylation inhibition 430, 440
GTP/GDP binding 28
Myc combined action in tumour formation 31–2
tyrosine kinase signalling 162–3, 1486
ras mutations 11, 27–8, 31, 299, 436, 606, 1469, 1525
cellular vaccine targets 1393
cholangiocarcinoma 575
correction by gene therapy 1527
gall bladder carcinoma 577
initiation step of carcinogenesis 1087
lung cancer 512
thyroid carcinomas 607–8
to constitutive activity 28
Ras (v-*ras*) 955
Ras-MAP kinase pathway 175, 1498
initiation by Grb2 168
tyrosine kinase signalling 162–3
Rathke's pouch 491
Rb 109, 908, 955, 959, 1364
cell cycle progression regulation 17, 30, 111, 903, 1526
progression across R-point 30, 1514–15
tumour-related defects 30, 1514–15
myoblast terminal differentiation induction 959
see also Rb–E2F pathway
Rb 28–9, 962–3, 967, 1108, 1525
loss of heterozygosity (LOH) 98, 248, 836
mutations
associated cancers 908
bone tumour predisposition 722
central nervous system tumours 1123–4
laryngeal squamous cell carcinoma 508
lung cancer 511–12, 908
osteosarcoma 724
parathyroid carcinoma 607
retinoblastoma 18, 35–6, 826, 836
small-cell lung carcinoma 504
transgenic/knockout mouse studies 908–9
brain tumour models 1131
RB1 (*Rb-1*)
knockout mouse studies 1131
mutations
familial retinoblastoma 737
pancreatic cancer 592–3
soft tissue sarcomas 739–40, 745, 760
Rb–E2F pathway 15–17, 30, 836, 1514
cell cycle regulation 15–17
papillomavirus E7 protein interaction 53, 337, 620, 626–7
Tax protein effects 82–3
RBP56/hTAFII68 738–9
reactive nitrogen oxide species 391, 395
reactive oxygen species 396
DNA adduct detection 298–9
DNA damage 89, 422, 1507–8
ionizing radiation generation 1360
mitochondrial production 94
REAL classification 506–7, 694, 812
receptor protein tyrosine phosphatases (RPTPs) 1034
receptor tyrosine kinases 162, 1486
activation 171
antiangiogenic therapy targets 1468
binary 171–4
cell growth regulation 25, 1487
cellular/subcellular localization 171
developmental regulation 169
distribution 169
focal adhesions 173
functional domains 162, 1486
inhibitors 1491–2, 1500–1
ligand mimics 171
ligands
availability/release 170–1

differential distribution 169
receptor interactions 169, 1486
signal transduction mechanism 165–8, 1486
signalling
growth factors 953, 1486
regulation 170–1
Smads 185–6
versatility 169–70
structure 169–70
superfamily 169, 1486
T cell receptor activation 172–3
therapeutic targeting 1486
see also growth factor receptor blockade
turnover 170–1
VEGF binding in angiogenesis 980
recombinant Trail protein (Apo2L) 132
recombinant-interleukin 2 (IL-2) *see* interleukin 2 (IL-2) therapy
RecQ1 100
RecQ4 (RecQL4) 100
RecQ5 100
RecQ-family helicases
Bloom syndrome 100–1
DNA repair 89–103
rectal cancer *see* colorectal cancer
rectum
development 545
gross appearance 545
transanal endoscopic microsurgical approaches 1561–2
recurrence 1036–7
reduced folate carrier (RFC) 1288–9
reduction reactions 277, 427
Reed–Sternberg cells 340, 693, 713–16
Reichert's cartilage (second branchial arch) 841
relative risk 260–2, 266
rem 308
renal angiomyolipoma 1161
renal cell carcinoma 646, 648–53, 820
aetiology 649–50
2-acetylaminofluorene 282
chemoembolization 1256, 1260
chromophobe 651–3
classification 648–9
clinical management 653
collecting duct 651–2
conventional (clear cell) carcinoma 649–53
electron microscopy 652
epidemiology 649
familial forms 649
Fuhrman nuclear grading 649
imaging 1149
immunohistochemistry 652
interleukin 2 (IL-2) therapy 1375
metastases
false-negative bone scan 1218
middle ear 849
molecular genetics 652–3
nutritional influences 394–5
papillary 651–3
paraneoplastic syndromes 648
pathology 648–9
gross features 650–1
microscopic features 651–2
preoperative computed tomography (CT) 1176
presentation 648
prognostic factors 653
sarcomatoid dedifferentiation 652
stem cell transplantation 1546–7
telomerase activity 113
TNM staging 649, 653
TSC2 mutations 30
unclassified 652
vimentin expression 469
renal cysts 820
computed tomography (CT) 1160–1
ultrasound 1197
renal malformations 573
renal pelvis 645
tumours 402, 653–4
renal tumours
computed tomography (CT) 1160–1, 1197
magnetic resonance imaging (MRI) 1206
multidrug resistance proteins (MRPs) expression 1354
percutaneous ablative therapy 1261
percutaneous biopsy 1247, 1250
ultrasound 1197–8
reovirus 1071–2
repairosome 94–5
replication protein A (RPA) 99–100, 100–1, 1363
DNA base-free (abasic) site repair 90–1
nucleotide excision repair 91
reporter genes for knock-in mouse studies 905
representational difference analysis (RDA) 62–4, 341
residual disease detection, telomerase activity 116
resistant starch 549, 552
respiratory epithelium 503–4
mucociliary escalator 504
see also Schneiderian membrane
respiratory tract anatomy 503–5
restin 242, 1469
antineoplastic/antiangiogenic activity 1476
restriction endonucleases 1522–3
restriction fragment length polymorphism (RFLP) 35–6
clonality assessment from X chromosome inactivation 606, 247–8
tumour-suppressor gene identification 31
restriction point (R-point) 14, 15–17, 18, 1512
cell cycle regulation 18, 1514–15
misregulation in cancer 19, 1514–15
resveratrol 424, 436
ret mutations 601
medullary thyroid carcinoma 601, 606
phaeochromocytoma 606
ret/PTC rearrangements, thyroid papillary carcinomas 605, 607–8
reticuloendothelial system 3
retina 825–6
retinal tumours 828–32
clinical management 838
retinal vascularization 977, 980
retinoblastoma 28–9, 35–6, 826
clinical management 838
differential diagnosis 829
endophytic 826
epidemiology 826
exophytic 826, 828–9
experimental induction 1128–9
extraocular extension 830
familial form 35, 737, 836, 907
genetic aspects 18, 35–6, 826, 836
loss of heterozygosity (LOH) 248, 836
metastasis 830, 837
mixed endophytic–exophytic 829
osteosarcoma association 19, 36, 724, 837
pathology 35, 828–30
prognostic factors 837
radiotherapy complications 827

retinoblastoma – *continued*
Rb knockout mouse studies 908–9
Rb mutations 18, 35–6, 826, 836
screening 826, 836
spontaneous form 35, 836, 908
spontaneous regression 829
retinoic acid 966–7
clinical aspects 1414–15
differentiation regulation 429, 959–60, 1106–7
retinoic acid receptor (*RAR*) 495
chromosomal locus 1314
retinoic acid receptor α *(RAR*α*)*
acute promyelocytic leukaemia 909
PML fusion 909, 1106–7
retinoic acid receptor $\beta 2$ *(RAR*$\beta 2$*)* 325, 331
promoter hypermethylation 325
retinoic acid receptors (RARs) 1409–10
retinoid receptor selective ligands 1426
retinoid X receptor 1409, 1411
PPAR heterodimers 1407–8, 1409
vitamin D receptor complexes 1408–9
retinoids 966–7
apoptosis regulation 132
chemoprevention 429–30, 436, 439, 440–1, 1434–5
head and neck tumours 496
liver cancer 1433
oral leucoplakia regression 1427–31
differentiation induction therapy 1409–10
combination treatment 1411–12
gene expression modulation 1409–10
interferon combined treatment 1411
clinical trials 1415
therapeutic approaches 22, 132, 1478
vitamin D_3 combined treatment 1411
retinol *see* vitamin A
retinopathy of prematurity 977
animal models 977
retroperitoneal structures, percutaneous biopsy 1248
retroviral vectors 1481
antisense oligonucleotide/ribozyme delivery 1442–3
DNA vaccines 1394
expression vectors 965, 1526–7, 1528
retroviruses 75, 965, 1525
cultured cell transformation 962
endogenous proviruses 76
genome structure 76
host-derived oncogene carriage 76
life cycle 75–6
long terminal repeating sequences (LTRs) 75–6
provirus integration into cellular DNA 75–6, 965
replication 75–6, 965
reverse mutation assays 352
reverse transcriptase 75–6, 965
reverse transcriptase polymerase chain reaction (RT-PCR) 746
rfa mutations 352
RFC 94
RGD sequence 211, 242–3
disintegrins 1477
therapeutic potential 1477
cytotoxic drug targeting 1481
rhabdoid tumour
extrarenal malignant 759–60
immunohistochemistry 745, 759–60
rhabdomyoma 777
rhabdomyosarcoma 466, 733, 752–4
bladder 654
chromosomal translocations 753
classification 849
clinical management 754, 849
heart 781
immunohistochemistry 745, 752–3
larynx 508
liver 576
middle ear 849
nose 503
orbit 834–5, 838
p53 mutations 29
prognostic factors 754, 849
staging 753
ultrastruture 746
rheumatological disease 519, 718
rhizoxin 1329
rhodamine 123 1048
rhodamine isothiocyanate 1031–2
rhodostomin 1477
riboflavin *see* vitamin B_2
ribonuclease P 1440
ribonucleotide reductase 1292
ribozymes 1439, 1440–41
clinical studies 1446
delivery 1442–3
design 1441–2
formulation 1442
gene expression inhibition 1441
hairpin 1440–1, 1445
hammerhead 1440, 1445
preclinical studies 1444–5
Richter syndrome 697
Rick 124–6
RIP 124, 154
RIP 13 (NCoR) 140–1
RIP 140 1453
RIP-2 (Cardiac) 124
risk assessment
DNA adduct evaluation 303
radiation exposure 314
risk ratio 260
RITA radiofrequency ablation device 1263
rituximab (anti-CD20) 472–3, 1371, 1382–3, 1488–9
clinical trials 1575–6
radioisotope-labelled 1385–6
RNA base nomenclature 1271–2
RNA enzymes *see* ribozymes
RNA polymerase II 88, 91–2
RNA tumour viruses 75–85
brain tumour animal models 1129
cellular sequences acquisition 27–8
haematological malignancies 1104–5
oncogenesis 27
see also retroviruses
RNA-based vaccines 1059–60, 1394
robotic surgical systems 1565–6
rodent carcinogenicity studies 365–73
alternatives to second rodent study 368–9
analysis 365–6
assay validation 369–70
dose selection 370–1
dose-limiting pharmacodynamic endpoint 371
genotoxic compounds 370–1
limit dose 371
maximum feasible dose 371
maximum tolerated dose (MTD) 365, 370
pharmcokinetic endpoint 370–1
saturation of absorption 371

false-positive findings 366–7
ICH guidance 366, 367–72
S1B 367–70
S1C 370–1
weight of evidence assessment 371–2
need for long-term studies 367–8
reproducibility 366–7
rodent strains 365, 367
study design 365–6
testing in 2 rodent species 367–8
tumour site concordance 367
validation 365
rodent chow diets 920
rodent fibroblast culture 961–2, 967
neoplastic transformation 961–2
oncogene expression assays 963–4
transformation assays 964
rodent inbred strains 365, 367
genetic drift 367
spontaneous tumour incidence 367
breast cancer models 924–5, 927
rodent mammary tumour induction *see* breast cancer models
rodent ulcer *see* basal cell carcinoma
rofecoxib 443, 1425, 1432
roller bottle culture 884–5
Rothmund–Thomson syndrome
bone tumour predisposition 722
RecQ-family helicase defects 100
round cell liposarcoma 469–70, 750
Rous, P. 25, 335, 1524
Rous sarcoma virus 76
experimental brain tumour induction 1128–9
oncogenic potential 27, 76
v-src 27, 76, 161
rutin 424

S

S phase 13–14, 17, 953–5, 1324
antimetabolite activity 1271
cell fraction labelling 957–8
checkpoint 18
chemotherapy target 20–1
topoisomerase II activity 1317
S100 gene family 869
S635 1129
S16020-2 1356
mechanism of action 1316
S-100 protein 466–7, 469
acoustic neuromas 846
brain tumours 1123–4
cartilage 722
chondrosarcoma 725–6
chordoma 727–8, 811–12
choroid plexus papilloma 808
dysembryoplastic neuroepithelial tumour 806–7
endolymphatic sac papillary tumour 848
Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
jugulotympanic paraganglioma 846
Langerhans cell histiocytosis 730
neuroblastic tumours 815
neurofibroma 815–16
salivary gland carcinoma 498
sentinel lymph node biopsy 471–2
soft tissue tumours 745
squamous cell carcinoma 481
synovial sarcoma 758
uveal melanoma 834
S-antigen
brain tumours 795
primitive neuroectodermal tumours 798–9
saccharin 655
saccule 841
safrole 391
saline infusion sonohystography 1193–4
salivary gland carcinoma 496–500
aetiology 498
clinical management 500
epidemiology 496–8
grading 498, 500
immunohistochemistry 498
lymphoepithelial 340–1
molecular genetics 498
pathology 498
presentation 498
staging 499–500
salivary gland development 491
Salmonella typhimurium mutation assay 352, 353–4
test strains 353
salmosin 1477
salt intake, dietary 395–6, 529–1
sample size 381–2
SAOS-2 1117–18
sarcoma 228–9, 465
aetiology/pathogenesis 736–7
associated genetic disorders 737
chromosomal translocations 470, 473–4
clinical management 760–3
local disease 760–1
locally recurrent disease 761–2
metastatic disease 762–3
epidemiology 736
grading 741–2, 760
heart *see* cardiac sarcoma
hepatic metastases 1196–7
immunohistochemistry 745
inert materials implantation induction 275–6
keratin expression 469
liver 564, 575–6
childhood sarcomas 576
lung 519
MDR1 expression 1351
minimal residual disease (MRD) detection 475
models 1117–19
molecular genetics 737–40
chromosomal translocations 737–9
tumour-suppressor genes/oncogene alterations 739–40
p53 knockout mouse studies 907
penis 676
prognostic factors 760
prostate 671–2
radiotherapy-induced 827
seminal vesicle 678
staging 741–3
uncertain differentiation 757–60
uterus 629–30, 634–5
viral induction in animal models 962
sarcomatoid dedifferentiation, renal cell carcinoma 652
SAS 750
soft tissue sarcomas 740
sassafras oil 391
satumomab pendetide (Oncoscint) 1223

satumomab pendetide (Oncoscint) – *continued*
sensitivity 1223–4
saturated fats 395
scaffold proteins 200–1
extracellular matrix 214–15
scatter factor 1083–4
SCH6636 1498
Schiller–Duval bodies 813
Schistosoma haematobium 345
bladder cancer association 345, 660
Schistosoma japonicum 345
liver/colorectal cancer association 345
Schistosoma mansomi 345
schistosomiasis 345–7
bladder cancer 335, 345, 660
chemotherapy 346
colorectal cancer 345
control measures 347
epidemiology 345
liver cancer 345
parasite life cycle 345
Schneiderian membrane 505
Schneiderian papilloma 505
Schwann cells 789
schwannoma 792, 816–17, 819–20, 1123
Antoni A pattern 816–17, 834
Antoni B pattern 816–17, 834
orbit 834
pancreas 587
scid mouse studies 1044
melanoma models 1116–17
metastasis models 1003, 1005
model characteristics 917
xenograft recipients 915–16
oncolytic reovirus therapy 1072
SCID-hu metastasis models 1003
scintimammography 1225
imaging protocol 1225
indications 1225
Sck 173
sclera 825–6
sclerosing epithelioid sarcoma 747
screening 462–3, 1417
breast cancer 462, 612–3, 1235–6
colorectal cancer 550–2, 1158–9,
computed tomography (CT) 1158–9
costs 462–3
gastric cancer, Japanese programme 533
genetic susceptibility markers 384–5, 462
lung cancer 1158–9
occupational cancers 419
oesophageal cancer 532–3
organ-confined invasive cancers 462
ovarian cancer 1195
precursor lesions detection 462
suitable tumours 462
test positive predictive value 462
test sensitivity/specificity 462
urothelial carcinoma 655–6
screening antitumour compounds 1043–5
cell-based drug screen 1044
highthroughput strategy 1045
historical models 1043
hollow-fibre assay 1044
mean graph development 1043
molecular targeting 1044
SRB assay 1044
xenograft models 918–20
dose selection 919
endpoints for assessing activity 919–20
toxicology 919
XTT assay 1044
scribble 215–6
scrotal cancer 273, 413, 1086–7
scrotal mass 1192
sebaceous adenoma 550
eyelid 828
sebaceous cell carcinoma 826
eyelid 828, 837
sebaceous gland tumours 828
second primary tumours, clonality analysis 250–1
SEK1 *see* MKK4
selectins 984–5, 1016–17, 1034
selective oestrogen receptor modulators 1460–3
selenium
chemoprevention 1431–2, 1434
dietary 424, 549, 672–3
glutathione peroxidase dependence 428, 438
immune response effects 441
self-renewing tissues 4, 7–8
cancer origins 10–11
semicircular canals 841
seminal vesicle
development 677
structure 677
seminal vesicle cancer 677–81
epithelial–stromal tumours 678
pathology 677–8
prognostic factors 678
sarcoma 678
secondary tumours 678
seminoma 465–6, 679–80
clinical management 685
immunohistochemistry 683
pathology 680–1
gross appearances 680
miscoscopy 680–1
radiotherapy sensitivity 1359
spermatocytic 681
testis 1192–3
semustine (methyl-CCNU) 1301
senescence (cellular senescence) 17, 109, 960, 1087
carcinogenesis defences 113
telomere shortening 17, 109–11
senkirkine 390
sensitivity of screening test 462
sentinel lymph node sampling 1232–3, 471–2, 1558–9, 1563
accuracy 1233
breast cancer 1232–3, 1563
melanoma 1232, 1563
penile cancer 677
principle 1232
Seocalcitol (EB 1089) 1414
serial analysis of gene expression (SAGE) 900
pancreatic cancer 591
serine proteinases 231–2
serine-threonine kinases 162
cell growth regulation (Ras targets) 28
serotonin 524–5, 595
Sertoli cell tumour
immunohistochemistry 683
testis 682–3
Sertoli cells 678–9
multidrug resistance proteins (MRPs) 1352

Sertoli–Leydig cell tumour 640
serum-free tissue culture media 881
Sevenless 161
severe combined immunodeficiency (SCID) 97
  γ_c-cytokine receptor/JAK–STAT signalling defects 149
  stem cell transplantation 1545, 1548
sex cords 678–9
sex-cord stromal tumours
  clinical management 642, 685
  molecular genetics 641
  ovary 635, 642
  pathology 640
  testis 682–3
Sézary syndrome 486
Sézary–Lutzner cells 486
SH2 162, 173–4, 953, 1410–11, 1487
  receptor tyrosine kinase signalling 165–6, 166–8, 171–3
SH3 162, 173–4, 1410–11
Shc 173, 212–13 1487
SHH (sonic hedgehog) 483–5, 820
shiitake (*Cortinellus shiitake*) 391
SHP 140
SHP1 173–4
SHP2 166–8
sialyl-LewisSXs 985
siamois 203–4
sickle cell disease 1548
sievert (Sy) 308
14–3-3-σ (*STRATIFIN*) promoter hypermethylation 324–5
signal transduction 5, 225–8
  cell adhesion molecules 983–4
  chemokine receptors 154–7
  chemopreventive strategies 429, 439
  receptor tyrosine kinases 164–8, 1486
    adaptor molecules 166
    docking proteins 166–168
    enzymes 166
    ligand activation 164–5
    tyrosine phosphorylation 164–6
  see also signalling pathways
signal transduction inhibitors 1497–503
  clinical trial design 1497
  combined therapeutic approach 1497–8
  genomic targets 1502
  rationale 1497
signalling pathways
  angiogenesis models 980
  apoptosis pathway interactions 131–2
  breast cancer progession models 866
  cell cycle regulation 15
  cytokines 147–57
  differentiation regulation 959–60
  epidermal growth factor receptor (EGFR) 1487
  growth factors 953
  integrins 211–12, 985
  interferons 150–2
  metastasis formation 1055–6
  radiotherapy responses 1364–6
  steroid receptors *see* steroid receptor signalling
  therapeutic targeting 1497–502
  tumour necrosis factor receptors 153–4
  tyrosine kinases 162–4, 1487
  see also signal transduction
silencer of death domains (SODD) 124–6, 154
silica 418
simian sarcoma virus 1128–9
simian T cell leukaemia virus (STLV) 77
simian virus 40 (SV40) 19, 49–50, 335, 962, 965
  brain tumours 794
  haematological malignancies 1104–5
  mesothelioma risk 767–9, 771
Sin3A 320–1
single photon emission computed tomography (SPECT) 1215
  bone scanning 1216
single-chain antibodies, growth factor receptor blockade 1492
single-nucleotide polymorphisms (SNPs) 380
  disease susceptibility associations 380
  high-throughput technologies 377
sinonasal carcinoma 505
  occupational exposure-related 417
  small-cell carcinoma 506
  undifferentiated 506
sinonasal lymphoma 507
  T cell/natural killer (NK) cell 507
Sipple syndrome *see* multiple endocrine neoplasia 2 (MEN-2)
size of tumour 1086
Sjögren's syndrome 519
SKC 1044
Ski 184–5, 190
skin
  anatomy 477–8
  development 477
skin cancer 477–88
  aetiology 478
  chemical carcinogens 273, 478
  clinical management 488
  epidemiology 264, 266
    skin/eye pigmentation association 478, 482
  metastatic 477
  models 1086–90
    chemical carcinogen/UV radiation 1116–17
    chemical carcinogenesis in mouse skin 1086–7
    in vitro tumour progression 1087–8
  molecular genetics 487
  pathology 478–87
  precursor lesions 479–81, 1435
  prevention 479
    chemoprevention 1435
  prognostic factors 487–8
  telomerase activity 113
  ultraviolet radiation exposure association 313, 478
skin tumours, benign 477
  pathology 478
Smad1 183, 185–6
Smad2 182, 183–6, 204
  tumour-associated mutations 188, 190
Smad3 182–6, 190, 204
  inactivating mutations 188–90
Smad4 183–5, 204
  tumour suppressor function 188–90
  tumour-associated mutations 188–190, 547
    see also DPC4
Smad5 183
Smad6 183, 186–7
Smad7 182–6
Smad8 183
Smads
  BMP receptor signalling 183
  MH1 domain 183
  MH2 domain 183, 188
  mutations in tumours 11, 188
  subfamilies 183
  transforming growth factor-β (TGF-β) receptor signalling 183–6
    CBP/p300 transcriptional co-activator 183–4, 184–5

Smads
  transforming growth factor-β (TGF-β) receptor signalling – *continued*
    cross-talk with other signalling pathways 185–6, 204
    transcription activation 183–4
    transcription factor interactions 184–6
    type I receptors 183
  tumour suppressor function 188–90
small intestine tissue organization 7–8
small lymphocytic lymphoma 1382–3
  B type (B chronic lymphocytic leukaemia) 696–7
    genetic features 696–7
    immunophenotype 696
  liver involvement 1101
small-cell carcinoma 599–600
  cervical cancer 617
  endocrine pancreas 587
  lung *see* small-cell lung cancer
  prostate 671
  sinonasal region 506
small-cell lung cancer 510, 600
  chromosomal abnormalities 512
  clinical presentation 515
  drug resistance 1351
  epidemiology 511
  immunohistochemistry 518
  Lambert–Eaton myasthenic syndrome 516
  metastases 515, 1001
  nofetumomab merpentan 1223
  oncogene changes 512
  p53 mutations 512
  paraneoplastic syndromes 515
  pathology 517–18
  RB mutations 512
  treatment 518
Smn 129–30
smooth muscle actin
  atypical teratoid/rhabdoid tumour 800
  leiomyosarcoma 751–2
  rhabdomyosarcoma 752–3
  salivary gland carcinoma 498
  soft tissue tumours 745
SMRT (TRAC2) 140–1
SN-38 1319
snuff/snuff-dipping 399, 403
socio-cultural risk factors 263
SODD (silencer of death domains) 124–6, 154
sodium chloride, dietary intake 395–6, 529–31
sodium diclofenac 1425, 1435
sodium nitrite 391–2
soft tissue tumours 733–63
  adipose tissue differentiation 749–51
  aetiology/pathogenesis 736–7
  benign 734
  classification 734
  clinicopathological features 746–60
  diagnosis 743–6
    biopsy 744–5
    clinical history 743
    histological assessment 745
    imaging 743–4
    immunohistochemistry 745
    molecular genetics 746
    physical examination 743
    thallium-201 scanning 1221–2
    ultrastrutural analysis 746
  endothelial/pericytic differentiation 754–5
  epidemiology 736
  fibroblastic/myofibroblastic 746–8
  fibrohistiocytic 748–9
  genetic disorders 737
  larynx 508
  magnetic resonance imaging (MRI) 1210
  malignant 734
  molecular genetics 737–40
  multidrug resistance proteins (MRPs) expression 1354
  nasopharynx 503
  oncogenesis 733–4
  osseous/cartilagenous differentiation 757
  peripheral neuroectodermal differentiation 755–7
  plain radiography 1148
    abdomen 1149
  prognostic factors 745
  skeletal muscle differentiation 752–4
  smooth muscle differentiation 751–2
  see also sarcoma
soft tissues 733
softwood dusts 505
solar keratosis *see* actinic keratosis
solar radiation *see* ultraviolet radiation
solitary eosinophilic granuloma 730
solitary fibrous tumour 743
  immunohistochemistry 745
  nasopharynx 503
solitary myeloma (plasmacytoma of bone) 730
solitary pulmonary nodule 509
  computed tomography (CT) 1158–9, 1161–2
  diagnostic thoracoscopy 1563
  ^{18}FDG positron emission tomography (PET) 1228
solutions, tissue culture 878–83
solvents 568
somatic mutation
  clonality analysis 247
  theory of carcinogenesis 967–8
somatostatin 524–5, 580
somatostatin analogues
  neuroendocrine tumour management 609
  radiolabelled
    diagnosis/imaging 608
    therapeutic use 609
somatostatin receptors 608–9
somatostatinomas 587, 595
  clinical features 586
sonic hedgehog (SHH) 483–5, 820
soot 273
SOS–Grb2 complex 162–3, 168, 211–12, 985
Soto syndrome 568
Southern blot analysis 35–6, 473, 692, 717–18
  transgenic mouse production procedure 892
soy sauce 392
SP-1 (Sp1) 137, 184, 186
specificity of screening test 462
spectral karyotyping (SKY) 899–900
sphingosine 1477–8
spinal tumours 1203
spindle assembly checkpoint 1517
spindle cell oral carcinoma 494
spinous layer of epidermis 8
splenic mass, percutaneous biopsy 1247–50
Sporomiella 1491–2
sporostatin 1491–2
squalamine 1471
squamous cell carcinoma 465
  anal canal 547
  cervix 617

clinical management 488
eyelids 827, 837–8
inflammatory cytokines expression 1088–9
keratin pearls 1403
larynx 508
lung 516–17
middle ear 848
models 1086–90
chemical carcinogenesis in mouse skin 1086–7
knockout mouse 1090
transgenic 1089
tumour progression 1088
xenografts 1088
nasopharynx 505
oesophagus 525, 528–5, 537–42
oral cavity 491–6
penis 676
skin 313, 477–8
aetiological factors 478
carcinoma *in situ* 481
histological variants 481
infectious agents 347–9
pathology 481–2
prognosis 487–8
urethra 661–2
squamous dysplasia
bronchial 513
penile preinvasive lesions 676
squamous papilloma
larynx/oropharynx 507
nasopharynx 505
trachea 507
SRB assay 1043–4
method 1044
src 27, 161, 1525
mutations 1314–15
breast cancer 27–8
colon cancer 27–8
v-src 27, 76, 955
Src protein kinases 161–2, 170, 211
focal adhesions 173
K^+ ion channel function 175
regulation 173–4
T cell receptor signalling 172–3
SRC proteins
nuclear receptor (NR) box (receptor-interacting domain) 139
oestrogen receptor coactivator function 138–140, 1451, 1453
proteasome-mediated degradation 142
structure 138–9
SRC-1 (ERAP-160; NcoA-1) 138–40, 142, 1453
SRC-2 (GRIP1; NcoA-2; TIF-2) 138, 140, 1452
SRC-3 (ATCR; RAC3;p/CIP; AIB1; TRAM-1) 138–40
SSX/SYT fusion 758, 770
ST1571 1547
staging 471–2
computed tomography (CT) 1162–6
laparoscopy 1559
surgical treatment planning 1558
thoracoscopy 1563
stapes 841–2
STAT1 150–2
STAT2 152
STAT3 149
STAT4 149
STAT5 149
STAT6 149
STATs 149, 1410–11
knockout mouse studies 149
signalling
binary receptor tyrosine kinasess 171–2
TGF-β signalling convergence 185–6
see also JAK–STAT signalling pathway
statistical analysis
phase III clinical drug trials 1572
xenograft studies 914–5
Stauffer syndrome 743
stauroporin 17, 1477–8, 1499
Stein–Leventhal syndrome 1139
stem cell transplantation 1545–53
allogenic 1546–8
graft versus host disease 1548
HLA mismatched 1546, 1548–9
indications 1546–8
non-malignant disease 1546–7
peripheral blood stem cells 1546
T cell depletion 1546
autoimmune disease management 1553
autologous 1545–6
conditioning regimens 1546
relapse management 1551
cord blood transplantation 1549–50
graft versus leukaemia effect 1550–2
graft versus tumour effect 1552
haploidentical 1550
graft versus host disease 1550
T cell depletion 1550
historical development 1545
host versus graft transplantation tolerance 1552
nonmyeloablative 1552–3
unrelated 1548–9
stem cells 4–5, 959, 1084
bone marrow 6, 10, 1103
breast tissue 9
carcinogenetic processes 11
clonogenic 7–8
differentiation studies 964
division process 5
embryonic *see* embryonic stem cells
epidermis 8, 9
haematopoiesis 1102–3
bone marrow ‘homing’ 1104
cell culture 1111
haematopoietic 1103
lymphoid 1103
liver 10
lymphoid 6
mesenchymal 10
myeloid 6
neuronal 10
pluripotent 5
proliferation rates 252–3
properties 5
small intestinal crypts 7–8, 252–3
totipotent 4
stents
biliary 595
colonic 1562
endobronchial 1176
endoluminal 1176
endovascular 1176
stereotactic surgery 1564
breast lesions 1564
neurosurgery 1564
sterigmatocystin 390

sterilization
  autoclave 876–7, 879
  tissue culture solutions 879
steroid cell tumours 600
  clinical management 609
  ovary 640
  pathology 603
  prognosis 608
steroid hormone-producing cells 599
steroid hormones 135, 599
  cell growth regulation 17, 25
  P-glycoprotein interaction 1048
  steroid receptor interactions 135
steroid receptors 135
  coactivator/corepressor interactions 138, 140
  conserved functional domains 135
  ligand-binding pocket 137
  signalling 135–44
  steroid hormone interactions 135
  transcription modulation 135
steroid response element (SRE) 135
Stewart–Treves syndrome 737
STK11 mutations 547
STK11/LKB1 mutations, pancreatic cancer 588–9, 592–3
stomach
  anatomy 524–5
  development 524
  digestive processes 524
  mucosal endocrine cells 524
  regions 524
STRAP 182–3
stratum basalis 477–8
stratum corneum 8
stratum spinosum 477–8
strawberry naevus 834
streak gonad 684
Streptomyces 1333
Streptomyces antibioticus 1280
Streptomyces caespitosus 1341
Streptomyces parvulus 1340
Streptomyces peucetius 1333–4
Streptomyces verticillus 1344
streptomycin 882, 1044
streptozotocin 595, 609
stromal cells 1025
  heterogeneity 1029
stromal interactions 1025–37
  contact-dependent 1031–4
  epithelial growth 1030–1
  hormone responsive tumours 1029–30
  in vitro models 1029
  metastasis 1034–7
  reciprocal stromal–epithelial 1031
stromal matrix 210
stromelysin 1 1034–6
stromelysin 3 1034–6
stromelysins 231–2, 1030–1
  malignant invastion 1034–6
Sturge–Weber syndrome (encephalo-trigeminal angiomatosis) 836–7
styrene 414
SU101 (lefluonomide) 1468, 1500–1
SU5416 1468, 1501
SU6668 1468, 1501
subculture (passaging) 883–4
subependymal giant cell astrocytoma 818
subependymoma 805
suberoylanilide hydroxamic acid (SAHA) 1407
  clinical studies 1412–13
subglottic carcinoma 508
subglottis 503
subgroup analysis 381–2
sublingual glands 491
submandibular glands 491
suicide gene therapy 1069–70
  brain tumours 1135
  bystander effect 1069–70
  immune response 1070
suicide genes 1069
sulforaphane 438
sulindac 423, 428, 430–2, 549, 552
sulphides, allium vegetables 424, 438
sulphorhodamine B *see* SRB assay
sunburn 1088
sunlight exposure 249–50
  occupational 414–5, 418
  skin cancer association 313, 478
  see also ultraviolet radiation
superficial cerebral astrocytoma *see* desmoplastic infantile ganglioglioma/astrocytoma
superoxide dismutase 422
suppressing agents 421, 429–32
supraglottic carcinoma 508
  CT follow-up 1177–8
supraglottis 503
suramin 1467–8, 1477
surfactant 504
surfactant apoprotein A 504
surgery 473, 1557–66
  clinical staging 1558
  curative therapy 1557–8
  palliative therapy 1557
  principles 1557–9
    operative 1558–9
    tumour dispersal prevention 1558
  radioimmunoguided 1564–5
  stereotactic 1564
surrogate endpoint biomarkers (SEBs) 1420–3
survival as end-point 919–20
survivin 122–3, 558
susceptibility markers 377, 380–4, 1423–4
  allele frequency–risk relationship 384
  cancer-prone mutations 19, 35–44
  DNA repair 382–3
  metabolic polymorphisms 381–2, 384
  screening strategies 384–5
  tobacco smoke carcinogen metabolism 401–2
suspension culture 884
  haematological cells 1109
  logarithmic-phase growth 888
  screening antitumour compounds 1044
SW480 1076
Syk 172–3
sympathetic nervous system neuroblastic tumours 794
symphytine 390
Symphytum officinale (comfrey) 390
synaptophysin 467
  brain tumours 795, 1123–4
  endolymphatic sac papillary tumour 848
  Ewing sarcoma/peripheral neuroectodermal tumour of bone (PNET) 726–7, 755–6
  islet cell tumours 586
  jugulotympanic paraganglioma 846
  neuroendocrine tumours 602
  primitive neuroectodermal tumours 798–9

small-cell carcinoma 671
cervix 617
lung 518
synchronous tumours, clonality analysis 250–1
syndecan-1 (syndecan; CD138) 214–15, 220
multiple myeloma 729–30
syndrome of inappropriate antidiuretic hormone secretion 515
synergistic drug combinations 1051
clinical protocol design 1052–3
combination index (CI) 1052
median-effect plots 1051–2
quantitation 1051–2
selectivity 1052
therapeutic goals 1051
synovial sarcoma 469, 758, 760
chromosomal translocations 470, 758
heart 781, 785
immunohistochemistry 745, 758
synpolydactyly 10
systemic lymphangiomatosis 723–4
SYT/SSX 758, 770

T
t(1;13)(p36;q14) 753
t(2;5)(p23;q35) 711
t(2;13)(q35;q14) 753
t(3;3)(p14;q29) 785
t(8;14)(q24.1;q32) 704
t(9;14)(p13;q32) 702
t(9;17)(q22;q11) 757
t(9;22)(q22;q12) 757
t(11;14)(q13;q32) 692, 698
t(11;18)(q21;q21) 700–1
t(11;22)(p13;q12) 759
t(11;22)(q24;q12) 738, 756
t(12;15)(p13;q25) 747
t(12;16)(q13;p11) 748–50
t(12;22)(q13;q12) 750, 757
t(14;18) 837
t(14;18)(q32;q21) 692, 699–700
diffuse large cell lymphoma 692
follicular lymphoma 692
t(17;22)(q22;q13) 748
t(21;22)(q22;q12) 756
t(X;17)(p11.2;q25) 759
t(X;18)(p11.2;q11.2) 758
T antigen 965, 967
transgenic brain tumour models 1131
T cell lymphoma 706–12
angioimmunoblastic 708–9
central nervous system 812
cutaneous *see* mycosis fungoides
enteropathy-type 712
Epstein–Barr virus 60–1, 340–1
hepatosplenic γ/λ 708
intestinal tract 548
lymphoblastic/acute lymphoblastic leukaemia (precursor T cell lymphoblastic lymphoma) 706
p53 knockout mice 907
peripheral 706–12
not otherwise specified 709–10
TGF-β TβRI mutations 186–7
WHO classification 694
see also adult T cell leukaemia/lymphoma; T cell/natural killer (NK) cell lymphoma
T cell receptor 172
activation 172–3
lymphoma 473, 692
precursor T cell lymphoblastic 706
rearrangements 706, 1361–2
structural aspects 172–3
T cell/natural killer (NK) cell lymphoma 706–12
nasal type 711–12
sinonasal tumours 507
T cells 172, 1056
γ_c cytokine secretion 147
activation 1056–7, 1389–90
antibody-based anti-tumour activity enhancement 1384
bone marrow precursor cells 6
co-stimulation 1056, 1060
depletion for stem cell transplantation 1546, 1550
graft versus host disease 1548
isolation from blood 1108–9
lymph node follicles 689–90
T-brain 214–5
TACE (TNF-α-converting enzyme) 152
taenia coli 545
TAF_{II} 30 1452–3
TAG 636, 983
TAK-1 (TGF-β-activated kinase) 204
takin 215
tal-1 1107
talc 636
talin 214, 985
talipes 573
tamoxifen 20, 746–7, 921, 1460–2, 1501–2
breast cancer 615, 1449, 1451, 1454–5, 1458–9
oestrogen receptor status 144
preventive treatment of high-risk women 1421–3, 1427, 1433–4, 1458, 1460
carcinogenic activity 276–7, 294–5
DNA adduct formation 296–7, 298–9
endometrial abnormalities assessment 1193–4
hepatic metabolism 294–5
mode of action 132, 439
oestrogen receptor interaction 439, 1454, 1460–2
resistance 140–1, 144, 1459
retinoids combined treatment 1411
side effects 276–7, 1434, 1460
endometrial polyp induction 627–8
endometrial tumour induction 632, 1139–40
tangeritin 424
targeting vectors 903–4
tarragon oil 391
TAS-103 1320–1
Tax protein 78–9
adult T cell leukaemia pathogenesis 83–4, 342, 1108
apoptosis avoidance 84
cell proliferation 83
fixation of genetic abnormality 83–4
expression *in vivo* 84
pleiotropic effects 83–4
trans-repression of transcription 81–2
transcriptional activation 80
trans-activation 80–81, 83
transcriptional inhibitors down-regulation 81
tumour-suppressor proteins inhibition 82–3
taxanes
combined drug therapy 1334–5
irinotecan sulphate (CPT-11) 1320
Vinca alkaloids 1331
Pgp/MDR1 transport 1328

taxol *see* paclitaxel
^{99m}Tc imaging 1215
  bone scanning 1216
  intracerebral 1220–1
  scintimammography 1225
  sentinel node localization 1232–3
  thyroid 1218
^{99m}Tc labelled monoclonal antibodies 1223
^{99m}Tc MIBI imaging 1220–1
^{99m}Tc pertechnate 1218
^{99m}Tc sestamibi 1225
^{99m}Tc sulphur colloid 1232–3
Tcf-4 556–7
TCF/LEF transcription regulators 206, 215
  Wnt signalling 196, 204
TE-85 1117–18
tea 436, 450, 655, 1435, 1482
TEC-11 1471
telomerase 19, 107–17, 495, 962–3, 967
  cancer *in vitro* progression models 117
  cancer-related activity 113–16
    diagnostic/prognostic applications 116
    telomerase-negative tumours 115–16
  hepatocellular carcinoma 571
  hTERT 110–11, 117
    immunohistochemical detection 116
  hTR 110–11, 117
    in situ hybridization 113–16
  overexpression in immortalized cell lines 19, 1087
  pancreatic cancer 590–1
  properties 109–11
  ribozyme targeting 1445
  therapeutic inhibition 116–17
  tissue distribution 110
  TRAP assay (telomere repeat amplification protocol) 111–12, 115–16
  urothelial carcinoma screening 656
telomere 107
  functions 107
    chromosomal integrity 107–8
    divisional clock 108–9
    end replication problem 108
  length 108–9
    ALT lengthening mechanism 115–16
    crisis 109
    tumours 109, 113–15
  structural aspects 107
temozolomide 1076–7
temporal bone 841–2
temporal bone tumours
  classification 842
  meningioma 847
  metastatic tumours 849
tenascin 8, 210
  antibody-mediated targeted drug delivery 1481
  integrins binding 211
tenascin-C 1026
  breast cancer-related expression 1026
teniposide 1316
tensin 173
TEP1 see PTEN
teratoma
  central nervous system 813
  differentiation 1403
  heart 777
  immature 640, 813
  immunohistochemistry 683
  mature 640, 813
    see also dermoid cyst
  nasopharynx 503
  orbit 835
  ovary 640
  testis 1192–3, 682
terminal deoxynucleotidyl transferase (TdT) 696
terminal differentiation 959, 1403–4
  epidermal cells 8–9, 477
  therapeutic approaches 966–7
    preventive strategies 429, 441
terpenoids 424
testicular cancer
  clinical management 685
  distant metastasis 1193
  epidemiology 683
  magnetic resonance imaging (MRI) 1210
  molecular genetics 684–5
  multidrug resistance proteins (MRPs) expression 1354
  prognostic factors 685
  risk factors 683–4
  ultrasound 1192–3
testicular teratoma 1192–3, 682
  p53 knockout mouse studies 907
testicular tumours 679–85
  immunohistochemistry 683
  metastatic 1193
  pathology 679–83
testis
  development 678–9
  structure 679
testosterone 276–7
  differentiation regulation 959–60
  mammary gland development 1026–7
2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) 287, 298, 414–15
tetracycline derivatives 1480
tetradecanoylphorbol acetate (TPA) 287, 1116
tetraethylenepentamine (TEPA) 1549–50
tetrahydrofolate 1287, 1290
Tetrahymena thermophila 1440
TFE3/ASPL fusion 759
TFIIH 91, 93
TFIIs 137–8, 1452–3
tfm (testicular feminization) 1027
TG.AC v-Ha-ras mouse model 368–9
thalassaemia major 1545–8
thalidomide 1467, 1548
thallium-201 scanning 1220–2
  brain tumours 1221
    interpretation 1221
    procedure 1221
  indications 1220
  musculoskeletal tumours 1221–2
    interpretation 1221
    procedure 1221
  principle 1220
thecoma 640
therapy 473
thermal ablation therapy 1261, 1565
thioguanine 1284
  mechanism of cytotoxicity 1287
  metabolism 1284–7
  resistance 1287
    DNA mismatch repair system defects 95
  retinoids combined treatment 1411
thioproline 427
thiopurine methyltransferase 1287
6-thiopurines 1284–7

thio-tepa (triethylenethiophosphoramide) 660, 1298
DNA adduct formation 1298
thoracoscopy 1562
diagnostic 1562–3
resection for malignancy 1563
staging 1563
thorium dioxide *see* Thorotrast
Thorotrast 311–12
angiosarcoma 575
cholangiocarcinoma 574
hepatocellular carcinoma 568
soft tissue sarcomas 737
urothelial carcinoma 653–4
thrombin 231, 1019
thrombocytes *see* platelets
thrombomodulin 657
thrombopoetin 1103
thrombospondin 1011–12
angiogenesis 971, 1088
therapeutic strategies 1470
antineoplastic synthetic peptides 1476
integrin binding 242, 1026
transforming growth factor-β (TGF-β) activation 180
thymic (mediastinal) lymphoma 704
thymidine kinase 354, 1276
thymidine labelling 957–8
thymidine phosphorylase 237, 239, 244
thymidine pulse labelling index (TLI) 957–8
thymidylate kinase 1272–3
thymidylate synthase 1273–6
antitumour compound molecular targeting 1044
thyroglobulin 467
thyroid hormone metabolism 599
thyroid anaplastic carcinoma 600–2
pathology 605–6
thyroid biopsy 1247
thyroid cancer 600
epidemiology 600–1
^{131}I imaging 1218, 1220
ionizing radiation exposure 311–12
metastases
cervical lymph nodes 1190–2
false-negative bone scan 1218
molecular genetics 607–8
multidrug resistance proteins (MRPs) expression 1353
screening 601–2
ultrasound 1190–2
recurrent disease detection 1192
thyroid follicular adenoma 600
Gsα mutations 606–7
pathology 604
thyroid follicular carcinoma 600
aetiology 601
clinical management 609
clonality analysis 251
invasive behaviour 604
pathology 604–5
PPAR-γ mutations 607–8
prognostic factors 608
thyroid follicular tumours 603–5
nodules 604
hyperplasia 604
pathology 604
treatment 609
thyroid hormone receptor, retinoid X receptor interactions 1409
thyroid hormone-producing cells 599
thyroid hormones 276–7, 599
differentiation regulation 1106–7
thyroid insular carcinoma 599, 605–6
thyroid lymphoma 601–2
thyroid metastatic carcinoma 601–2
thyroid nodular goitre 605
clonality analysis 251
thyroid nodule
colloid 601–2
follicular 604
^{131}I imaging 1218
papillary hyperplastic 605
percutaneous ablative therapy 1261
^{99m}Tc pertechnate imaging 1218
ultrasound 1218
thyroid papillary carcinoma 600–2, 605
clonality analysis 251
pathology 605
tall cell variant 605
thyroid papillary lesions 605
benign neoplasms 605
hyperplasia 605
hyperplastic nodule 605
thyroiditis 601–2
tiazofurin 1411
Tie-1 1469
Tie-2 244, 1469
TIF-2 *see* SRC-2
tight (occludens) junctions 9, 215, 983, 1081–2
TIMP-1 241
TIMP-2 241
TIMP-3 241
TIMP-4 241
tingible body macrophages 689
tissue culture 873–88
applications 886
biological limitations 887–8
cell cycle synchronization 888
logarithmic-phase growth 888
media/serum effects 887
monodispersed cells 887
passage number 887
trypsin-induced artefacts 887
cell line selection criteria 886
cell lines 883
in vivo/in vitro assays 886
normal cells maintenance 886
tissue culture equipment 873–8
autoclave 876–7
biological safety cabinet 873–6
cell counter 878
culture dishes/flasks 878
frozen cell storage facilities 877
incubator 877
inverted microscope 877
pipetting devices/pipettes 878
sterile items 878
warm room 878
tissue culture media 873, 880–3
amino acids 880
antibiotics/antimycotics 882
buffering capacity 881
cell freezing solutions 882–3
cell removal solutions 882
coenzymes 880
glucose/carbohydrate 880
ingredients 880
inorganic salts 880

tissue culture media – *continued*
lipid sources 881
nucleic acid derivatives 881
reducing agents 881
salt solutions 879–80
serum 881
serum-free 881–2
vitamins 880
tissue culture solutions 878–83
salt solutions 879–80
sterilization 879
water 878–9
tissue culture techniques 883–6
basic culture 883
cell counting 884
cell freezing 884
cloning 885–6
large scale/high-density culture 884–5
primary culture 885
subculture (passaging) 883–4
suspension cultures 884
viability examination 884
tissue fixation 464
tissue inhibitor of metalloproteinases 1 (TIMP-1) 591
tissue inhibitor of metalloproteinases 2 (TIMP-2) 232
tissue inhibitors of metalloproteinases (TIMPs) 558, 626–7, 1478
angiogenesis 231, 241
malignant invasion/metastasis 231–2
pancreatic cancer 591
tissue organization 209
carcinogenesis 10–11
control 10
extracellular matrix function 216–18
polarity establishment 215–16
tissue plasminogen activator (tPA) 558
angiogenesis 241
malignant invasion/metastasis 231–2
tissue processing 464
tissue regeneration capacity
renewing tissues 5
stem cells 4–5
tissues 3–11
classification 3–4
titanium 275
TK (thymidine kinase) 354
TLS/*CHOP* fusion 750
TMP3/*TMP4*/*ALK* rearrangements 747–8
TNM staging 471
bladder cancer 658
gastric cancer 541
oesophageal cancer 537
oral squamous cell carcinoma 496
renal cell carcinoma 649, 653
renal pelvis/ureter urothelial tumours 653–4
urethral tumours 661–2
tobacco smoke 18, 377–8, 389, 509, 1081, 1085–7
animal studies 403–4
carcinogens 301, 405, 509
chemistry 404–7
particulate phase 404–5
cigarette smoke condensate (CSC) 404
contribution to occupational cancers 417–18
environmental exposure 402–3, 417–18
exposure assessment
arylamine–haemoglobin adducts 377–8
DNA adducts 378–9
NNK metabolites 408
mainstream 40–5
occupational exposure 414–15, 418
sidestream 402–5
tumour promoting activity 407
tobacco, smokeless/chewed 399
chemical composition 407
oral cancer (squamous cell carcinoma) 403, 478, 481, 492
verrucous variant 494
tobacco smoking 33, 260, 262–3, 336–7, 342, 399–411, 509, 794
aerodigestive tract cancers 402
animal studies 403–4
bladder cancer 402, 407, 655
bronchopulmonary carcinoid tumours 510
cancer chemoprevention 410–11
cervical cancer 620
cessation 409, 1417
colon cancer 549
commercial aspects 399–400
DNA adduct formation 299–301, 407–9
genetic polymorphisms 304
lung cancer risk assessment 303
endometrial cancer risk reduction 632
epidemiology 399
cancer causation 401–3
related cancer deaths 400–1
hepatocellular carcinoma 421, 565–6
laryngeal keratosis 507–8
laryngeal squamous cell carcinoma 508
lung cancer 264–6, 400–3, 407, 511–12, 1431
duration of exposure 384, 400
mechanism of tumour induction 407–10
nasopharyngeal carcinoma 505–6
oesophageal cancer 528
oral cancer (squamous cell carcinoma) 402, 407, 492–3
p53 mutations 299, 407–9
dysplastic tracheobronchial mucosa 251
pancreatic cancer 402, 407, 511, 588
passive smoking 260, 402–3, 509
renal carcinoma 402, 649–50
reverse smoking (chutta) 492
salivary gland carcinoma 498
smoked products 399
TOC mutations 541
tongue cancer 492
tonsils 503
TOP-53 1316
topetecan 1318
topoisomerase I 1278, 1313, 1334
gene locus 1314–15
levels, CPT-11 resistance 1319–20
SN-38 inhibition 1540
topoisomerase I inhibitors 1317–20
determinants of cellular sensitivity 1319–20
resistance 1319–20
topoisomerase II 1313, 1315, 1334
gene loci 1314–15
Her2/*neu* coexpression in breast cancer 1314–5
topoisomerase IIα 1314–15
heat shock response 1316–17
molecular genetics 1314–15
selective inhibitors 1315–16
topoisomerase IIβ 1314–15
molecular genetics 1314–15
selective inhibitors 1315–16
topoisomerase II inhibitors 1315–17, 1339
categories 1315
determinants of cellular sensitivity 1316–17

p53 protein activity 1317
topoisomerase inhibitors 95, 361, 1313–21, 1334
dual inhibitors 1320–1
Herceptin combined therapy 1314–15
topoisomerase I 1317–20
topoisomerase II 1315–17
topoisomerases 1046, 1313, 1334
classes in mammalian cells 1313–14
gene loci 1314
tositumomab, radioiodine-labelled 1385
total calorific intake 394–5
Toxoplasma gondii 1291
toxoplasmosis 1220–1
TRAC2 (SMRT) 140–1
trace elements 423–4
trachea 504
tracheal tumours 509
TRADD (Tradd) 123, 153–4
TRAF2 interaction 154
TRAF2 153–4
TRADD interaction 154
TRAF3 153–4
TRAF5 153–4
TRAFs 153–4
TRAIL ligand 152
Trail protein (Apo2L)
'decoy' receptors 124–6
therapeutic approaches 132
trail receptor-1 (DR4; Apo2) 123, 124–6
trail receptor-2 (DR5) 123, 124–6, 131–2
TRAM-1 *see* SRC-3
TRAMP (transgenic adenocarcinoma mouse prostate) 1096–7
trans-catheter therapy *see* chemoembolization
transanal endoscopic microsurgery 1561–2
transcription 1522
activation
JAK–STAT signalling pathway 148
oestrogen receptor 136–42
p53 protein 896
TGF-β receptor Smad signalling 183–4
tyrosine kinases 162–3
myc regulation 28
regulation by DNA methylation 319–20
gene silencing 321–2
molecular modulators 320–2
regulation by histone modifications 317–20
steroid hormone receptor modulation 135
transcription factor IIH (TFIIH) 91–3
transcription factors 141, 959
antimicrotubule agent high concentration effects 1327
differentiation regulation 960
growth factor activation 953
homeobox gene regulation 10
Smad interactions 184–6
TGF-β regulation of expression 181
transcription-coupled DNA repair (TCR) 91–4, 1510
defect in Cockayne syndrome 92–93
transcutaneous CT-guided needle biopsy 1166–9
transducisome 214
transfection 1523
cell culture studies 886
transferrin 1012
transformation 26–7, 1084
breast cancer models 865–6
cultured cells 962–3, 967
rodent fibroblasts 961–2, 967
gap junction-mediated intercellular communication 1032
in vitro assays 964
oncogenes 28, 962–3, 967
transformation competent genes 965–6
tumour antigen expression 1389
tumour viruses 27, 962
transforming growth factor-α (TGF-α)
angiogenesis 237, 1490
breast tissue regulation 9–10
tumour cells 25
epidermal growth factor receptor (EGFR) interaction 1486–7
hepatocellular carcinoma 571
oesophageal adenocarcinoma 542
pancreatic cancer 1488
proliferation stimulation 1488
small intestinal crypts 8
transforming growth factor-β (TGF-β) 14, 179–92, 1074, 1124
activation 180
angiogenesis 237, 971, 1028–9, 1072–3
cellular responses 180–1
cell proliferation inhibition (growth arrest) 180–1, 186
colon/gastric cancer 96
cyclin-dependent kinase (cdk) inhibitor induction 180–1
differentiation regulation 959–60
endothelial cell production 1012–13
expression 180
extracellular matrix interactions 180
extracellular matrix protein expression regulation 181
gap junction-mediated intercellular communication 1032
gene expression regulation 181
haematopoiesis inhibition 1104
hepatocellular carcinoma 571
invasion/metastasis 232
latent complex 180
radiotherapy responses 1365
small intestinal crypts proliferation inhibition 8
structure 179
TGF-β receptor complex 182
tumour development/cancer progression stimulation 186–190, 1087
microenvironment effects 191–2
transforming growth factor-β1 (TGF-β1) 179
expression in tumour cells 180, 190–1
local immunosuppression in tumorigenesis 191
TGF-β receptor interactions 182
tumour angiogenesis induction 191
transforming growth factor-β2 (TGF-β2) 179
TGF-β receptor interactions 182
transforming growth factor-β3 (TGF-β3) 179
TGF-β receptor interactions 182
transforming growth factor-β (TGF-β) receptors 182–3
ALK-1/TSR-1 182
ALK-2/ActR1/Tsk7L 182
cancer-related mutations
colon cancer 96
head and neck cancer 495
pancreatic cancer 592–3
nonSmad signalling 186
Smad signalling 183–6
TβRI 182, 186–7
TβRII 182–3
DNA mismatch repair 96
expression in carcinomas 187–8
HNPCC-associated somatic mutations 186–7, 550
TGF-β complex 182
protein interactions 182–3
signalling 182
type I 182, 183
type II 182

transforming growth factor-β (TGF-β) superfamily 179
transgenic models 865–6, 891–903
  brain tumours 1131–3
  breast cancer 1094–5
  chemopreventive agent/antigenotoxin testing 425, 445, 1426
  gene therapy models 1523
  genotoxicity testing 357–8
  haematological malignancy 1110–11
  historical aspects 891–2
  limitations 901
  melanomas 1116
  metastasis models 1003
  oncogenes 895–6, 899
  p21bcr/abl expression 1110–11
  polyoma middle T antigen (PMTA) 1003
  production by pronuclear injection 892–3
    founder mouse production efficiency 892
    large DNA fragments expression 892–3
  production via homologous recombination in ES cells 893–5
    DNA transfer by electroporation 893–4
    ES cell injection into blastocytst 893–4
    gene targeting 894–5
    germ-line transmission 893–4
    screening procedures 895
  prostate cancer 1096–7
  squamous cell carcinoma 1089–90
  tumorigenesis 365, 368–9, 895–6, 899
  tumour-suppressor genes 896–7, 899, 903
transglottic carcinoma 508
transit cells 4–5
  amplifying 5
  dividing 4–5
  small intestinal epithelium 7–8
transitional cell carcinoma *see* urothelial carcinoma
transitional cell ovarian tumour (Brenner tumours) 637, 639–40
translation 1522
transplatin 1303
transrectal ultrasound 1558
  prosate cancer 672–3
transthoracic CT-guided needle biopsy 1166
transvaginal ultrasound
  endometrial cancer screening 632–3
  female pelvis 1193
  ovarian cancer 642
  uterine cancer 1193
TRAP assay (telomere repeat amplification protocol) 111–12, 115–16
  diagnostic/prognostic applications 116
trastuzumab (Herceptin) 21, 467, 472–3, 594, 615, 1371, 1383, 1488–9, 1497, 1500
  clinical trials 1575–6
  cytotoxic drug combined therapy 1313–14, 1489
  side effects 1383
tretinoin 1435
tributyrin 1406
  clinical studies 1413
trichoepithelioma 827
trichofolliculoma 827
tricholemmoma 827
trichostatin A 1407
  all-*trans*-retinoic acid combined treatment 1411–12
triethylenemelamine 1298–9
triethylenethiophosphoramide *see* thio-tepa
trifluoperazine 1048
trigone 645–6
trimethoprim 1291
trimetrexate 1291
TRIP-1 182–3
troglitazone 1407–8, 1413–14
trypsin-EDTA solution 882
trypsin-induced artefacts 887
tryptorelin 1455
TSC1 39–40, 818, 836–7
TSC2 30, 818, 836–7
TSG101 30
TSRI265 1477
Tsussilago farfara 390
tuberin 836–7
tuberous sclerosis (Bournveille disease) 39–40, 44, 836–7
  cardiac rhabdomyoma 777
  genetic aspects 39–40, 818, 836–7
  pathology 818
tubulins 1323–4
  mutations 1328
tubulogenesis model 977–9
tumorigenesis see carcinogenesis
tumour antigens
  cellular vaccine targets 1392–4
  immune response 1389–90
tumour cell lines
  cellular vaccines 1391
  major histocompatibility (MHC) molecule expression 1056–7
  metastasis models 1056–7
    sublines 1057
    see also Esb/Esb-MP cells
  multidrug resistance proteins (MRPs) expression 1353
  P-glycoprotein expression 1351
  tumour-associated antigens 1056–7
tumour growth delay assays 919
tumour markers 472–3, 966
  bladder cancer 659–60
  erbB-2/neu 966
  ovarian cancer 636
  pancreatic cancer 590–1
  see also biomarkers
tumour microenvironment 191–2, 1063, 1088, 1485–6
tumour necrosis factor α (TNF-α) 152–3, 1360–1
  angiogenesis 229, 237, 239, 971, 1028–9
  antitumour immune response 938, 1074
  apoptosis induction 153
  differentiation induction 1410
    retinoid combined treatment 1411
  endothelial cell activation 1013–4
  integrin expression inhibition 1477
  invasion/metastasis 1034–6
  radiotherapy responses 1364–5
  strutural aspects 152
  therapy 1372, 1376–7
    toxicity 1376–7
    vitamin D_3 combined treatment 1412
  vascular targeting 1471
tumour necrosis factor α (TNF-α)-converting enzyme (TACE) 152
tumour necrosis factor β (TNF-β; lymphtoxin-α) 152–3
  angiogenesis 229
  strutural aspects 152–3
tumour necrosis factor receptor 1 (TNFR1) 123, 124–6, 153
  apoptosis induction 153–4
  TRADD interaction 153
tumour necrosis factor receptor 2 (TNFR2) 153
  apoptosis induction 153–4
  cell proliferation regulation 153–4
  TRAF interactions 153–4
tumour necrosis factor receptor superfamily 152
tumour necrosis factor (TNF) receptors 153
  apoptosis triggering 123–6, 1047–8

adapter protein binding 123
tumour resistance mechanisms 124–6
'decoy' receptors 124–6
protein-interaction domain 123
signal transduction 153–4
tumour viruses
brain tumour animal models 1128–9
cultured cell transformation 962
haematological malignancies 1104–5
historical aspects 27
ribozyme targeting 1445
see also DNA tumour viruses; RNA tumour viruses
tumour-associated antigens 1056, 1380
cellular vaccine targets 1392–4
metastasis model 1057–8
tumour cell line expression 1056–7
tumour-associated macrophages 239
tumour-infiltrating lymphocytes (TIL)
interleukin 2 (IL-2) combined therapy 1372, 1375–6
tumour cell interactions 938
mathematical modelling 941–2
tumour-suppressor genes 19, 28–31, 35, 41, 195–6, 288–9, 335, 891, 955, 1524–5
alteration/mutation-related cancers 29
central nervous system tumours 1123–4
haematological malignancy 1105, 1108
head and neck cancers 494–5
hepatocellular carcinoma 571
lymphoma 691
pancreatic cancer 591–3
post-transplantation lymphoproliferative disorders 718
carcinogenesis 30, 293
chemical-induced 288–9
chemopreventive intervention targets 436, 441–2
gene targeting studies 896–7
gene therapy 1070–1, 1524
historical aspects 28
identification techniques 31
knockout mouse studies 891, 903–5
mismatch repair genes 30
mutations 25–27, 195, 204–5, 1524
frame-shift 94
normal cellular function 30
protein product signal transduction 1525–6
recessivity 28, 955, 1524
replacement 1526–7
clinical applications 1527
preclinical studies 1526–7
rationale 1526
silencing by DNA methylation 321–2
nickel carcinogenesis 328
transgenic mouse studies 896–7, 903
tumour–tissue interactions 1025–37
tumstatin 1476
tweak (DR3; weasle) 124–6
twin 203–4
two-hit hypothesis 31, 35
see also multistep carcinogenesis
TYK2 148, 152
tylosis 40, 541
tympanic cavity *see* middle ear
tympanic membrane 841–2
tyramine 392
tyrosinaemia 568
hepatocellular carcinoma 566
tyrosine kinases
cell growth-related phosphorylation 27, 953
cellular functions 175
cytoplasmic proto-oncogene coding 28
historical background 161–2
nonreceptor 162, 174–5
functional domains 162
role in cancer 175–6, 1488
signalling 161–76
mechanism 162–4
superfamily 162
transcription activation 162–3
see also receptor tyrosine kinases
tyrosine phosphatases 170, 173, 467
chemokine receptor signal transduction 156
PTEN function 907
T cell receptor 172

U

ubiquitination
β-catenin regulation 196, 201
cyclins 1513
oestrogen receptor 142
see also proteasome-mediated degradation
UCN-01 1499
UCN-1010 21–3
UDP-glucuronic acid 427
UDP-glucuronyl transferase 427
ulcerative colitis 249–50
cholangiocarcinoma association 574
colorectal cancer association 550, 557
endoscopic screening 551
colorectal dysplasia 554
extrahepatic bile duct tumour association 577
ulcers, chronic 478, 481
Ulex europaeus lectin
collecting duct carcinoma 652
vascular neoplasms 810–11
ultimate carcinogens 277, 294
ultrabiothorax 203–4
ultrasound 308, 1147, 1149, 1187–98
activity units 308–9
biliary tumours 1197
biological effects 308
cavitation-related tissue damage 311
biopsy guidance 1190
breast lesions 1190, 1238–41
screening applications 1236
carcinogenesis studies 313
computed tomography comparison 1188
contrast agents 1189–90, 1198
cyrosurgical ablation guidance 1265
diagnostic applications 1190–8
Doppler 1189
harmonic imaging 1187
hepatic tumours 1195–7
intraoperative 1197
nephroblastoma (Wilms' tumour) 647
ovarian cancer 1194–5
percutaneous guidance
ablative therapy 1262–3
alcohol injection 1264
biopsy 1247, 1250
microwave coagulation therapy 1265
principles 1187–9
protection standards 314
real-time imaging 1188–9

ultrasound – *continued*
renal tumours 1197–8
soft tissue tumours 743–4
sources of exposure 309
spatial resolution 1187–8
staging applications 1558–9
technological advances 1198
testicular cancer 1192–3
three-dimensional techniques 1198
thyroid cancer 1190–2
transducers
frequency 1187–8
miniature probes 1198
uterine cancer 1193–4
see also endoscopic ultrasound; transrectal ultrasound; transvaginal ultrasound
ultraviolet radiation 33, 307, 1081
biological effects 308
brain tumour models 1128
carcinogenesis 311, 313–14
DNA adduct formation 298–9
DNA damage 89, 310, 481–2, 1508
cyclobutane pyrimidine dimer formation 89, 310
DNA photoproduct repair 91, 280, 310
dose/activity units 308
extracellular matrix effects 1088–9
eye/ocular adnexal tumours 826–7
fluence 308
lip cancer 493
melanoma induction 33, 313, 478–9, 485, 1115
in vivo models 1116–17
preventive strategies 421, 479
protection standards 314
skin cancer 264, 478–9
basal cell carcinoma 482
squamous cell carcinoma 481–2
sources of exposure 309
sunburn 1088
wavelength 308
umab (fully human antibodies) 1380
UMP–CMP kinase 1272–3
umuDC-containing plasmids 352–3
upper gastrointestinal tract tumours 523–43
endoscopic resection 1562
uracil glycosylase 90
uranium miners 311–12, 512
ureter
anatomy 645
development 645
ureteric tumours 653–4
urethane 1092
breast cancer models 925
urethra
anatomy 646
development 645
urethral tumours 660–2
clinical management 662
female urethra 660–1
clinical features 660–1
histology 661
prognosis 661
staging 661
male urethra 661–2
clinical features 662
histology 662
prognosis 662
staging 662
uric acid 422
uridine phosphorylase 1273
urinary bladder *see* bladder
urinary tract development 645
urinary tract infection 655
urinary tract tumours 645–62
telomerase activity 113
urine cytology
bladder cancer follow-up 660
DNA adducts detection 301
urothelial carcinoma screening 655–6
urokinase 231, 558
urokinase plasminogen activator (uPA) 231–2, 1034–6
urokinase plasminogen activator receptor (uPA-R)
angiogenesis 241
therapeutic strategies 1470–1
malignant invasion/metastasis 232
uroplakin 657
urothelial carcinoma 654, 656–8
aetiological factors 654
arylamines/aniline dyes 33, 655
cyclophosphamide 655
phenacetin 655
radiation 655
tobacco 655
urinary tract infection 655
bladder 654–60, 1488
clinical management 660
epidemiology 654
flat lesions 657–8
grade 656–8
immunohistochemistry 657
kidney 646
micropapillary lesions 658–9
molecular genetics 657–8
occupational exposure 655
papillary neoplasms (papillomas) 657–9
pathology 654
gross features 656
microscopic features 656–9
prognostic factors 658–60
renal pelvis/ureter 653–4
sarcomatoid 657–9
screening 655–6
small cell 658–9
TNM staging 653–4, 658
tumour markers 659–60
urethra 661–2
urothelial carcinoma *in situ* 657–9
vascular invasion 659
urothelial carcinoma in situ 657–9
urothelium (transitional epithelium) 645–6
immunohistochemistry 657
Usurpin 124–6
uterine leiomyoma 634
pathology 629–30
uterine leiomyosarcoma 634–5
pathology 629–30
uterine tumours 627
endometrial biopsy 1193
nutritional influences 394–5
stromal tumours 634–5
ultrasound 1193–4
transabdominal 1193
transvaginal 1193
utricel 841
UV index 308

UVA
biological effects 308
melanoma association 478
UVB
biological effects 308
ozone absorption 309
skin cancer association 478
UVC, biological effects 308
uveal melanoma 826
choroid 833, 838
ciliary body 833
classification 833–4
epidemiology 826
genetic aspects 826
immunohistochemistry 834
iris 832–3
metastasis 837
predisposing lesions 826–7
prognostic factors 837
racial pigmentation associations 827
uveal naevi, progression to melanoma 826–7
screening 827
uveal tract 825–6
uveal tract tumours 832–4
clinical management 838
uvrB deletions 352–3

V

V79 354
v-*abl* 963–4
v-*erbB* 176, 955
v-*myc* 955
v-*raf* 955
v-*ras* 962–3, 967
v-*sis* 955
v-*src* 27, 76, 161
vaccines 1389–99
active specific immunotherapy (ASI) 1061–2
cell-based *see* cellular vaccines
genetic *see* DNA vaccines
immune response-related strategies 1389–91
live tumour cells 1060
pancreatic cancer 594
principles 1059
RNA-based 1059–60, 1394
tumour cell–dendritic cell fusion products 1060–1
vaccinia viruses 1071–2
vagina 617
vaginal cancer 337
valinomycin 1048
vascular addresses 236
vascular cell adhesion molecules (VCAMs) 244–5
malignant invasion 230
vascular endothelial cell growth factor (VEGF) 836–7, 1360–1
angiogenesis 238–9, 243, 939, 971, 999–1000, 1028, 1072–3, 1088, 1469
signalling pathways 980
therapeutic approaches 21, 1073, 1384, 1470
antiangiogenic therapy
coupled toxins 1466
epidermal growth factor receptor (EGFR) blockade 1490
monoclonal antibodies 1466
antisense oligonucleotide targeting 1443
hypoxia-related expression 239, 1470
knockout mouse studies 238
receptor tyrosine kinase binding 980
retinal vascularization 977
vascular endothelial cell growth factor A (VEGF-A) 238
vascular endothelial cell growth factor B (VEGF-B) 238
vascular endothelial cell growth factor C (VEGF-C) 238
vascular endothelial cell growth factor D (VEGF-D) 238
vascular endothelial cell growth factor receptor 1 (VEGFRI-1; flt-1) 238
antiangiogenic therapy 1468
hypoxia response elements 239
knockout mouse studies 239
VEGF interaction 980
vascular endothelial cell growth factor receptor 2 (VEGFRI-2; flk-1; KDR) 238
antiangiogenic therapy 1468
antisense oligodeoxynucleotide (ODN) interactions 1441–2
knockout mouse studies 239
VEGF interaction 980
vascular endothelial cell growth factor receptors 3 (VEGFRI-3; flt-4) 238
vascular endothelial cell growth factor receptors
receptor tyrosine kinase inhibitors 1501
tumour endothelium 1465
vascular malformations, central nervous system 810
vascular permeability factor (VPF) 229
vascular spread 999–1000
vascular targeting 1465, 1471
vascular tissue
development 733
morphology 236
tumours 236–7
therapeutic strategies 237
vasculogenesis 235
vasostatin 242, 1469
Vasp 215
Vav 173, 1487
V(D)J recombination
DNA double-stranded break formation/repair 96–7
non-homologous end joining 97
Rag-1/*Rag-2* mutations 917–18
scid mouse model 917
vegetables *see* fruit and vegetables
verapamil 1048, 1338, 1350
Verluma (nofetumomab merpentan) 1223
Verner–Morrison syndrome 586
verrucous carcinoma
bladder 660
oral cavity 494
vertebral body haemangioma 1261
vesiculo-vacuolar organelles 1010–11
vestibule 841
VHL 30, 820, 836–7
head and neck cancers 495
methylation in tumours 31
renal cell carcinoma-related changes 652–3
vidarabine 1281–3
video-assisted thoracoscopic surgery (VATS) 1562
solitary pulmonary nodule excision 1563
videomicroscopy *see* intravital videomicroscopy
villi 7
vimentin 466–7, 469
angiomatoid fibrous histiocytoma 748–9
atypical teratoid/rhabdoid tumour 800
brain tumours 795
chordoma 811–12
choroid plexus papilloma 808
endolymphatic sac papillary tumour 848
epithelioid sarcoma 758–9
fibrosarcoma 747
meningioma 810, 847

vimentin – *continued*
nephroblastoma (Wilms' tumour) 648
renal cell carcinoma 652
soft tissue tumours 745
squamous cell carcinoma 481
synovial sarcoma 758
vinblastine 958, 1329
combined drug treatment 1327–8
mechanism of action 1323, 1329
antitumour compound screening 1043
resistance 1048, 1353–5
HEp2 cell line 1351
side effects 1329
Vinca alkaloids 21
combined drug therapy 1334–5
historical aspects 1329
mechanism of action 1329
resistance 1076, 1353
taxane combined treatment 1331
transport 1328
vincristine 756, 958, 1329
mechanism of action 1323, 1329
resistance 1047, 1353–5
HEp2 cell line 1351
side effects 1321
vinculin 173, 214–15, 985
vinorelbine 1329
docetaxel combination 1331
mechanism of action 1323, 1329
paclitaxel combination 1331
vinyl chloride 737, 783
angiosarcoma of liver 413, 575
hepatocellular carcinoma 568
mechanism of carcinogenicity 286
DNA adduct formation 296–7, 300–1
vinzolidine 1329
VIPomas 587
clinical features 586, 595
viral vectors 1523
adeno-associated viruses 1442–3
adenovirus *see* adenovirus
antisense oligonucleotide/ribozyme delivery 1442–3
DNA vaccines 1395
gene therapy 1522, 1527–8
brain tumours 1135
endostatin angiogenesis inhibition 1073
p53 496, 1070–1
thymidine kinase–ganciclovir system 1536
retroviruses 965, 1442–3, 1481, 1526–8
virtual colonoscopy 1558
Virtual Human Project 1565
virus-like particle (VLP)-based vaccines 337
vitamin A (retinol)
antigenotoxic properties 423–4, 440–1
deficiency 441, 528
early lung cancer treatment 1432
skin cancer chemoprevention 1435
vitamin B_2 (riboflavin)
antigenotoxic properties 424
deficiency 528
vitamin B_6, antigenotoxic properties 424
vitamin B_{11} *see* folic acid
vitamin B_{12}, antigenotoxic properties 424, 432
vitamin C (ascorbic acid) 530, 1432
antigenotoxic properties 424, 427–8, 437–8
vitamin D deficiency 672–3
vitamin D receptor 1408–9
vitamin D_3
retinoids combined treatment 1411
terminal differentiation induction 429, 441, 1408–9
tumour necrosis factor α (TNF-α) combined treatment 1412
vitamin D_3 analogues 1414
vitamin D_3 receptor 1411
vitamin E 396, 672–3, 1432, 1434
antigenotoxic properties 424, 428, 437–8, 441
immune system responses 441
vitamins
antigenotoxic activities 423–4, 428
deficiency, oral cancer pathogenesis 492–3
Vitaxin 1468–9, 1477
vitronectin 210
extracellular matrix 210
degradation during malignant invasion 231
haematopoietic cell interactions 1104
integrin binding 242, 1026
VM gliomas 1127–8
vocal cords 503–4
Vogelstein's model (Vogelgram) 32–3
von Hippel–Lindau syndrome 44, 836–7, 1206
endolymphatic sac papillary tumours 847
genetic aspects 37–8, 836–7
pancreatic tumours 584
pathology 820
renal cell carcinoma 649–50
retinal capillary haemangiomas 831
VHL mutations 30, 820, 1470
von Recklinghausen disease *see* neurofibromatosis type 1
von Willebrand factor 242
angiosarcoma 575–6
voxels 1156
VP-16 *see* etoposide
vulva 617
vulval cancer
clonality analysis 251–2
human papillomavirus (HPV) 337
vulval intraepithelial carcinoma (VIN) 251–2
VZV thymidine kinase–araM prodrug activation system 1539

W
WAF1 mutations 1123–4
WAGR syndrome 647
Waldenstrom macroglobulinaemia 1382–3
Waldeyer's ring 503
non-Hodgkin's lymphoma 506–7
warm room 878
warts 336–7
water bath 878
Watson syndrome 38
WD repeats 182–3
WDHA syndrome 586
weasle (DR3; tweak) 123–6
Weibel–Palade bodies 729, 1010–11
Werner syndrome 87–88, 117
cellular genomic instability 102
clinical features 101
RecQ-family helicase defect 100–3
see also WRN
Whipple procedure 594
white blood cell–DNA adducts 378–9
WHO lymphoma classification 694–5
WIF-1 (Wnt-inhibitory factor-1) 199
Wilms' tumour *see* nephroblastoma

wingless 196
Wnt genes 196–8
Wnt inhibitors 199
Wnt proteins 196–8
  Frizzled receptor interactions 198
Wnt receptors 198–9
Wnt signalling 195–207
  aberrant activation in cancers 195–6, 550, 556
  additional regulators 204
  β-catenin
    intracellular level regulation 199–203
    nuclear gene expression regulation 203–4
    transcription activation mechanism 204
  cell surface receptors 198–9
  Drosophila melanogaster 196
  functions in embryonic development 195–6
  oncogenic activation of pathway 204–7
  target genes 203–4
  TGF-β signalling cross-talk 185, 204
  transduction pathway 196–204
Wnt-1 196, 1093
Wolfring glands 826–7
wood dust exposure 418, 505
Working Formulation lymphoma classification 693–4
workplace surveillance/monitoring 418–19
WRN 100–1
  DNA repair 102–3
  genomic instability 102
  Ku interaction 102
WRN 101
WT-1 30, 1525
  EWS fusion 759
  mutations
    central nervous system tumours 1123–4
    Denys–Drash syndrome 647
    WAGR syndrome 647
    Wilms tumour (nephroblastoma) 30, 40–1
WT-2 mutations, Beckwith–Weidemann syndrome 647
Wyburn–Mason syndrome 836–7

X
X chromosome inactivation
  clonality assessment 606, 247–9
  patch size considerations 248–9
  skewed 248
X-linked lymphoproliferative disorder (Duncan's syndrome) 59–60
X-rays 1359–60
  brain tumour risk 794
  DNA damage 89, 310
  linear energy transfer (LET) 307
xenograft studies 914
  applications 918–20
  breast cancer models 1094
  experimental agent assessment 918–20
    dose selection 919
    excision assays 919
    in situ assays 919
    survival estimates 919–20
    toxicology 919
    tumour growth delay estimates 919
  experimental design 914
    choice of model 915
    data analysis 914–15
  gastrointestinal tract tumour models 1091
  graft establishment 914
  human brain tumour cells 1127
  immune-deficient rodent recipients 915–18
  in vivo metastasis models 1003, 1005
  inoculation sites 918
  limitations 920–1
  lung cancer models 1092
  melanoma models 1116–17
  metastasis models 918, 920, 992–3
  number of animals required 914–15
  nutritional studies 920
  osteosarcoma models 1118
  prostate cancer models 1096
  squamous cell carcinoma models 1088
  statistical power 914–15
  tumorigenesis assay 918
Xenopus, Wnt target genes 203–4
xeroderma pigmentosum 87–88, 117, 299, 382, 1508
  DNA polymerase mutations 103
  skin cancer predisposition 478, 481–2
xeroderma pigmentosum complementation group proteins (XPA to XPG) 91, 99–100
  nucleotide excision repair 91
  transcription-coupled repair (TCR) 91
    Cockayne syndrome cells 93
XIAP 122, 131
ximab (chimaeric antibodies) 1380
Xnr-3 203–4
XP proteins *see* xeroderma pigmentosum complementation group proteins (XPA to XPG)
XPA-deficient mouse 300, 368–9
XRCC2 99
XRCC3 99
XRCC4 1361–2
XTT assay 1044
  method 1044
XY females 44

Y
yeast artificial chromsomes (YACs) 892–3
yolk sac tumour
  central nervous system 813
  immunohistochemistry 683
  ovary 640
  testis 682

Z
Zap 172–3
ZD0473 1303
ZD1839 (Iressa) 1497, 1500
ZD4190 1468
Zeiss glands 826–7
zinc 275, 424, 528
Zollinger–Ellison syndrome 40, 586–7, 609
ZSG/EWS fusion 738–9
zumab (humanized antibodies) 1380
zyxin 985